Pattern Recognition in Intracranial EEG Signals and Prediction of Memory

Dissertation

zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt
von

Marlene Derner
aus
Berlin

Bonn, 2020

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn.

Promotionskommission:

- Erstgutachter: Prof. Dr. Christian Bauckhage
- Zweitgutachter: Prof. Dr. Stefan Wrobel
- Fachnahes Mitglied: Prof. Dr. Thomas Schultz
- Fachfremdes Mitglied: Priv.-Doz. Dr. med. Juergen Fell

Tag der Promotion: 02.12.2020

Erscheinungsjahr: 2021

Danksagung

An dieser Stelle möchte ich mich herzlich bei allen bedanken, die mich bei der Anfertigung dieser Arbeit unterstützt und motiviert haben.
Danke an Prof. Christian Bauckhage, für die angenehme und unkomplizierte Betreuung sowie den konstruktiven Austausch und Danke an Prof. Stefan Wrobel für die Übernahme der Zweitbetreuung.
PD Dr. Juergen Fell danke ich ganz besonders für die Möglichkeit, Teil seiner Arbeitsgruppe zu sein und an spannenden Projekten mitarbeiten zu können sowie für die wissenschaftliche Unterstützung. Besonders die wertvollen Ratschläge, inspirierenden Diskussionen und die stets offene Bürotür für Probleme jeglicher Art waren enorm hilfreich und motivierend.
Zudem danke ich meiner gesamten Arbeitsgruppe für die vielfältige fachliche Diskussion und außerfachliche Unterhaltung. Danke Leila Chaieb, Jan Baumann, Amirhossein Jahanbekam, Yvonne Höller, Roy Cox, Marcin Leszczynski und allen anderen Kollegen für die immer lockere und herzliche Arbeitsatmosphäre und euren ansteckenden Enthusiasmus.
Für ihre Unterstützung danke ich außerdem ganz herzlich meiner Familie. Meinen Eltern, die mich während meiner gesamten Ausbildung tatkräftig unterstützt und ermutigt haben. Besonders möchte ich mich an ihrem Interesse an meiner Arbeit bedanken. Meinem lieben Mann Matthias für sein Verständnis und seine Geduld sowie die moralische Unterstützung und meinem kleinen Johann, ihr seid stets an meiner Seite und das Größte für mich.
Abschließend möchte ich nochmals allen Danke für die Hilfe beim Erstellen meiner Promotionsschrift sagen, Danke Sabine Landvogt, Steffen Höhne, Matthias und natürlich Juergen für das Korrekturlesen und Danke Leila für die sprachliche Hilfe!

Abstract

This thesis presents a method for pattern recognition in neurocognitive data, in particular intracranial electroencephalographic (iEEG) data. The approach aims to reveal mechanisms underlying cognitive processes. This means that the algorithm has not only been designed to achieve above chance prediction results, but also to offer a better understanding and new insights into the functionality of the brain.

A support vector machine algorithm has been developed which deals with the complex data structure, the high number of potential datapoints and features, as well as the typically small and unbalanced sample sizes. In particular, the periodicity of phase measures is taken into account for statistics and training of the model. Background information about cognitive processes provides an informative basis for feature selection. Time series analysis is used for feature extraction and circular statistics cope with the periodic characteristics of the data entering feature preselection.

Then the algorithm was applied to iEEG data recorded in presurgical epilepsy patients during a continuous word recognition task. In two studies, memory formation was successfully predicted based on iEEG measures from rhinal cortex and hippocampus, two memory-related brain regions. Different iEEG measures (i.e. absolute phases, phase shifts and power values) were compared for their predictive capabilities. The results obtained by training the algorithm with iEEG data reveal the superiority of absolute phases compared to the other measures and their importance for memory processes. Hence, the presented method is able to provide valuable insights into basic mechanisms of brain functions.

Finally, a further application comprising memory enhancement methods is presented. Here, the results of the previous application are turned into assumptions for further research. In particular, 5 Hz auditory beat stimulation was applied during an associative memory task. It was shown that phase locking was increased and that memory performance was altered depending on absolute phase values. These findings confirmed the importance of oscillatory phases for memory formation and the informative value of the previous outcomes. Taken together, the presented algorithm is able to expose key information patterns derived from neurocognitive data and might be used in memory enhancement applications.

Contents

1 Introduction **1**

2 Related Methods **7**

2.1 Regression 7

2.2 Big Data and Deep Learning 9

2.3 Support Vector Machine 11

2.4 Summary 14

3 Theoretical Background **17**

3.1 Neurocognitive Framework 17

3.1.1 The Human Brain 17

3.1.2 EEG Signals 18

3.1.3 EEG Oscillatory Phases and Memory 18

3.2 Time Series Analysis 19

3.2.1 EEG Recording 20

3.2.2 Extraction of Phase and Power Values 21

3.3 Circular Statistics 25

3.3.1 Mean Resultant Vector and Mean Direction 26

3.3.2 Rayleigh Test 26

3.3.3 Fisher's Method 27

3.3.4 Circular Kruskal-Wallis Test 27

3.4 Support Vector Machine 27

3.4.1 Cross-Validation 29

3.5 Summary 30

4 Pattern Recognition in EEG Data **31**
4.1 Artefact Rejection . . . 33
4.2 Feature Selection . . . 34
4.3 Classification . . . 35
4.4 Optimization . . . 36
4.4.1 Slack Variable . . . 36
4.4.2 Sample Size . . . 37
4.4.3 Time Windows . . . 39
4.4.4 Generalized Linear Models . . . 40
4.5 Summary . . . 41

5 Prediction of Successful Memory Encoding Based on Single-Trial Rhinal and Hippocampal Phase Information **43**
5.1 Material and Methods . . . 44
5.2 Results . . . 48
5.3 Summary . . . 59

6 Prediction of Memory Formation Based on Stimulus-Related Phase Shifts **63**
6.1 Material and Methods . . . 63
6.2 Results . . . 65
6.3 Control Analyses . . . 69
6.4 Summary . . . 74

7 Auditory Beat Stimulation to Control Phase Dynamics and Modulate Memory Performance **75**
7.1 Introduction . . . 76
7.2 Material and Methods . . . 78
7.3 Results . . . 84
7.4 Summary . . . 88

8 Conclusion **95**

Bibliography **99**

List of Figures

3.1 Phase and amplitude from a harmonic oscillation on the limit cycle. . 22

4.1 Workflow of the proposed method. 32
4.2 Adjustment of sample sizes 39
4.3 Merged time intervals 41

5.1 Paradigm and Electrode implantation site 46
5.2 Results of the Rayleigh tests. 50
5.3 Circular Kruskal-Wallis test 52
5.4 Mean phase differences 52
5.5 Examplary phase values 54
5.6 Individual classification accuracies 55
5.7 Individual classification accuracies for prestimulus intervals 55

6.1 Chosen frequencies and time points for phase shifts 66
6.2 Prediction accuracies 68
6.3 Distribution of phase shifts 70
6.4 Signal-to-noise-ratios 72
6.5 Frequency-resolved average prediction accuracies 73

7.1 Associative learning paradigm and beat stimulation. 79
7.2 Behavioral results. 85
7.3 Phase-locking results. 87
7.4 Phase distributions for the different beat stimulation conditions. . . . 89

List of Tables

5.1 Chosen frequencies and time points 57
5.2 Chosen frequencies and time points in prestimulus range 58

6.1 Chosen frequencies and time points in phase shifts 67

1 | Introduction

Pattern recognition is the automated discovery of regularities and pattern in data through computer algorithms. Prediction algorithms enable model building to predict future outcomes. Together they are used in a wide variety of research areas, and have many applications including the fields of medicine and neuroscience. In particular, these promising approaches enable researchers to potentially solve complex questions by investigating large quantities of data or datasets with special structures. When applying machine learning methods to neurocognitive data, particular challenges (i.e. small sample sizes, complex data structures, high-dimensionality) have to be handled in order to develop sufficient methods for understanding how the human brain and body are working or how they respond to a disease and its intervention (clinical or therapeutic), in order to optimize medical treatment.

Neurosciences focus on research of the nervous system. Along with the structural analysis of the brain and investigation of neural diseases like dementia, Alzheimer's and Parkinson's disease, underlying functions and basic mechanisms of cognitive processes are studied. Uncovering how the brain works (e.g. regarding consciousness or memory formation) is one of the key goals of the neurosciences (see e.g. Kandel et al., 2012). General knowledge about the brain's different regions and their main functions can be used as starting points for further and more detailed analyses. To image brain activity, there are different commonly used methods. Functional magnetic resonance imaging (fMRI) visualizes active areas indirectly via changes in blood flow and offers a high spatial resolution but at the expense of temporal resolution. On the contrary, electroencephalography (EEG) measures electric brain activity directly with a high temporal resolution but at the cost of spatial resolution. There are non-invasive EEG recorded from the scalp surface and intracranial EEG (iEEG) measured directly within brain structures. As cognitive processes take place

within milliseconds, EEG, especially intracranial EEG recordings, offer useful data to explore underlying mechanisms of brain operations (e.g. Hämäläinen et al., 1993; Niedermeyer and da Silva, 2004). Although iEEG offers the most precise spatial information (compared to surface EEG), recordings cannot easily be acquired because it is an invasive imaging method and patient groups are often rather small. Furthermore, also the electrode coverage is limited in iEEG and might not be suitable for some studies. EEG features like event related potentials (ERPs), signal power and oscillatory phases are known to play an important role in neural processing (e.g. Bruns, 2004). Studying these characteristics is thus essential to get to the bottom of cognitive functions. Particularly phase information as a circular unit of measurement requires careful handling during analysis procedures to account for its periodic structure. Hence, applying statistics, especially pattern recognition algorithms, to neurocognitive data pose a challenge, not only because of the often quite small size of datasets, but also because of the data structure itself.

A classical analysis approach is statistical modelling using regression. Regression analysis fits a model function by minimizing the errors between observations and estimations. The function (e.g. linear, logistic) is defined by the model design. After parameter estimation, significance of regressors can be determined and new values can be prognosticated (see e.g. Steyerberg, 2009). Although until now, regression models are the most commonly used models for clinical purpose, they are not the method of choice in the presented approach. Apart from commonly using the whole dataset, this algorithm is limited by the choice of the model function and the required statistical assumptions (e.g. linear relationship between dependent variable and regressors and homoscedasticity for linear regression). Therefore, it is not flexible enough for complex data structures that may appear in neurocognitive data.

Current classification algorithms profit from growing speed and storage capacity of modern computing systems. During recent years, data science has become a trending topic using big data and neural networks or deep learning. It can be applied in all kind of research areas as well as cross-disciplinary and also finds its application in everyday life (e.g. predicting price developments in stock market, improving Google search results and personalized advertisements). The growth of capacity of data storage and computational power also had a huge impact on natural sciences. Computational neuroscience forms a research area connecting cognitive

neuroscience with computer science for example by combining image processing and machine learning techniques. The special characteristics of deep neural networks is that they can automatically learn from raw data without further coded rules by simulating the way biological neurons make connections (see e.g. Schmidhuber, 2015). Still, this method is not suited for the presented analysis approach, because deep learning requires large datasets that are often hard to gather when working with patient data. Aside from this restriction, interpreting training results with regard to basic mechanisms is problematic because thousands of weights can be used to reach a solution without showing what is going on exactly inside the neural network.

Alongside other statistics and classification methods like decision trees and nearest neighbour classification, support vector machine (SVM) is another algorithm that became possible and important with growing computational power. A SVM classifier separates groups by transforming the data to a higher dimension to make it linearly separable. It has the ability to use complex, high-dimensional training data and generalize its results to unseen samples (see e.g. Press et al., 2007). Therefore it is a powerful tool that offers the opportunity to be used in a wide range of biomedical applications. It provides the possibility to use small datasets which makes it suitable for training with neurocognitive data. In recent research, SVM is mainly applied using many features, combining behavioural and neurological measures. But using a high number of features to classify small datasets bears the risk of overfitting[1] and makes the interpretation of underlying mechanisms challenging. Prediction based on neuroscience data can target different objectives. In clinical research, many studies aim to develop methods that are able to predict certain outcomes or differentiate between specific classes (see Chapter 2 for related work). Whether for personalized approaches for optimal patient treatment or for mind reading applications, the result is more important than the underlying mechanism. This might likewise be the case when classification algorithms offer the possibility to prevent violent behaviour with aggression therapy or provide additional support for children with estimated low reading abilities. Hence, to achieve these aims, it can be very helpful to combine many different features including behavioural measures and diverse neuroimaging data arising from fMRI and EEG recordings. Also, the number of features can be

[1] When a model is fitted too specifically to the present dataset (for example same set for training and testing), it might detect characteristics that do not generalize to new data

quite high to improve classification performance, while a reduced number makes it easier to interpret the importance of single features. With regard to practical applications, it is important to keep an eye on the cost-benefit ratio and consider if neurocognitive measures add information that is worth the extra costs (i.e. for expensive MRI data probes).

The approach presented in this thesis aims to develop an algorithm that provides as much biological information as possible in neurocognitive datasets. By doing so, it intends to offer new insights into the functionality of the brain. This means for the development of the method that a lower prediction performance is acceptable in favour of a better understanding of the underlying mechanism. Therefore a minimal number of features is preferable. Regression analysis which is still the most widespread method in clinical studies is not suited for this purpose because it is not flexible enough to handle complex data structures. Additionally, in existing studies, models are often fitted using the complete dataset and internal validation (like bootstrap) is used (see Chapter 2.1). It is important to keep that fact in mind with regard to overfitting and generalization. The training (and validation) data should be clearly separated from the test dataset for model building to get reliable validation results. Algorithms like SVM and deep learning are able to classify more complex datasets. Deep neural networks have become increasingly popular in maybe all interdisciplinary fields and profit a lot from growing storage capacities and computational power. They are very good in automatic pattern recognition to make predictions or cluster data and are mostly able to generalize to new data (e.g. Schmidhuber, 2015). But deep learning is not applicable to the neurocognitive datasets that shall be examined with the method developed in the following. This is because of several reasons. Deep learning typically requires large datasets. Especially when data is collected from patients, it might be difficult to recruit a large sample due to a low number of potential participants. When thinking about using data collected in different studies or over a longer time span in different hospitals, apart from finding a congruent subject base it is often problematic to access this data because of ethical reasons and data protection regulations (Boyd and Crawford, 2012). Additionally, the way deep learning and neural networks handle data is not suited for the presented approach. This machine learning technique combines different features or even parts of features to new data to predict and cluster the

original data. In many cases this can be very effective resulting in high prediction accuracies. But the interpretability of underlying mechanisms is really complicated, not only caused by the number of features, but the algorithm itself. It is hard to understand how the data was processed from the algorithm and which conclusions regarding basic principles can be drawn from the results. Hence, the chosen algorithm for further development in this thesis is a SVM. In most applications it is also trained using a lot of features to get best prediction results. But it is possible to use small datasets and when using a low number of different features, interpretation is easily practicable.

While regression analysis does not offer the necessary flexibility to handle complex data and datasets are mostly too small to result in reasonable models for deep learning, SVMs have to manage several challenges as well when working with neurocognitive data. Typically, EEG data is recorded with a high sampling rate and together with all neurologically relevant frequencies its high-dimensionality offers many possible features (cf. Hämäläinen et al., 1993). At the same time, datasets are typically quite small, because they are recorded in patient groups and gathering big data fails at data congruency and ethical reasons or data protection regulations (cf. Boyd and Crawford, 2012). To avoid overfitting, features have to be selected carefully. Additionally, it is important to account for unbalanced data for training and testing the algorithm as well as for special data structures (i.e. circular data). The method that is developed in this thesis will reduce computational time and the number of features to choose from by combining datapoints, use circular statistics for feature preselection, and test how to cope best with unbalanced class sizes. Moreover, the circular data structure is respected when training the SVM.

In this thesis, machine learning ideas in neural data domains are developed, machine learning questions are posed on neural data and it is dealt with its computational aspects. First, some background about different machine learning methods and their application in medicine and neuroscience as well as their interpretation concerning basic mechanisms is provided (Chapters 2 and 3). Then, a method for EEG data analysis is developed, which handles the circularity of phase measures, the challenge of small datasets and aims to reveal as much biological background as possible (Chapter 4). Finally, the presented method is applied to real brain data (Chapters 5 and 6) and a conclusion and prospect of future development are given

(Chapters 7 and 8).

The content of this dissertation has been published in large part in these three articles:

- Prediction of successful memory encoding based on single-trial rhinal and hippocampal phase information (Höhne et al., 2016)
- Prediction of memory formation based on absolute electroencephalographic phases in rhinal cortex and hippocampus outperforms prediction based on stimulus-related phase shifts (Derner et al., 2018b)
- Modulation of Item and Source Memory by Auditory Beat Stimulation: A Pilot Study With Intracranial EEG (Derner et al., 2018a)

2 | Related Methods

In the following chapter, different methods common to both classification and prediction are described (i.e. regression, deep learning, SVM). In addition to the individual methodological descriptions, some examples of how these techniques have been applied in recent neuroscientific research are given. Additionally, the limitations of applying these algorithms in the context of describing brain functions based on EEG data, and finally why a SVM is implemented in this thesis, are also explained.

2.1 Regression

Regression is a classical analysis approach that is commonly used for modelling in clinical research (see e.g. Steyerberg, 2009). Regression models estimate the relationship between a dependent variable y and one or more independent variables $x_1, ..., x_n$ (regressors). The model function f is fitted by minimizing the error ϵ between the observed values and the estimation described through the regressors.

$$y = f(x_1, ..., x_n) + \epsilon \tag{2.1}$$

There are many studies analysing diverse data using regression models to gain information about the association between certain behaviour and neurocognitive features. Additionally the significance of correlation values is examined in many cases. The focus of these studies is often on improving the prediction of certain behaviours, for example, to be able to individualize patients' treatment or to minimize costs to society.

There are studies that aim to answer the question of treatment response. E.g. voxelwise regression of fMRI data during the presentation of different face expres-

sions showed significant association between the pretreatment activity in amygdala and anterior cingulate cortex and treatment response in generalized anxiety disorder (Whalen et al., 2008; Nitschke et al., 2009). A whole brain regression analysis with differential fMRI responses to angry vs. neutral faces revealed that treatment response to cognitive behavioural therapy in social anxiety disorder can be predicted from this data (Doehrmann et al., 2013). For generalizability of the results, cross-validation was used to build the generalized linear model (GLM) and test the prediction based on brain activation. Its significance was assessed using permutation tests. Subject-specific contrast images were used for the analysis with a voxelwise threshold and topological correction procedure to account for false positives and limit the false discovery rate. Furthermore, amygdala activation to emotional face expressions shows a significant correlation to reduction of symptoms in depressed patients in a random-effects multiple regression model (Canli et al., 2005). And using regression analysis of fMRI responses to fearful and neutral facial expressions, a significant association between activity in the amygdala and ventral anterior cingulate and the response to behaviour therapy for post-traumatic stress disorder was detected (Bryant et al., 2008). Apart from the fact, that these studies use fMRI data that is non-invasive and offers a lot of possible features compared to (i)EEG data, regression models were often built using complete datasets, which reduces the evaluation to internal validation and limits the generalizability to new data. They focused on prediction of outcomes and did not aim to explain underlying mechanisms.

Other studies concentrate on future incidents which have effects on society. Aharoni et al. (2013, 2014) used Cox proportional-hazards regression models to study the relation between activity in the anterior cingulate cortex (ACC) during a go/no-go control task (GNG) together with other potential predictors (including age, alcohol or drug abuse and GNG error rate) and the months to a future rearrest of released criminal offenders. The discrimination accuracy (how well groups are separated) was assessed via ROC (receiver operation characteristic) curves that follow true positive and false positive fractions. They evaluated the coefficients in the model using an internal validation by resampling the Cox distribution in a bootstrapping sequence and found that only the error-related ACC response was robust to bootstrap resampling. Maurer et al. (2009) showed that ERPs in EEG data recorded in kindergarten children improve the prediction of their reading ability in primary school in a multi-

ple linear regression model. By adding neurophysiological measures to behavioural measures as predictors, the explained variance increased significantly. Furthermore, while using a logistic regression, the number of correctly classified children was significantly higher, when neurophysiological values were included. Even a correlation between fMRI activity in reward-related regions of the brain and popularity of songs (measured by the number of sales) was shown using logistic regression (Berns and Moore, 2011). In these studies neurophysiological measures improved models significantly but behavioural measures were still used for modelling which masks the importance of underlying brain activity.

The aim of the different research projects using neurological data and regression analysis is mainly to improve existing models (e.g. for advances in patients' treatment). Many of these studies report the significance of correlation values and associations between predictors and the dependent variable. Nevertheless, they often use the complete dataset and internal validation to build the model and do not aim to give specific prediction accuracies. Thereby, the risk of overfitting and the question of generalization to new datasets remain open. Additionally, regression models are fixed to preassigned model functions and statistical assumptions (e.g. linear relationship between dependent variable and regressors and homoscedasticity for linear regression). Neurocognitive data cannot always fit these requirements caused by small datasets and complex data structures. Hence, although regression models are still the most commonly used models for clinical purpose, they are not the method of choice in the presented approach.

2.2 Big Data and Deep Learning

Caused by the continuously growing capacity of data storage and computational power the collection and analysis of massive amounts of data in all sorts of domains gets easier and easier. Data science deals with these "new" records. The so called big data (i.e. datasets too large for traditional analysis methods) can be used in diverse machine learning procedures. The further development of artificial intelligence (AI) benefits from these huge datasets and the advances of high power computing systems and pattern recognition algorithms. Deep learning is a subset of AI inspired by the way the brain is structured. It uses multi-layered neural networks to recognize

patterns in representations of images, text, speech and other data. One can find its daily application for example in automated recommendations and advertising on websites or in translation apps. Deep neural networks can automatically learn from raw data without further coded rules and their predictive accuracy increases when provided with more data. Simulating the way biological neurons make connections in the brain, underlying patterns and correlations in the training data are identified. To do so, weights are assigned to the connections between the artificial neurons to define the responses to different features and data is back-propagated to check for mistakes (e.g. Schmidhuber, 2015).

For example Ramsundar et al. (2015, 2017) used deep learning on broad collections of pharmaceutical data (big data) to predict interactions between targets and small molecules for drug discovery.

For different mind reading approaches brain activity data and deep neural networks are combined. Enabling direct communication from neuronal activity to a computer (e.g. using a brain-computer interface (BCI)) might help people with disabilities or neurological diseases or for example improve image search in a fast and natural way.

By training a neural network using EEG data recorded while participants were looking at images of different object classes, Spampinato et al. (2017) showed that automated discrimination between visual categories based on brain signals is possible with an accuracy of 83 %. FMRI pictures contain general information regarding what the person is thinking about. Different mind reading approaches based on fMRI data are described from Wen et al. (2018) using deep learning to encode and decode watched movies with natural stimuli and from Shen et al. (2019a,b) who reconstructed images people were looking at from visual cortical activity using a pretrained deep neural network for the same input images. Their algorithm also generalizes to artificial shapes and alphabetical letters although trained only on natural images. Speech reconstruction using deep learning trained on invasive electrocorticography (ECoG) from the human auditory cortex recorded in presurgical epilepsy patients was demonstrated from Akbari et al. (2019). Angrick et al. (2019) used a deep neural network to reconstruct audible waveforms from ECoG captured while patients read words aloud. Anumanchipalli et al. (2019) translated neural activity into speech acoustics when sentences were read or spoken aloud but also

for silently mimed sentences via neural network. A BCI was trained using a deep neural network to classify EEG signals from imagery left and right hand movement (Zhang et al., 2019). Lawhern et al. (2018) presented a neural network to construct EEG-specific models where automatic feature extraction and classification generalize across different ERP and oscillatory-based BCI tasks. The analysed paradigms covered performance improvement through correction of incorrect BCI output (i.e. of a speller), and classification of actual or imagined movement.

Especially for clinical research, it is hard to gather large amounts of data. Groups of participants (i.e. patients) are often very small and data cannot be combined easily. Homogeneity of data is needed for comparability and reproducibility. Also, combining data from different studies or using data which might have been recorded unwittlingly or used for other than original purposes is critical because ethical guidelines have to be met and patients' privacy has to be protected. Because of this, medical datasets are being isolated in institutions and strict regulations concerning data sharing are applied. Still, releasing data could help in designing personalized drugs and research all kinds of diseases (Boyd and Crawford, 2012; Manovich, 2012).

Furthermore, deep learning does not only stick to biological backgrounds anymore in favour of brute force computing to get high performance algorithms. The brain's biology makes any interconnection between neurons within a certain distance possible. It is much more complex than any existing artificial neural network where the number of layers, connections and directions of propagations is discrete. Also, today's approaches need much more training input than the human brain to understand concepts. Nevertheless, it is still possible that research developing AI brings new insights into how the brain works. Still, interpreting training results from deep learning with regard to basic mechanisms is very hard because several thousands of weights can be used to reach a solution without showing what is going on exactly inside the neural network. Therefore, deep learning is not suited for the presented analysis approach.

2.3 Support Vector Machine

A support vector machine is another type of algorithm that became possible and important with growing computational power. SVM classifies data via a hyperplane. It

searches for a linear function (the hyperplane) that separates the data in classes with maximal free space between hyperplane and the nearest datapoints. This means it is a "large margin classifier" dividing data by a clear gap that is as wide as possible to make the classification of new data objects reliable. The datapoints determining the separating plane are called support vectors. The support vector machine uses linear classification (based on the linear hyperplane function). To be able to classify non-linear data, a kernel function transforms the data to higher dimensions where it is linearly separable. Here, the hyperplane that separates the data classes linear is found and can be transformed back to lower dimensions via the kernel function. Additionally, other parameters can be adjusted to allow for exceptions (i.e. datapoints in wrong classes). These slack variables make the algorithm more flexible. To find the best function to separate the training data, the allowed classification errors are penalized during the optimization process. Via this method, overfitting can be avoided and the number of support vectors needed to describe the hyperplane is smaller (see also Chapter 3.4).

The SVM is a powerful tool in a wide range of biomedical applications because of its ability to use complex, high-dimensional training data and generalize its results to unseen samples. It can also be used for small datasets. It offers the possibility to investigate development of mental diseases, responsiveness to medical treatment and social and cognitive functioning. Some examples for recent research in neuroscience using SVM algorithms are the following:
Soon et al. (2008) examined the predictability of motor decisions before the subject made this decision consciously and found that the outcome can be encoded based on brain activity in prefrontal and parietal cortex up to 10 s before the decision enters awareness. They applied a GLM with 26 regressors as well as a linear SVM with 10-fold cross-validation (leave-one-out) to search for predictive voxels in an unbiased fashion. They used a "searchlight" algorithm to determine which local brain pattern clusters in the fMRI data were predictive for subjects' decisions. The clusters contained voxels in a radius of 3 voxels ($\sim$ 64 voxels used as features for each cluster) with fixed regularization parameter $C = 1$. Also via a SVM and leave-one-out cross-validation, Costafreda et al. (2009) were able to identify responders to cognitive behavioural therapy based on fMRI data recorded during presentation of different sad facial expressions. And Hoeft et al. (2011) predicted future reading

gain (vs. no improvement) in dyslexia based on multivariate pattern analysis of whole brain activation patterns in fMRI and diffusion tensor imaging data (which visualizes brain structural connectivity) during phonological processing using a SVM and leave-one-out cross-validation for generalization. Behavioural measures were not predictive, but the multivariate pattern analysis in regions selected based on prior knowledge revealed right prefrontal brain mechanisms during a reading task as effective predictors. Further, recursive feature (i.e. voxel) elimination was used to find an optimal pattern with maximal performance to distinguish between the two classes. Koutsouleris et al. (2009) aimed to find whole-brain neuroanatomical abnormalities as valuable biomarkers for early detection of psychosis. They used multivariate pattern classification on structural MRI data to identify subjects at high risk of disease transition. Generalizability of prediction results was estimated via cross-validation and classification of an independent cohort.

Caused by the high-dimensionality of fMRI data, predictions can be made based on a high number of features. While this measure cannot record individual neurons but rather shows blood flow, it is still able to differentiate very small brain areas with voxel size about 1 mm. As a result, even when using many features the examined brain regions can be quite small anyway. Additional dimensionality reduction can be performed using principal component analysis (PCA). But results are less interpretable with regard to underlying basic mechanisms when using a lot of features or combination of features like from PCA. MRI maximizes spatial information at the cost of temporal resolution whereas EEG provides high temporal resolution at the loss of spatial resolution (e.g. Hämäläinen et al., 1993). Hence, EEG does not offer the same number of features from areas that are close to each other than MRI does and interpretation of underlying mechanism from predictions based on many locally different features is difficult.

There are also studies using discriminant analysis to quantify the ability of certain features to separate different classes and determine the accuracy of correct classification via these models. A stepwise discriminant analysis can be applied to identify features that enter and stay in the model. For example Molfese et al. (2001) analysed how accurately ERP measures and different behavioural features discriminate between groups of different reading ability. Fell et al. (2008) investigated which mediotemporal EEG measures (e.g. ERP, power, phase synchronization and

rhinal or hippocampal phase-locking) predict memory formation most effectively in a continuous word recognition task based on data from first word presentations, revealing that inter-trial phase-locking outperformed other EEG measures. And by applying parametric discriminant analysis, Fell et al. (2011) found rhinal alpha and hippocampal theta power as best predictors for subsequent memory among different power frequency ranges from the prestimulus domain. It has to be considered that these studies mainly include all data for modelling and thereby generalizability and reliability of classification accuracy is to be handled with care.

2.4 Summary

In clinical research, investigators aim to develop highly specialized models that are most appropriate for predicting a certain outcome or for correctly classifying data. Up until now, regression analysis remains the most widespread method in medical research fields. These studies often combine behavioural and neurocognitive data and use whole datasets to get significant regression models (e.g. Whalen et al., 2008). Furthermore, the analyses are restricted by statistical assumptions and design functions (e.g. linear relationship between dependent variable and regressors and homoscedasticity for linear regression). Hence, they are often not best suited to handle the complex data structures that can arise in the neurocognitive domain. SVM and deep learning are able to manage large datasets and many features but in most cases require a lot of data or features to make accurate predictions. The number of features, but also the deep learning algorithm itself, presents a problem when determining the basis of neurocognitive functions. There are a lot of applications where the final prediction outcome is more important than its interpretability concerning underlying mechanisms. This is the case, for example, in personalized patient treatment (e.g. Nitschke et al., 2009; Doehrmann et al., 2013) or mind reading approaches (e.g. Wen et al., 2018; Shen et al., 2019b). In these cases, training the algorithms with many different features is useful so as to achieve higher prediction accuracies. In the following sections, a method is developed that aims to uncover basic brain mechanisms based on EEG data (e.g. the importance of phase information for memory processes). To enable generalization and to obtain reliable results, the test and training datasets should be clearly separated so that no internal

validation occurs. In turn, a reduced number of features makes it easier to interpret the importance of single features and a lower prediction performance is acceptable in favour of a better understanding of the underlying mechanisms. Apart from the difficult interpretability, deep learning is not suitable for this purpose as it typically requires large datasets that are often not available when working with patient data. As SVM can be trained using small datasets, this algorithm has been chosen for further development in this thesis. It handles the complex structure of EEG data. Additionally, the challenge of small, unbalanced sample sizes combined with potentially high-dimensionality of the data is addressed.

3 | Theoretical Background

3.1 Neurocognitive Framework

The algorithm presented in this thesis is developed for intended use with neurocognitive data. To this end, background information concerning the basic structures of the brain, EEG acquisition and the relationship between cortical oscillations and memory can be useful for the choice of analysis methods and feature selection. These are broadly described in the following chapter.

3.1.1 The Human Brain

When analysing the human brain, general knowledge about its different regions and their main functions can be used as starting point for feature selection.

The brain basically consists of the cerebrum, the cerebellum and the brainstem. The brainstem is relevant for basic functions like the regulation of heartbeat, respiration and eating. It also contains structures for the sensory system and motor control especially of the face. The cerebellum is important for coordination of motoric functions and essential aspects of motor learning. The cerebrum has two hemispheres which in general are responsible for the opposite site of the body each. Cortex, basal ganglia and limbic system belong to the cerebrum. The basal ganglia are situated at its base and include parts of the motor system. In the cortex, the outer layer of the neural tissue, four cerebral lobes can be distinguished for each hemisphere: frontal, temporal, occipital and parietal lobe. The frontal lobe is the largest lobe including the primary motor cortex and regions for speech, decision making and short-term memory. The occipital lobe receives visual information which is connected to sensory memory. The parietal lobe is important for spatial navigation and processes

sensory information. The temporal lobe handles acoustic input and contains structures for semantic and memory. The limbic system is situated in the mediotemporal lobe in each hemisphere and its most important structures are the hippocampus and entorhinal cortex which play an essential role in memory formation (see for example Thompson, 2001).

3.1.2 EEG Signals

EEG is an imaging method (in the wider sense) that is particularly used for diagnostic purpose in epilepsy or sleep disorders as well as during narcosis monitoring. It has a lower spatial resolution (especially in surface EEG) than other imaging methods like MRI but a higher temporal resolution which is within milliseconds - the time range of cognitive processes. Additionally, it measures electric brain activity (changes in membrane potentials) directly instead of blood flow for example. EEG signal characteristics are mainly described through different frequency components and their amplitudes (which are typically higher in lower frequencies).

Several frequency bands in EEG signals are distinguished depending on their biological meaning. Alpha frequencies (7 Hz - 13 Hz) are associated with a relaxed waking state with closed eyes. During certain sleep states and mental processes beta frequencies (13 Hz - 30 Hz) are present. Frequencies in the gamma range ($>$ 30 Hz) occur during cognitive processes and might represent the formation of neuronal networks. The delta frequency band (0.5 Hz - 4 Hz) is mainly present during deep sleep in healthy adults. Theta frequencies (4 Hz - 7 Hz) can be found in healthy adults during sleep and rhythmic theta-activity is associated with mental processes like memory tasks (Walter, 2005). (The exact division can slightly vary depending on the sources).

3.1.3 EEG Oscillatory Phases and Memory

Oscillatory phases of local field potentials and electroencephalographic signals play an important role in neural processing. They can facilitate or inhibit neural activity and communication within a certain time window or processing stage (e.g. Womelsdorf et al., 2007; Fell and Axmacher, 2011) by interacting with neural membrane potentials and modulating neural excitability and discharge times (Elbert and

Rockstroh, 1987; Fröhlich and McCormick, 2010; Anastassiou et al., 2010).

Several studies investigated their impact on perceptual and cognitive operations. Prestimulus phases in the alpha frequency band of scalp EEG were found to influence perception of visual stimuli (Busch et al., 2009; Mathewson et al., 2009). Additionally, by manipulating local phases and phase differences between regions with transcranial alternating current stimulation visual and acoustic detection is changed too, suggesting a causal role of phase dynamics (Neuling et al., 2012; Helfrich et al., 2014). Phase information is also well-known to be an essential characteristic of memory operations. Several neural operations within the medial temporal lobe (especially within the rhinal cortex and hippocampus) are associated with long-term memory (e.g. Sederberg et al., 2007; Fell et al., 2008; Lopour et al., 2013; Burke et al., 2014). Stimulus-related phase resets are found in low-frequency oscillations during memory tasks (e.g. Rizzuto et al., 2003; Mormann et al., 2005; Haque et al., 2015). In memory processes, phases of hippocampal theta oscillations regulate the direction and magnitude of synaptic plasticity. In rats, it was shown that synaptic efficacy (which is linked to learning and memory) is enhanced when electrical stimulation is applied at the peak of hippocampal theta oscillations and depressed when the stimulation is implemented at the trough (Pavlides et al., 1988; Huerta and Lisman, 1993). Moreover, it was found that mediotemporal phase information reflecting the stability of phases across trials (e.g. rhinal and hippocampal inter-trial phase-locking and rhinal-hippocampal synchronisation) is superior to amplitude information or event-related potentials for differentiation between subsequently remembered and forgotten trials (Fell et al., 2008; Lopour et al., 2013). Effective stimulus processing and memory formation may also be accompanied by phase shifts, a response of neuronal structures to stimulus presentations (e.g. Achuthan and Canavier, 2009).

This summarized description of the relationship between oscillatory phases and memory can also be found in Höhne et al. (2016) and Derner et al. (2018b).

3.2 Time Series Analysis

Neurophysiological signals (in this case EEG signals) have a natural temporal ordering of datapoints with one observation or sample belonging to a single point in time at each sampling point. Methods used to analyse such time series data belong to the

mathematical domain of time series analysis. It aims to describe signals, identify changes caused by specific stimuli or circumstances, eliminate irrelevant information (e.g. filtering to remove noise) and predict future signal values based on previously available data. Commonly used methods are regression analyses for curve fitting and signal smoothing or calculation of moving averages. Auto- and cross-correlation can be applied to analyse the relation between signals or between different time points within a signal. Furthermore, analysis in the time and frequency domain is important in time series analysis, which also includes spectral analysis. Characteristics of useful approaches (i.e. Butterworth filtering and Hilbert, Fourier and continuous wavelet transform) are described in the following chapter.

3.2.1 EEG Recording

EEG data is typically recorded in different channels either with multiple surface electrodes with their locations mainly based on the 10-20-system as well as extended or reduced modifications of the 10-20-system or intracranial with implanted subdural grid electrodes or depth electrodes with several contacts (e.g. Niedermeyer and da Silva, 2004). IEEG has a higher spatial resolution than scalp EEG and can detect rapid local changes that might remain undetected in scalp EEG due to spatial averaging. IEEG is an invasive method and is in humans only used for medical purpose (for example presurgical evaluation in epilepsy patients, when the seizure onset zones cannot be precisely determined with noninvasive examinations). But often it is possible to additionally perform some research paradigms to collect data. When the continuous EEG signal is recorded, a sequence of discrete time data is generated by sampling in equidistant time intervals. The EEG voltage recorded in one channel gives the difference between the voltages of two electrodes which are the recording electrode and a certain electrode chosen as reference electrode. The clinical scalp EEG is usually recorded with a sampling rate of 256 Hz - 512 Hz but the sampling rate can be substantially higher in research applications. IEEG is often recorded at even higher sampling rates because subdural signals contain components of higher frequencies. This is necessary based on the *Nyquist-Shannon-theorem* (Shannon, 1949) which states that a continuous time-signal can be completely reconstructed from a discrete time-signal when the sampling rate is larger than two times the highest frequency contained in the signal. This means to reconstruct a signal with

a maximal frequency of f_{max} a sampling rate $> 2f_{max}$ is needed and the distance T between recorded points of the discrete-time signal must fulfil $T < 1/(2f_{max})$.

3.2.2 Extraction of Phase and Power Values

An EEG signal is composed of different components which comprise regular variations (periodic sinus oscillations of different frequencies), event-related signal changes and random noise. Because the signal is continuously shifting, data points which are temporally close to each other are more closely related than observations further apart. Important measures associated with neural activity are phases (e.g. phase synchronization, phase-locking) and amplitudes (e.g. power). These signals typically change frequency characteristics over time and time-frequency analysis becomes important. It provides complex-valued fields (i.e. two-dimensional) over time and frequency domain simultaneously representing amplitude and phase. Techniques include Hilbert transform, continuous wavelet transform and short-time Fourier transform. Traditional Fourier transform is mostly appropriate for longer signal segments and often not the first choice for short segments of neural activity. With matched time-frequency resolution, these three spectral analysis approaches are equivalent concerning phase and amplitude (Bruns, 2004).

Coordinates in the phase space can be used to describe the state of a harmonic oscillation at a specific point in time (Fig. 3.1). The phase value describes the position of a periodic oscillation definitely. Because of the periodicity of the sine or cosine oscillation the phase value between 0 and 2π recurs after each period. The EEG signal f can be described with its amplitude A and phase value ϕ at every point in time as:

$$f(t) = A(t)\, e^{i\phi(t)} = A(t)\left(\cos(\phi(t)) + i\sin(\phi(t))\right) \tag{3.1}$$

which results in

$$\tan(\phi(t)) = \frac{\sin(\phi(t))}{\cos(\phi(t))} = \frac{Im(f(t))}{Re(f(t))} \tag{3.2}$$

The phase gives the angle and the amplitude gives the corresponding absolute value (distance from origin) in these polar coordinates (Pikovsky et al., 2001).

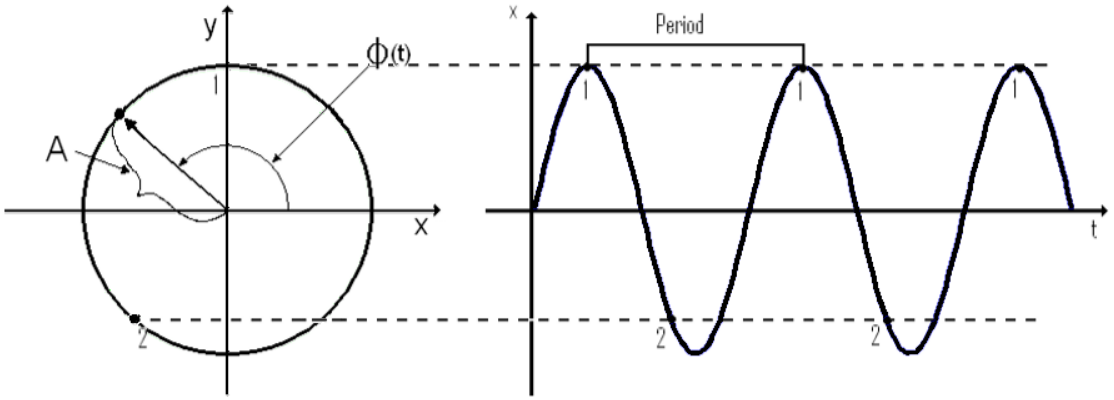

Figure 3.1: Phase $\phi(t)$ and amplitude A from a harmonic oscillation on the limit cycle.

Butterworth Filter and Hilbert Transform

For the algorithm presented in this thesis, Butterworth filtering and Hilbert transform is used to decompose the signal into neighbouring frequency components and computing the analytic signal. Continuous wavelet transform is an alternative method to get a time-frequency representation and is described at the end of this section. To extract measures from different frequencies, EEG responses can be filtered in a chosen frequency range by a second-order two-pass Butterworth filter with a certain bandwidth. The Butterworth filter is a recursive digital filter that suppresses the unwanted signal components. The filter order determines the sharpness of the filter in the transition band (i.e. an increasing filter order increases the attenuation beyond the cut-off frequency). It is given by the highest power of the transfer function in the numerator or denominator. The transfer function describes the dependency between input and output signal. With a normalized cut-off frequency of $f = 1$ the transfer function $H(s)$ for a second-order Butterworth filter is

$$H(s) = \frac{K}{s^2 + \sqrt{2}s + 1}. \tag{3.3}$$

It can be transferred to other cutoff frequencies (Meyer, 2006). Importantly, this filter preserves the phases that will be used as features in the classification. Additionally, the roll-off around the cutoff frequency is sufficiently sharp for low filter orders and without ripples. This allows a combination of acceptably small edge

effects with reasonably high filter accuracy.

To obtain the phase values the complex discrete-time analytic signal $w(t)$ is determined by the Hilbert transform of the signals. The complex Hilbert transform $s_H(t)$ is added to the real part of the original signal $s(t)$.

$$w\left(t\right) = s\left(t\right) + is_H\left(t\right) \tag{3.4}$$

This imaginary part is a version of the original signal with a shift of $\pi/2$. It is given by a convolution with the function $1/\pi t$ (Kiencke, 1998):

$$s_H\left(t\right) = s\left(t\right) * \frac{1}{\pi t} = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{s\left(\tau\right)}{t-\tau}d\tau \tag{3.5}$$

(Fast) Fourier transform (see below) can be used to determine the analytic signal. The Fourier transform has the characteristic that the Fourier transform F of the convolution of two functions is the product of the Fourier transforms of these functions (Föllinger, 2003). This means for the Hilbert transform $s_H(t)$:

$$F\left(s_H\left(t\right)\right) = F\left(s\left(t\right) * \frac{1}{\pi t}\right) = F\left(s\left(t\right)\right) \cdot F\left(\frac{1}{\pi t}\right). \tag{3.6}$$

After this calculation the inverse Fourier transform is determined to obtain the Hilbert transform (Pikovsky et al., 2001).

Based on the complex signals $w_{j,k}$, the phases $\phi_{j,k}$ for each contact and the phase differences $\Delta_{j,k}$ between two contacts ($c1$ and $c2$) are extracted for each time point j of each trial k.

$$\phi_{j,k} = \arctan\left(\frac{Im\left(w_{j,k}\right)}{Re\left(w_{j,k}\right)}\right) \tag{3.7}$$

$$\Delta_{j,k} = \phi_{j,k}\left(c_1\right) - \phi_{j,k}\left(c_2\right) \tag{3.8}$$

The phases span the range $[0, 2\pi)$ with zero representing the peak and π the trough of the oscillation. Additionally, power values $Pow_{j,k}$ can be extracted from these complex signals.

$$Pow_{j,k} = Re\left(w_{j,k}\right)^2 + Im\left(w_{j,k}\right)^2 \tag{3.9}$$

Fourier Transform

With the Fourier transform a time-dependent signal can be converted to the frequency domain. It has a perfect frequency resolution (for an unlimited signal) but no time information. The obtained spectrum describes how much each frequency contributes to the average power of the signal. It can be used for spectral analysis of time series. For a continuous function, the Fourier transform is defined through

$$\hat{f}(w) = \int_{-\infty}^{\infty} f(t)\, e^{-iwt} dt \tag{3.10}$$

And the original signal can be determined by the inverse transform

$$f(t) = \int_{-\infty}^{\infty} \hat{f}(w)\, e^{iwt} dw \tag{3.11}$$

with w the angular frequency (units are radians per second) (Whitaker, 2000).

For a discrete signal (like a recorded EEG signal with equal time steps) the discrete Fourier transform (DFT) is applied (Meyer, 2006). For a signal $s[n]$ with N sampling points, the discrete DFT is:

$$S[m] = \sum_{n=0}^{N-1} s[n]\, e^{-i\frac{2\pi mn}{N}},\ m = 0, ..., N-1 \tag{3.12}$$

and the inverse DFT:

$$s[n] = \frac{1}{N} \sum_{m=0}^{N-1} S[m]\, e^{i\frac{2\pi mn}{N}},\ n = 0, ..., N-1. \tag{3.13}$$

To reduce the computing time (from $O(N^2)$ to $O(N \cdot log_2 N)$), Fast Fourier transform is applied. It uses the symmetry of the spectrum as well as an optimized arithmetic. The discrete Fourier transform has a limited time-frequency resolution $\Delta_F = 1/T$ (with segment length T). This means that the frequency resolution is very coarse for short segments (e.g. $\Delta_F = 10$ Hz for $T = 100$ ms). A finer time-frequency resolution can be achieved, for instance, with Wavelet-based methods.

Continuous Wavelet Transform

Continuous wavelet transform is another method to get representations of the signal in the time-frequency domain. The signal $s(t)$ is convolved with filter kernels $w_f(t)$ called wavelets.

$$s_w(t) = s(t) * w_f(t) = \int_{-\infty}^{\infty} s(\tau) \cdot w_f(t-\tau)\, d\tau \tag{3.14}$$

The wavelets are complex-valued oscillations multiplied by envelopes of different shapes:

$$w_f(t) = a_f(t) \cdot e^{i2\pi ft} \tag{3.15}$$

with each representing a certain frequency band centred around f. To modify the frequency band, the chosen so called mother wavelet is varied by scale and translation parameter. They correlate with the signal if it contains a similar frequency. The obtained complex signal is equivalent to the analytic signal of the Hilbert transform. A commonly chosen mother wavelet is the Morlet wavelet with a Gaussian envelope

$$w_f(t) = e^{-t^2/2\sigma^2} \cdot e^{i2\pi ft} \tag{3.16}$$

where σ is the width of the Gaussian $\sigma = \frac{n}{2\pi f}$ and n the number of cycles. With the constant ratio f/σ the bandwidth is proportionally increasing with frequency. This might be useful to match the classical brain-signal frequency bands (Bruns, 2004). In this case also, the convolution can be calculated via the Fourier transform where the Fourier transform of the convolution is the product of the Fourier transforms of the two functions (i.e. the signal and the wavelet).

3.3 Circular Statistics

Phase values are circular values meaning that they have a periodic structure. Common statistics could lead to inaccurate or even absurd results, therefore circular statistics are used which take into account the periodicity of the data. Unit vectors and trigonometric functions are applied in many calculations and in particular, the Von Mises distribution, which is the circular analogue of the normal distribution, which underlies most circular statistical tests.

3.3.1 Mean Resultant Vector and Mean Direction

When the average of a circular dataset would be calculated as arithmetic mean, the result might be incorrect (for example $2\pi - a$ and a has a true mean of 0 but the arithmetic mean is π, which is exactly the opposite direction). The mean direction can be calculated by transforming the angle values θ_i, $i = 1, ..., N$ in unit vectors $r_i = \begin{pmatrix} cos\,(\theta_i) \\ sin\,(\theta_i) \end{pmatrix}$ and averaging the transformed vectors $\bar{r} = \frac{1}{N}\sum_{i=1}^{N} r_i = \begin{pmatrix} a \\ b \end{pmatrix}$. The mean direction $\bar{\theta}$ of the dataset is then calculated as the direction of the mean resultant vector $\bar{r}$:

$$\bar{\theta} = \begin{cases} \arctan(b/a), & a > 0, \\ \arctan(b/a) + \pi, & b \geq 0, a < 0, \\ \arctan(b/a) - \pi, & b < 0, a < 0, \\ \pi/2, & b > 0, a = 0, \\ -\pi/2, & b < 0, a = 0, \\ undefined, & b = 0, a = 0. \end{cases} \tag{3.17}$$

The length of the mean resultant vector lies between 0 and 1 and provides a measure for the directional concentration of the data. The circular variance is calculated as $S = 1 - \bar{R}$ (Berens, 2009).

3.3.2 Rayleigh Test

A significant Rayleigh test indicates that phases are not uniformly distributed but exhibit significant phase accumulations and is especially powerful in identifying unimodal differences from a uniform distribution (Berens, 2009). The p-value of a Rayleigh test is calculated based on the length of the mean resultant vector $\bar{R}$ of the data points x_i, $i = 1, ..., n$.

$$\bar{R} = \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}\cos\left(x_i\right)\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n}\sin\left(x_i\right)\right)^2} \tag{3.18}$$

$$p = \exp\left(\sqrt{1 + 4n + 4\left(n^2 - \left(\bar{R}n\right)^2\right)} - (1 + 2n)\right) \tag{3.19}$$

3.3.3 Fisher's Method

Fisher's method is a statistical procedure testing a hypothesis for a collective based on the results of independent statistical tests for the individuals of the collective (Neuhäuser, 2011). It utilizes the uniform distribution of each p-value in $[0, 1]$ under the null hypothesis. This means that $-2 \ln p_j$ is χ^2 distributed with 2 degrees of freedom and consequently $-2 \sum_{j=1}^{k} \ln p_j$ is χ^2 distributed with $2k$ degrees of freedom. With $\chi^2_{2k,1-\alpha}$ indicating the $(1-\alpha)$-quantile of the χ^2 distribution with $2k$ degrees of freedom, Fisher's method gives a significant result for:

$$-2 \sum_{j=1}^{k} \ln p_j \geq \chi^2_{2k,1-\alpha} \Leftrightarrow \prod_{j=1}^{k} p_j \leq \exp\left(-\frac{\chi^2_{2k,1-\alpha}}{2}\right) \tag{3.20}$$

3.3.4 Circular Kruskal-Wallis Test

Significant differences of phase values and inter-electrode phase differences between trials of different conditions can be determined based on a non-parametric multi-sample test for equal circular medians similar to a Kruskal-Wallis test for linear data (Berens, 2009). In this test procedure, the overall median θ is calculated based on all N observations in the combined sample. Then the number of negative differences m_i in $(-\pi, \pi]$, $i = 1, ..., k$, in every sample i as well as the total number of negative values in the combined sample $M = m_1 + ... + m_k$ are determined. The test statistic P is compared to the $1 - \alpha$ quantile of the χ^2_{k-1} distribution and calculated as

$$P = \frac{N^2}{M(N-M)} \sum_{j=1}^{k} \frac{m_j^2}{n_j} - \frac{NM}{N-M}. \tag{3.21}$$

3.4 Support Vector Machine

An SVM algorithm will be developed in this thesis. The SVM separates a dataset $(\vec{x_i}, y_i)$, $i = 1, ..., n$ in two classes via a hyperplane with $\vec{x_i}$ being a k-dimensional feature vector and y_i indicating the corresponding class -1 or 1. When trained, the sign of the function $f(\vec{x})$ should predict the value of y. For linearly separable data in n dimensions there exists an $n-1$ dimensional hyperplane which can be

described as

$$f(\vec{x}) = \vec{w} \cdot \vec{x} + b = 0 \tag{3.22}$$

with $\vec{w}$ a normal vector to the hyperplane and b an offset. Then the decision rule will be

$$y = \begin{cases} -1, & \text{for } f(\vec{x}) < 0 \\ 1, & \text{for } f(\vec{x}) > 0. \end{cases} \tag{3.23}$$

Additionally, the hyperplane with the largest margin will be determined by maximizing the distance to the points nearest to the hyperplane. It is possible to find two parallel bounding hyperplanes that separate the data with a distance as large as possible and fulfil

$$y_i(\vec{w} \cdot \vec{x_i} + b) \geq 1 \tag{3.24}$$

For this so called fat plane structure the distance between the bounding hyperplanes is twice the margin

$$2 \times \text{margin} = 2(\vec{w} \cdot \vec{w})^{-1/2} \tag{3.25}$$

To find the maximum-margin hyperplane, which is the hyperplane that lies halfway between the bounding hyperplanes, $\|\vec{w}\|$ has to be minimized subject to (3.24), $i = 1, ..., n$. Some of the data points lie exactly on one of the bounding hyperplanes ($f(\vec{x}) = \pm 1$) and are the so called support vectors of the solution.

To get rid of the strict assumption of linearly separable data the kernel trick is used. A function φ maps the k-dimensional feature vectors into an embedding space of a (much) higher dimension N. The idea is that the nonlinearly separable data can be separated via a linear hyperplane in the higher N-dimensional space. It is not necessary to know the mapping $\varphi(\vec{x})$ but use the kernel function $K_{ij} = K(\vec{x_i}, \vec{x_j}) = \varphi(\vec{x_i}) \cdot \varphi(\vec{x_j})$. This results in

$$f(\vec{x}) = \vec{w} \cdot \varphi(\vec{x}) + b = \sum_i \alpha_i y_i \varphi(\vec{x_i}) \cdot \varphi(\vec{x}) + b = \sum_i \alpha_i y_i K(\vec{x_i}, \vec{x}) + b \tag{3.26}$$

The kernel function must be symmetric and have nonnegative eigenvalues. Additionally, any multinomial combination of kernel functions is a kernel function. For a given kernel function $K(,)$, $K(\varphi(\vec{x_i}), \varphi(\vec{x_j}))$ is one for any φ. And $K(\vec{x_i}, \vec{x_j}) = g(\vec{x_i})g(\vec{x_j})$

is always a kernel for any function g. A linear kernel function is given as $K(\vec{x_i}, \vec{x_j}) = \vec{x_i} \cdot \vec{x_j}$.

To allow for data points that are not correctly classified by the hyperplane, a slack variable ξ with $\xi = 0$ for correctly and $\xi > 0$ for wrongly separated data is added. Based on a chosen regularization parameter λ in the range $0 < \lambda < \infty$, which varies accuracy and possible robustness of solutions, the optimization problem changes to:

$$\begin{aligned} \text{minimize: } & \frac{1}{2}\vec{w} \cdot \vec{w} + \lambda \sum_i \xi_i \\ \text{subject to: } & \xi_i \geq 0, \\ & y_i(\vec{w} \cdot \vec{x_i} + b) \geq 1 - \xi_i, \text{ i=1,...,n} \end{aligned} \tag{3.27}$$

This slack variable allows there to be some solution, also for not linearly separable data (Press et al., 2007).

3.4.1 Cross-Validation

When using cross-validation, the classification algorithm is trained iteratively with subsets of the complete data set to avoid overfitting and allow for generalization to new data. To do so, the data is split into a chosen number of equal folds (e.g. five folds for five-fold cross-validation). Then, one subset each is hold out as test or validation set and the remaining (e.g. four) folds are used to fit the model. The folds can even be chosen in an extreme way, where all data except one observation is used for training (leave-one-out). Prediction performance for each model is determined using the hold-out observations. Model fitting and validation is repeated for each fold. The overall accuracy is then estimated as the average performance over all accuracies. This means, all data is used for training and testing, but each time the model is tested with data that was not included in training. This way, the optimistic bias arising from overfitting (when the same data is used to train and test the model) is reduced (Bishop, 2006).

3.5 Summary

The background information and theoretical foundations previously outlined in this chapter have been used in the development of the SVM algorithm for intended application to neurocognitive EEG datasets. Understanding the background of how the neurocognitive data is derived allows for the choice of measures that can be used as features, to be better informed. For instance, information regarding the function of different brain structures can indicate which EEG recording channels are of interest, and in turn the relationship between EEG oscillations and memory processes help identify EEG measures that are relevant for the specific process (e.g. power or phase values). Time series analysis is used for feature extraction and the algorithm is concerned with the periodic structure of the data by computing circular statistics. Further details of the SVM algorithm are given in the next chapter.

4 | Pattern Recognition in EEG Data

In this chapter the prediction algorithm for intended application to neuronal datasets is described. It has been specifically designed for pattern recognition in EEG, expressly that of iEEG data. The method has been roughly described in Höhne et al. (2016) and Derner et al. (2018b).

Prediction and classification algorithms can address categorization of data to any chosen number of categories. The classification method has to be developed accordingly for binary or multiclass classification. In medicine and neuroscience, there are many applications where it is interesting to differentiate between only two different states (for example predicting memory formation remembered vs. forgotten (e.g. see Chapters 5 and 6), perception of a signal near a threshold (e.g. Mathewson et al., 2009) or positive vs. negative response to a treatment (e.g. Costafreda et al., 2009)). The methods outlined in this thesis focus on analysing EEG data with a two-class classification using a support vector machine. The most recent brain models concentrate increasingly on individual subjects instead of one overall model. Here, the same algorithm is applied for each subject but separately on a subject by subject basis. The results are still generalizable, but an underlying mechanism may become more evident on a single subject basis that might be hidden otherwise (i.e. when combining several subjects) due to inter-individual differences. Based on this information, features depending on frequencies and time points are selected individually for each subject. Importantly, the circular characteristic of EEG phases is considered.

After preprocessing of the EEG data, time windows and frequencies with statistically significant phase clustering across patients are identified. Then for each pa-

tient, time periods and frequencies for which the absolute phases and inter-electrode phase differences differ between both conditions are determined. Finally, a SVM is trained using the phases and phase differences from the most significant time windows and frequencies. The workflow of the proposed method is sketched in Fig. 4.1. In particular, the algorithm is concerned with the periodic structure of the data, the high number of potential datapoints and features, as well as the typically small and unbalanced sample size. Importantly, the overall aim was to employ a minimal set of features for successful prediction, for ease of exposition, on the one hand, and as such an approach is most closely related to possible practical applications (e.g. controlling one of the features by deep brain stimulation), on the other hand.

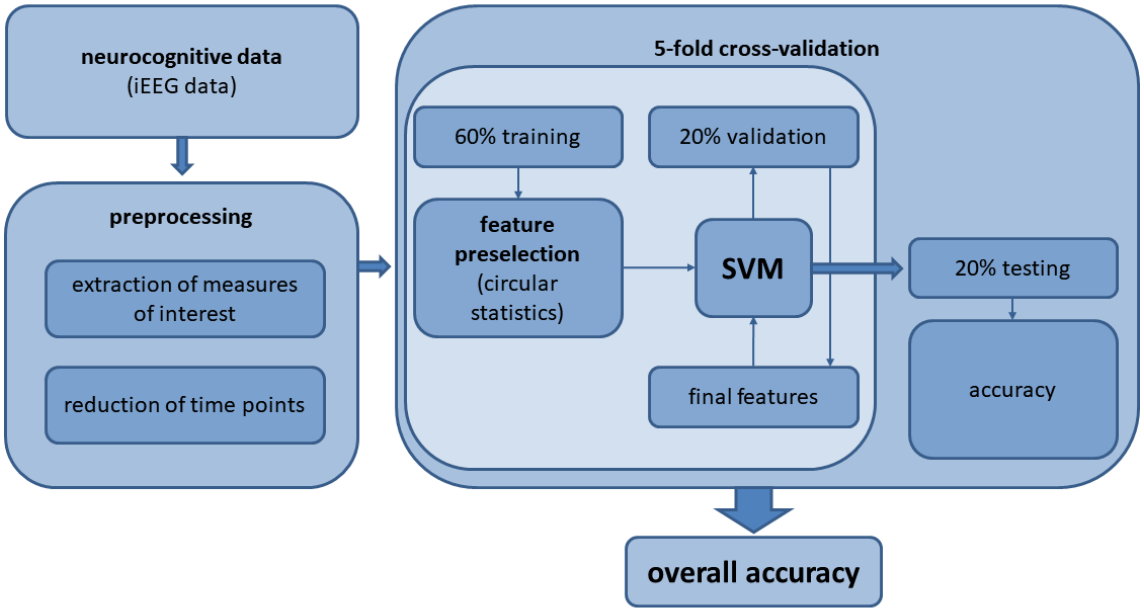

Figure 4.1: ***Workflow of the proposed method.*** *Neurocognitive data is preprocessed for extraction of measures of interest. The prediction procedure uses 5-fold cross-validation. Features are preselected in the training dataset (60 %) and the final features are determined based on prediction results of the validation data (20 %). The overall prediction accuracy is calculated as the average of prediction accuracies of each fold.*

4.1 Artefact Rejection

An EEG signal is typically filtered after recording to eliminate typical artefacts. A bandpass-filter is applied to remove artificial high or low signal components caused for example by movement or potentials generated by muscle cells. A notch filter is used to remove artefacts caused by power lines.

EEG data also often contain artefacts caused by muscle activity (blinking, chewing, and jaw muscle activity) as well as pathologic activity (e.g. epileptic spikes in iEEG data recorded in epilepsy patients). Trials that included abnormally high amplitudes as well as abrupt rises or falls should be removed by an artefact rejection. To keep signal quality as consistent as possible, automated artefact rejection algorithms should be preferred to manual artefact rejection. Here an algorithm implemented in MATLAB (MATLAB Version 8.2; The MathWorks Inc.) is chosen. Based on time intervals during which neural processes take place and keeping in mind that preprocessing might bring edge effects, trials are selected for a time window from 1000 ms before to 2500 ms after stimulus onset. The mean $\bar{x}$ (or $\bar{x_g}$) and standard deviations s (or s_g) for each segment is calculated for the data points x_i, $i = 1, ..., n$ as well as for the gradients $g_i = x_{i+1} - x_i$, $i = 1, ..., n-1$ (differences between two consecutive data points).

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{x_g} = \frac{1}{n-1}\sum_{i=1}^{n-1}(x_i - x_{i+1}) = \frac{1}{n-1}\sum_{i=1}^{n-1} g_i, \tag{4.1}$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2, \quad s_g^2 = \frac{1}{n-2}\sum_{i=1}^{n-1}(g_i - \bar{x_g})^2,$$

If one segment shows data points or gradients diverging more than five standard deviations from the mean this segment is eliminated for all contacts. So each data point x_i within a segment has to fulfil

$$|x_i - \bar{x}| < 5s \wedge |g_i - \bar{x_g}| < 5s_g \tag{4.2}$$

If data still shows artefacts (observed by visual inspection) after the automatic artefact rejection it should be considered to exclude a patient from further analysis and only keep the data from the other patients for classification.

4.2 Feature Selection

The SVM algorithm will use a five-fold cross-validation to determine the overall prediction accuracies. For feature selection, an additional validation dataset is used. For each subject, the dataset is randomly divided in a set of 60 % training, 20 % validation and 20 % test trials.

To extract measures from different frequencies, EEG responses are filtered by a second-order two-pass Butterworth filter with a bandwidth of 1 Hz. For SVM classification, frequencies from 0.5 Hz to 50 Hz in 0.5 Hz steps were considered because these frequencies cover the biological relevant frequency bands that occur in EEG signals. The complex discrete-time analytic signal is determined by the Hilbert transform of the signals to extract the phase values of each contact and phase differences between two contacts each. Additionally, power values can be extracted from these complex signals. In order to avoid edge effects, EEG responses were segmented from -1000 ms to 2500 ms with respect to stimulus onset, and after filtering and Hilbert-transform 500 ms at both sides are discarded. The free FieldTrip toolbox for MATLAB is used for filtering of the signal and extraction of phase values (Oostenveld et al., 2011). As measures for each time point and frequency are extracted in this step, the statistical tests that are described in the following can be calculated in parallel for each time-frequency point. To further reduce computing time, also filtering and Hilbert transform for each frequency can be parallelized for feature extraction beforehand.

For circular statistics the free CircStat toolbox for MATLAB is used (Berens, 2009). First to get an impression of the phase values contained in the dataset a Rayleigh test (function circ_rtest) is performed for each time window and each filter frequency separately for trials of both conditions. A significant Rayleigh test indicates that phases are not uniformly distributed but exhibit significant phase accumulations. To identify overall effects Rayleigh tests are performed for each of the k subjects individually and p-values p_j, $j = 1, ..., k$, are then combined using Fisher's method.

This step does not provide any information about potentially contained differences between conditions but about frequencies with high phase accumulations. This information could be used to reduce the frequency bands considered for feature

selection and thereby reduce the calculation time for the entire algorithm. When background information about the data is knows (for example from previous studies or common knowledge) this step might be omitted.

Frequencies and time intervals with significant differences of phase values and inter-electrode phase differences between trials of different conditions were selected for each patient based on a non-parametric multi-sample test for equal circular medians similar to a Kruskal-Wallis test for linear data (function circ_cmtest). These tests are performed on the training data set only (60 % randomly selected trials). In other words, there is no overlap between the 60 % of data used for testing for differences in median phase direction and the test data used for classification. Based on the overall result of the Rayleigh test the frequency range for classification is chosen (frequency range with most pronounced phase accumulations is chosen). For each patient and each measure the frequencies and time windows with the 10 most significant differences between conditions were preselected as features. To further reduce the number of preselected features a validation data set (20 % trials) is used and a support vector machine with a linear kernel classifying the trials into both categories is applied to this data set. Based on the highest prediction accuracies in the validation data one time-frequency point for each measure is selected as final features for classification of the test data (remaining 20 % trials) for each patient and measure. Importantly, frequency and time points are independently selected for each measure (e.g. phases, phase differences, phase shifts or power values). The workflow of the proposed method is sketched in Fig. 4.1.

4.3 Classification

One specific frequency and time point will be selected for each patient and measure based on significant differences between the measures for both conditions in the training trials.

Phase is a circular quantity, hence to avoid calculation errors and account for the periodicity of the data, the real and imaginary part $Re(\phi(t))$ and $Im(\phi(t))$ of the complex representation of the phases are entered as features to a support vector machine instead of the phase values $\phi(t)$ itself, resulting in a doubling of the number of features. (In other words, the phases and phase shifts are deconstructed into their

sine and cosine contributions.)

Because neurocognitive data (i.e. EEG data) is often unbalanced and has small trial numbers the classification procedure is performed using five-fold cross-validation with adjusted numbers of randomly chosen training trials. The data is down sampled by randomly choosing n training trials for the condition with the higher number of trials with n being the number of trials in the other condition. (Other methods for adjusting the number of training trials were also tested; see Chapter 4.4). The overall classification accuracy is reported as the averaged accuracies from these five cross-validation runs. Classification efficiency is evaluated by a non-parametric label permutation approach (Maris and Oostenveld, 2007). Group labels are randomly shuffled 1000 times and then these surrogate trials are classified again for all five-folds using the same classification procedure as for the original trials. The statistical significance of above chance classification performance (95 % threshold) is evaluated by ranking the mean accuracy of the real data within the accuracies obtained from the label shuffled data. Exactly stated, in order to conform to the 95 % threshold, the accuracy of the real data had to rank within the 50 highest accuracies of the 1000 label shuffled data. The support vector machine is applied to new trials and classification accuracies are evaluated for each patient separately.

4.4 Optimization

After deciding how the algorithm should handle the circular EEG data and which measures will be selected as features, the algorithm can be optimized. To do so, the algorithm was trained using a real dataset and prediction accuracies for different parameter choices were compared.

4.4.1 Slack Variable

When training a SVM with a linear kernel (min $1/2\vec{w}\cdot\vec{w}+\lambda\sum_i \xi_i$; see Chapter 3.4), a hyperplane that separates the data as well as possible while having the largest minimum margin is searched. The λ parameter is a regularization parameter which determines how much misclassification is allowed during optimization. A small λ gives a large margin but allows more outlier. In opposite, a large λ chooses a

smaller margin in favour of more correctly classified training points. Tuning λ is an important step. Trying to classify each training sample correctly or allowing a too large training error can reduce generalizability of the classifier to new data. λ should be chosen dependent from the data. This can be realized via a grid search where the SVM is trained with different instances of λ and the optimal parameter is chosen via validation data (additional to the final test set and distinct from the training data). In the application described in this thesis, $\lambda = 1$ was used.

4.4.2 Sample Size

Handling different group sizes is crucial when applying machine learning algorithms. Especially for small trial numbers, training results can be highly corrupted by unbalanced group sizes. For example, when differentiating between two classes with 90 % of trials in the first class and only 10 % in the other, a training algorithm would achieve an accuracy of 90 % by simply assigning all trials to the first group. Hence, for the presented algorithm, sample sizes will be adjusted before entering the SVM. Besides, it is possible to account for different group sizes in the test dataset by evaluating the classification performance for each class separately (sensitivity (true positive rate) and specificity (true negative rate) in a binary classification) and calculating the overall prediction accuracy as mean between these individual accuracies.

For adjusting group sizes, three different methods were compared using a dataset consisting of real iEEG data (for the entire analysis of this data, see Chapters 5, 6).

Resampling

The number of trials m in the smaller sample s_{small} can be increased to the number of trials n in the larger sample s_{big} by randomly selecting trials from s_{small} and adding them to this group (duplicating $n - m$ trials until the sufficient group size n is reached).

Downsampling

In a similar manner to resampling, the number of trials n in the larger sample can be reduced to the size m of the smaller sample by randomly selecting trials from s_{big}

and discarding them ($n - m$ trials) from the training dataset for classification.

Upsampling

The number of trials in the smaller sample (s_{small}) can also be augmented by adding artificially constructed trials to that group. These trials should of cause fulfil all characteristics of the original dataset and fit as good as possible in this particular group. In the approach examined for this comparison, for each new datapoint, two datapoints from s_{small} were selected in a random way. To do so, the phase value a_1 of the first datapoint was chosen randomly from the complete dataset s_{small} and the phase value a_2 of the second datapoint was chosen among the five nearest neighbours to the first point. Then, a new datapoint was inserted randomly which lies between the two original datapoints. Importantly, a random distance between the two points is chosen but the special characteristic of the circular data is maintained (i.e. the new datapoint lies on the unit circle).

When $a_1 \in [0, 2\pi)$ and $a_2 \in [0, 2\pi)$ and $a_1 < a_2$ the minimal distance between the two phase values is determined as $min(d_1, 2\pi - d_1)$ with $d_1 = a_2 - a_1$. To get a new simulated datapoint a random part of this distance $d = d_1$ or $d = d_1 - 2\pi$ respectively, is added to a_1 to get a new phase value: $a_1 + rd$, $r \in [0, 1]$. This new datapoint lies between the two original datapoints and on the unit circle. With the new phase value and the distance 1 of the unit circle, its sine and cosine attributes can also be determined.

Comparison Based on Real Data

To compare the three different methods for adjustment of group sizes, a real iEEG dataset was used. Three different features (rhinal phase, hippocampal phase, rhinal-hippocampal phase difference) were selected and entered separately into a SVM (deconstructed into their sine and cosine contributions). The samples were adjusted using resampling, downsampling or upsampling and the resulting prediction accuracies were compared.

For all three tested features, there was no notable difference between the three different adjustment methods (repeated measures ANOVA all $p > 0.7$). The overall accuracies ($\sim$ 64 %) as well as sensitivity ($\sim$ 65 %) and specificity ($\sim$ 63 %) gave

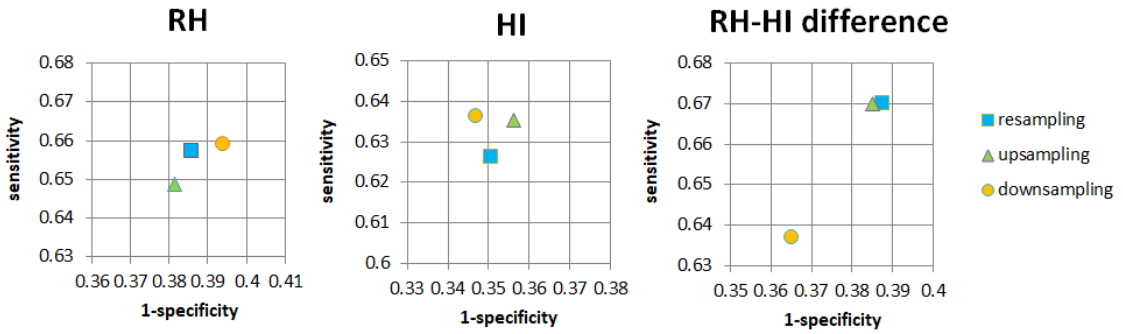

Figure 4.2: ***Adjustment of sample sizes.*** *Average prediction accuracies (sensitivity and specificity) in ROC area for three different adjustment methods. A SVM was trained with rhinal phase, hippocampal phase or rhinal hippocampal phase differences of adjusted samples each. There were no notable differences between the three methods (repeated measures ANOVA all $p > 0.7$).*

approximately equivalent results (see Fig. 4.2). This means that each method could be adopted. Upsampling is not the first choice because it adds artificial datapoints that are just assumed to fit the real dataset. Down- and resampling both select random portions of the data and provide more reliable results when the procedure is repeated several times. Because downsampling is the more classically used approach in neuroscience, this method is chosen for sample size adjustment.

4.4.3 Time Windows

Besides feature extraction for each selected frequency in parallel, the calculation effort might be optimized by reducing the number of data points used for the prediction algorithm. EEG signals are often recorded with a high sampling rate but are changing continuously with time depending on the actual state. In the following step the number of time points is reduced to minimize calculation effort and smooth the signal. For this purpose, phase and power values are averaged for non-overlapping successive time windows for each trial and frequency. EEG signals typically have a minimal sampling rate of 200 Hz. Based on this sampling rate, the prediction accuracy of the SVM algorithm was determined for averaged values from time windows with a duration of 5 ms (original resolution), 10 ms and 20 ms (which would be related to averaging across a whole 50 Hz cycle).

A SVM was trained separately for five different features of a real dataset: rhinal phase, hippocampal phase, rhinal-hippocampal phase difference, rhinal power, hippocampal power (see Chapter 5 for the complete analysis of this dataset). Prediction accuracies averaged across these five features were determined and compared. Sensitivity and specificity as well as the number of results significantly above chance level were also taken into account. (During this step, no high prediction accuracies are expected because the algorithm is trained with single features only.)

Average prediction accuracy for 5 ms time intervals was 60.0 %, 59.9 % for 10 ms and 58.7 % for 20 ms intervals. These results do not seem to show large differences but the tendency remains the same when regarding sensitivity (61.6 % for 5 ms, 61.2 % for 10 ms and 60.5 % for 20 ms) and specificity (58.7 % for 5 ms, 58.6 % for 10 ms and 56.9 % for 20 ms) separately (see Fig. 4.3). Additionally, a repeated measures ANOVA showed significant differences tested across all 27 subjects of the dataset ($F_{2,78} = 3.1993$, $p = 0.0489$; paired t-tests 5 ms vs. 10 ms $p = 0.72$, 10 ms vs. 20 ms $p = 0.09$, 5 ms vs. 20 ms $p = 0.05$). When considering the number of patients with prediction accuracies above chance level, results for the 20 ms time window (40.7 %) were clearly worse than those for the 5 ms (45.2 %) or 10 ms (45.9 %) time windows.

Based on these results averaging values in time windows with duration of 10 ms were chosen for the algorithm (250 windows in total for segments with 2500 ms length). With regard to the prediction algorithm, these values seem to contain as much information as the original data points (original resolution with 5 ms length for the analysed dataset) but reduce the calculation time for feature selection considerably (already reduced by half for a relatively low sampling rate of 200 Hz).

4.4.4 Generalized Linear Models

Since regression analyses like generalized linear models are still the most common models in clinical research, the developed method is compared to this traditional analysis method by building a linear model using a real dataset. Here, the SVM achieves clearly higher prediction accuracies than GLM (> 60 % vs. ~ 50 %). (See Chapters 5.2 and 6.2 for results reported based on SVM and 6.3 for results based on GLM).

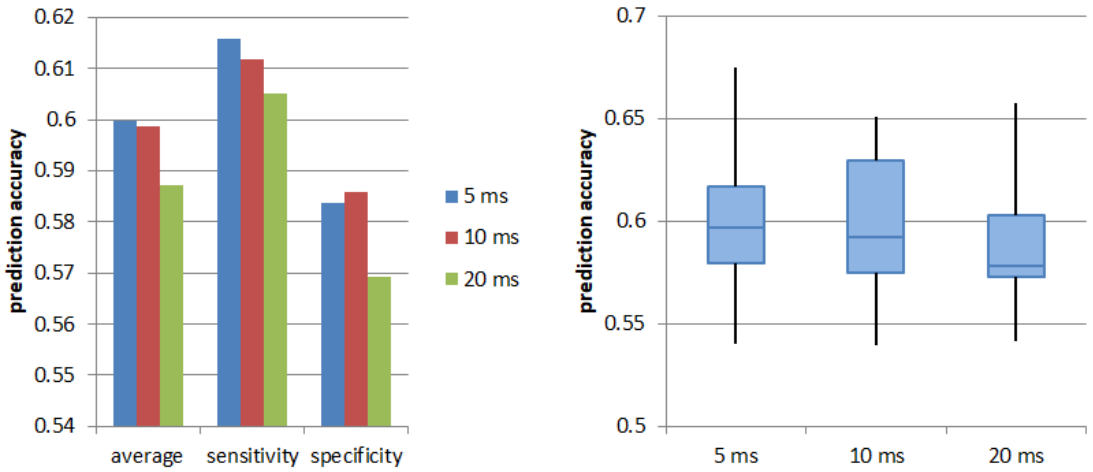

Figure 4.3: ***Merged time intervals.*** *Mean prediction accuracies across five single features (rhinal phase, hippocampal phase, rhinal-hippocampal phase difference, rhinal power, hippocampal power). Datapoints that entered the algorithm as features were averaged across time intervals of different length (5 ms, 10 ms, 20 ms). Left: mean overall accuracies as well as results for sensitivity and specificity. Right: boxplots for prediction results across all 27 subjects*

4.5 Summary

The SVM algorithm presented in this thesis is designed to cope with the complex data structure and to avoid overfitting of the typically high-dimensional EEG datasets. The features used are selected individually for each patient and feature preselection is based on tests comparing phase accumulation and phase directions derived from circular statistics (i.e. Rayleigh and Kruskal-Wallis tests) for different frequencies and time windows of averaged datapoints. The periodic structure of the data is also taken into account when the phase features enter the SVM decomposed in their sine and cosine parts, and the small and unbalanced sample size is considered through downsampling, cross-validation and repeated runs. Importantly, the algorithm is designed to reveal as much biological information as possible within the neurocognitive data and uncover basic mechanisms. Hence, in the next chapter, its application to a real iEEG dataset is presented.

5 | Prediction of Successful Memory Encoding Based on Single-Trial Rhinal and Hippocampal Phase Information

In this chapter, the previously described algorithm that is developed to analyse neurological data and investigate basic functions is applied to a real iEEG dataset. The data was recorded from the mediotemporal lobe of 27 presurgical epilepsy patients who were asked to perform a continuous word recognition paradigm. The algorithm shall examine the relation between phase values and memory formation. Phase information during the encoding trials is used as features for the support vector machine to predict subsequent remembering vs. forgetting. (The content of this chapter has been previously published in Höhne et al. (2016).)

Fell et al. (2008) investigated how closely different mediotemporal EEG measures are related to memory formation. They analysed the dataset regarding average event-related potentials, changes in rhinal and hippocampal power spectrum and measures reflecting the stability of phase values and phase differences across trials (i.e. inter-trial phase-locking and phase-synchronization between rhinal cortex and hippocampus). They found that memory formation comes along with phase regulation in the rhinal cortex and hippocampus shortly after stimulus onset and that phase-locking values showed the largest difference between subsequently remembered and forgotten trials compared to the other examined measures. Against the background of facilitation of neural processing and cellular plasticity through oscillatory phases via spike-field coupling and thereby promotion of memory formation

(e.g. Fell et al., 2011; see Chapter 3.1.3), there should be ideal and inappropriate phases for successful memory encoding. The results from the previous study reveal a stronger phase accumulation for subsequently remembered compared to forgotten trials but give no information if the phases are centred around different or the same values. Based on this information, it is an open question whether single-trial phase values per se can predict successful memory formation and trials can be correctly classified into remembered and forgotten based on selected phase values (or if the phase regulation is rather a reaction to something else like the word presentation).

First, encoding-related responses for remembered and forgotten words are analysed to explore the effect of absolute single-trial phase values as relevant predictor of subsequent memory performance. Then, it is examined if prediction based on single-trial phase values is superior to prediction based on single-trial power.

Results summary: Prediction of successful memory encoding was possible in the majority of patients (23 out of 27) based on absolute single-trial phase values with an overall classification accuracy of 69.2 % across all subjects. Importantly, only three features were chosen, one rhinal phase, one hippocampal phase and one rhinal-hippocampal phase difference. Successful prediction is also possible when features are selected from the prestimulus interval only (19 out of 27 patients with above chance prediction and an overall accuracy of 65.2 %). Prediction based on absolute single-trial phases outperforms prediction based on single-trial power. Also combining absolute phase measures with single-trial power did not increase prediction accuracies significantly.

5.1 Material and Methods

Patients

The EEG signals were recorded in 31 right-handed presurgical epilepsy patients (14 females, 16 - 61 years, mean age 40) who ranged in duration of their pharmacoresistant unilateral temporal lobe epilepsies from 4 to 57 years (mean 23 years). Based on magnetic resonance imaging scans or post-surgical histological examinations, 16 patients suffered from unilateral hippocampal sclerosis (right: 11, left: 5), 9 from unilateral extrahippocampal lesions without signs of hippocampal sclerosis (right: 6,

left: 3), 3 from unilateral hippocampal sclerosis with additional extrahippocampal lesions on the same side (right: 1, left: 2) and 3 patients had no clear lesion. All but two patients underwent subsequent epilepsy surgery after implantation. At the time of recording, all patients received anticonvulsive medication (plasma levels within the therapeutic range). Each patient performed a continuous word recognition task while recording as part of the presurgical routine and gave written informed consent to participate in the study and for the use of the data for research purposes. The study was approved by the local ethics committee of the Medical Faculty at the University of Bonn.

Experimental Paradigm

A continuous word recognition paradigm was performed (Fig.5.1 A). A total of 450 frequent German nouns (300 different words) were presented consecutively of which 150 words were displayed only once and the other were shown with one repetition. The lag between the first and the second presentation of a word varied between a short lag of 3 to 6 words in 50 % of the trials and a long lag of 10 to 30 words in the other 50 %. The word stimuli were presented in white colour on a black background and lasted 300 ms with an individually adjusted inter-stimulus interval of 1600 ms ($n = 6$), 2000 ms ($n = 16$) or 2700 ms $\pm$ 200 ms ($n = 9$) based on the subjects' abilities (evaluated based on their performance in a few pilot trials). The adjustments aimed to enable patients to respond to each word using one of two keys which they pressed with their right and left forefingers. They had to decide after each presentation if the word has been displayed before (old) or not (new). Depending on whether a word was correctly identified (i.e. correctly recognized as "old") or not (i.e. wrongly labelled as "new") at the second presentation, responses to the first presentation were classified as subsequently "remembered" or "forgotten". The EEG recordings related to the first presentation of words shown with one repetition were analysed and used for prediction applying the presented SVM algorithm. If only a small number of words were recognized correctly (bad performance with less than 30 correctly recognized "old" or "new" words) or if ERPs were defective through sharp waves or spikes, recordings were repeated with a parallel version of the continuous word recognition task on the following day. In these cases, the data of the second recordings were used for the analyses.

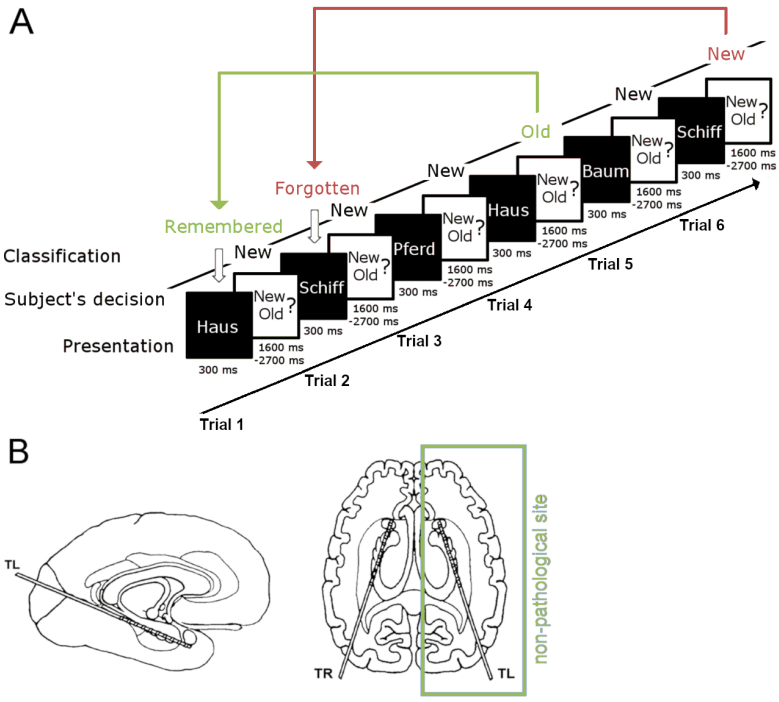

Figure 5.1: ***A: Continuous word recognition paradigm.*** *Subjects had to remember common German words and trials of the first presentation were classified as subsequently remembered or forgotten based on the response to the second presentation.* ***B: Electrode implantation site.*** *Only recordings from contacts contralateral to the ictal onset zone were included in the analysis.*

EEG Recordings

Intracranial depth electrodes with 10 platinum contacts were implanted stereotactically along the longitudinal axis of the hippocampus. Electrode contact placement was ascertained based on the individual MRIs and comparison with standardized anatomical atlases (e.g. Duvernoy, 1988). Each patient had at least one contact in the rhinal cortex and one in the hippocampus. IEEG was recorded using a sampling rate of 200 Hz, the signals were referenced to linked mastoids and bandpass-filtered from 0.01 Hz (6 dB/octave) to 70 Hz (12 dB/octave). Only recordings from contacts contralateral to the ictal onset zone were included in the analysis (Fig. 5.1 B). Then, for each patient one rhinal and one hippocampal electrode contact was chosen based on the MRI data and average ERPs. The rhinal (RH) contact was defined as anatomically located within the anterior parahippocampal gyrus with the largest mean amplitude (new words) of the negative component between 200 ms and 600 ms (N400 component, Grunwald et al. e.g. 1999). The hippocampal (HI) contact was defined as located within the hippocampus and showing the largest mean amplitude (new words) of the positive component between 300 ms and 1500 ms (P600 component, Fernández et al. e.g. 1999; Ludowig et al. e.g. 2008).

Because lateralization of verbal memory in MTL (mediotemporal lobe) 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.

Artefact Rejection

The automated artefact rejection algorithm was applied to remove the segments with abnormal high amplitudes or abrupt signal changes from the hippocampal and rhinal contacts. On average, 14 % of the trials were removed as a result of the artefact rejection. After the automatic artefact rejection, visual inspection revealed that the data from four patients still showed artefacts and therefore they were excluded from further analysis resulting in 27 patients for classification.

Extraction of Phase and Power Values

Phase and power values as well as the phase differences between rhinal cortex and hippocampus were extracted for each time point of each trial from -500 ms to 2000 ms with respect to stimulus onset using a second-order two-pass Butterworth filter with 1 Hz bandwidth in a frequency range from 0.5 Hz to 50 Hz (0.5 Hz steps) and Hilbert transformation as described in Chapter 3.2.2.

For each measure, values were averaged for non-overlapping successive time windows of 10 ms duration for each trial and frequency. For the resulting 250 time windows per frequency, Rayleigh tests were performed separately for "later remembered" and "later forgotten" trials for rhinal and hippocampal phase values as well as for rhinal-hippocampal phase differences.

Prediction of Subsequent Memory

Features to predict subsequent memory (remembered vs. forgotten trials) were given by one rhinal phase, one hippocampal phase and one rhinal-hippocampal phase differences. Features were selected as described in Chapter 4.2.

A non-parametric multi-sample test for equal circular medians for "later remembered" vs. "later forgotten" trials was performed for each time-frequency point of

the training data set for each patient. For each measure, from the 10 most significant time-frequency points, one was selected as final feature for classification of the test data (20 % trials) based on the highest prediction accuracies of the SVM classification in the validation data set (20 % trials). The five-fold cross-validation classification algorithm with adjusted numbers of trials was applied to the real and imaginary part of the phases resulting in a doubling of the number of features. The significance of classification accuracies was evaluated for each patient based on non-parametric label permutation statistics.

To compare prediction capabilities of power to absolute phase values, the same procedures which were applied to the rhinal and hippocampal phase values and phase differences were independently applied to these measures (i.e. frequency and time points were selected individually for each phase and power measure). Importantly, power values are no circular quantity and therefore the power values were entered to the classification algorithm unmodified, also, power values are only calculated for the rhinal and hippocampal channel but not as difference between these two contacts.

Additional Classifications

For the classification based on absolute phase values and phase differences two different frequency ranges were considered and their results were compared. First, based on the frequency range with most pronounced phase accumulations in the Rayleigh test, the restricted frequency range from 0.5 Hz to 13 Hz was considered for feature selection. Second, all frequencies up to 50 Hz were included. Additionally, the time windows limited to the prestimulus interval (-500 ms to 0 ms) were used for feature selection.

5.2 Results

Behavioural responses

On average, of all repeatedly presented words, 66.7 % ± 21.3 % (mean ± s.d.) were later successfully remembered as old (hits). Presented new words were wrongly categorized as old (false alarms) in 23.8 % ± 30.7 % of all cases. Hits 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 ms ± 161 ms; forgotten: 882 ms ± 232 ms; paired t-test $t_{30} = 0.175$, $p = 0.86$).

Phase Accumulation

First, frequency bands with significant phase accumulations across subjects were identified for rhinal and hippocampal phase values and rhinal-hippocampal phase differences within the frequency range from 0.5 Hz to 50 Hz (0.5 Hz steps). P-values of Rayleigh tests for each patient were combined using Fisher's method for each time-frequency point (non-overlapping 10 ms time windows) (Fig. 5.2). For all three measures, rhinal and hippocampal phases and phase differences, accumulations were mostly prominent in the low frequency range up to 13 Hz, mainly in the time range between -200 ms and 800 ms. Further accumulations were found for phase differences in the gamma frequency range between 40 Hz and 50 Hz. An accumulation of the rhinal-hippocampal phase difference points to a synchronization with a consistent coupling phase between the two regions (phase lags were clustered around zero). The extensive examination of possible influence of volume conduction is not discussed in this thesis (since it does not provide important information for the presented SVM algorithm) but data and control analyses addressing the possible influence can be found in detail in the supplementary material in Höhne et al. (2016).

Differences Between Conditions

Then, a non-parametric multi-sample circular analogue to the Kruskal-Wallis test was performed on the training data trials for each patient individually to identify frequencies and time intervals with significant differences in median phase direction between "later remembered" and "later forgotten" trials for each patient. Based on the overall results of the Rayleigh test, the two frequency ranges up to 13 Hz and up to 50 Hz were considered for the analysis. For the frequency range up to 13 Hz, significant ($p < 0.05$) differences between conditions were found in 2.9 ± 1.9 frequencies per measure (rhinal and hippocampal phase values and rhinal-hippocampal phase differences) with a mean length of significant intervals of 36 ms ± 24 ms on average

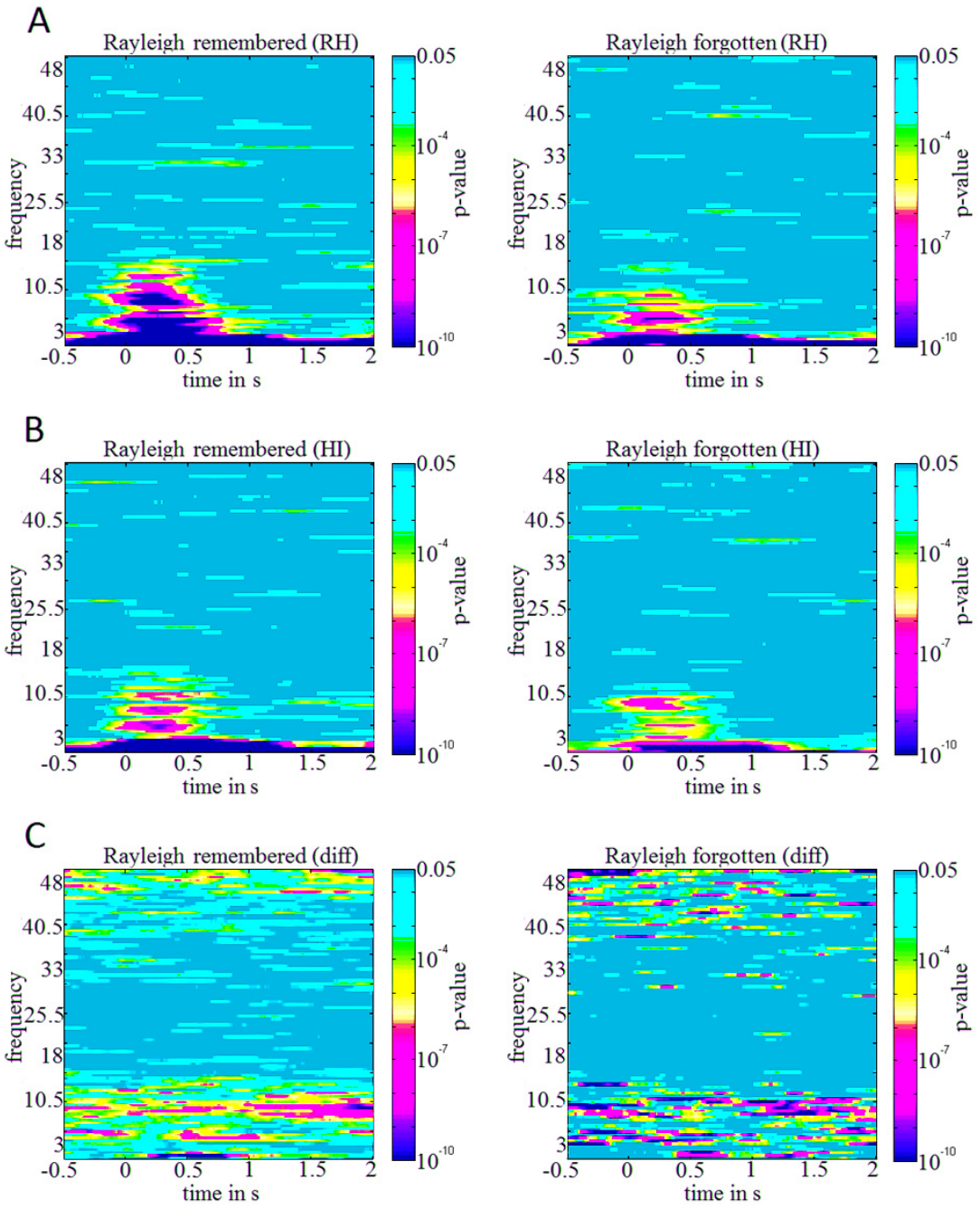

Figure 5.2: ***Results of the Rayleigh tests.*** *Fisher combined p-values of Rayleigh tests for phase values within rhinal cortex (**A**), hippocampus (**B**) and the phase differences between rhinal cortex and hippocampus (**C**) under the conditions "later remembered" (left column) and "later forgotten" (right). Colours indicate p-values according to a logarithmic scale, with all values* > 0.05 *coloured in dark cyan.*

across patients. Patients showed significant differences in 8.9 ± 6.1 frequencies with a mean length of 38 ms ± 14 ms regarding frequencies up to 50 Hz.

The test results for one exemplary patient are shown in Fig. 5.3 (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. Fig. 5.4 shows the differences between conditions averaged over trials for these three frequencies. 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 $-\pi$ 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. The exemplary patient is chosen in a way that it shows comparable results to the other subjects and does not represent extreme outcomes. Results of two other patients (one with more and one with less pronounced differences) can be found in the supplementary material in Höhne et al. (2016).

Classification Results

One rhinal phase, one hippocampal phase and one rhinal-hippocampal phase difference were chosen as features for classification from individual time-frequency points for each patient based on the results of the circular version of the Kruskal-Wallis test and classification accuracies in the validation data sets. The frequency range up to 13 Hz was considered for feature selection based on the results of the Rayleigh test. Alternatively, features were selected from an extended frequency range up to 50 Hz. Table 5.1 gives a list of selected frequencies and time points for each patient for both options when considering the whole time range.

The selected phase values for the exemplary patient are shown in Fig. 5.5 (frequencies as above; please see supplementary material in Höhne et al. (2016) 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

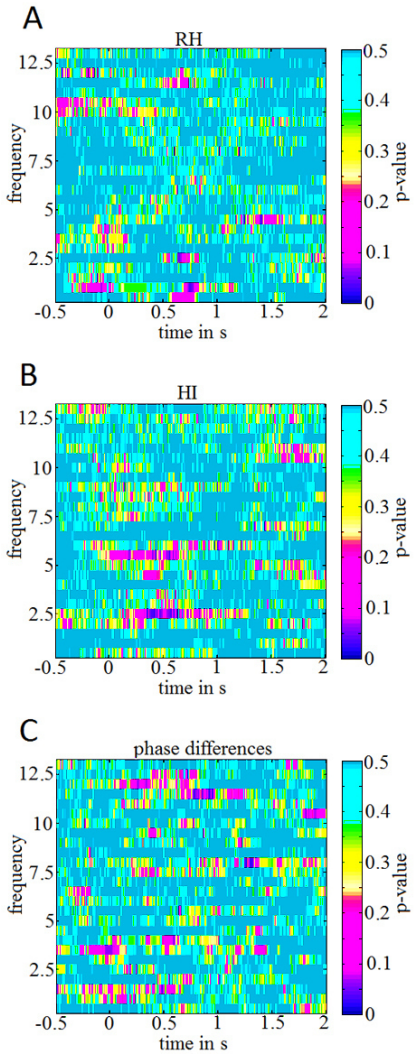

Figure 5.3: ***Circular Kruskal-Wallis test.*** *P-values for the tests "later remembered" vs. "later forgotten" for one exemplary subject (pat13) for frequencies up to 13 Hz for the phase values within rhinal cortex (**A**), hippocampus (**B**) and rhinal-hippocampal phase differences (**C**).*

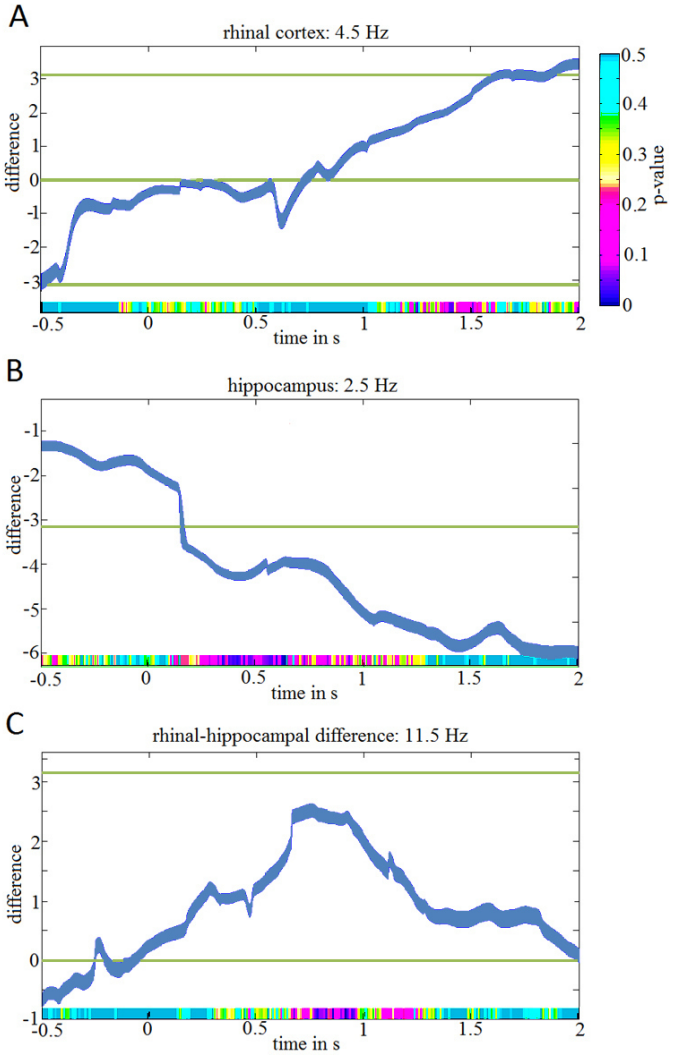

Figure 5.4: ***Mean phase differences*** *between "later remembered" and "later forgotten" over time for one exemplary subject (pat13). Mean phase differences for the rhinal phase (A), hippocampal phase (B) and rhinal-hippocampal phase difference (C) averaged over trials at the frequency that was selected for classification. Line width shows circular variance reduced by factor 5. The coloured line at the bottom indicates the p-values of the tests for differences between conditions. Green lines mark zero and $\pm\pi$.*

versus 3.61 ± 1.32 and rhinal-hippocampal phase differences at 6.07 ± 1.25 versus 0.65 ± 1.30 for prestimulus intervals.

Table 5.2 gives a list of the frequencies and time points chosen as features when feature selection was based on time windows limited to the prestimulus range. It may be possible that the Butterworth filtering has caused some temporal smearing of the poststimulus activity into the prestimulus domain. Filtering simulated signals with the chosen filter characteristics shows that such temporal smearing may extend up to half the cycle length of the filter frequency (e.g. 100 ms for 5 Hz). Based on these results, 39 (24.1 %) of the 6 x 27 = 162 values listed in Table 5.2 may be affected.

Applying the presented support vector machine, the overall classification accuracy of correct classifications into "later remembered" and "later forgotten" (averaged over all 27 subjects) for the frequency range up to 13 Hz was 66.2 %. For this frequency range, 21 subjects achieved individual classification results significantly above chance based on non-parametric label permutation statistics. The average accuracy was 67.9 % when regarding only these subjects with above chance classification. By using features from all frequencies up to 50 Hz the overall classification accuracy reached 69.2 %. Above chance results were achieved for 23 subjects with an average accuracy of 70.6 % for these patients. The individual accuracies for each patient and both frequency ranges are shown in Fig. 5.6.

Next, the time range for feature selection was limited to the prestimulus interval. Overall classification accuracy reached 61.2 % for the frequency range up to 13 Hz and above chance results were achieved for 15 patients. Including the frequencies up to 50 Hz for feature selection, above chance results were achieved for 19 subjects with an overall classification accuracy of 65.2 %. The corresponding individual accuracies are shown in Fig. 5.7.

The ability of each of the three different measures to predict successful memory formation was evaluated based on classification performance including only one measure at a time selected from the complete time range. For the frequency range up to 13 Hz, rhinal-hippocampal phase difference predicted successful memory performance most accurately (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: $F_{2,52} = 0.545$; $p > 0.5$).

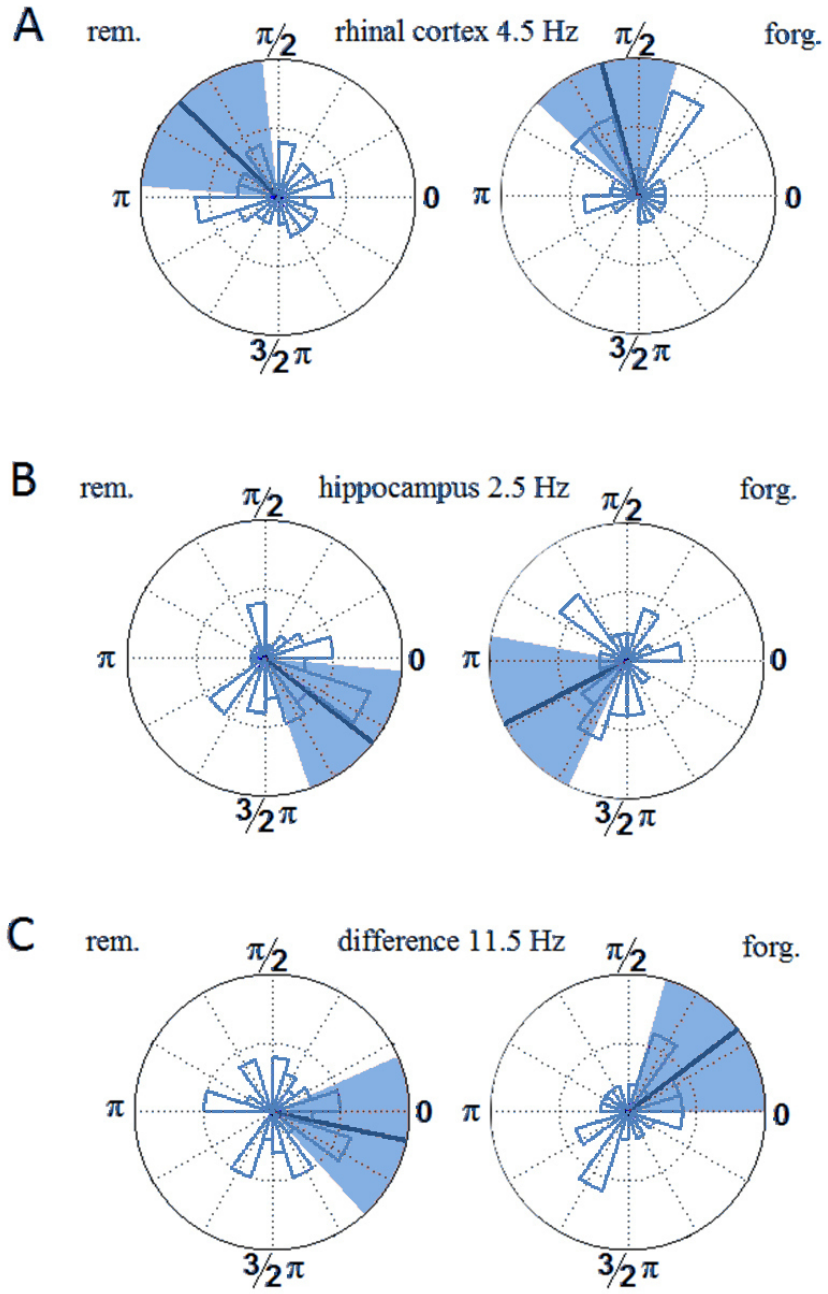

Figure 5.5: ***Phase values for one exemplary subject*** *(pat13) from the rhinal cortex (A), the hippocampus (B) and the rhinal-hippocampal phase difference (C) used for the training of the classifier. The figure shows rose diagrams of the values for the features selected for frequencies up to 13 Hz. 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 blue line; the angular deviation is displayed by shaded areas.*

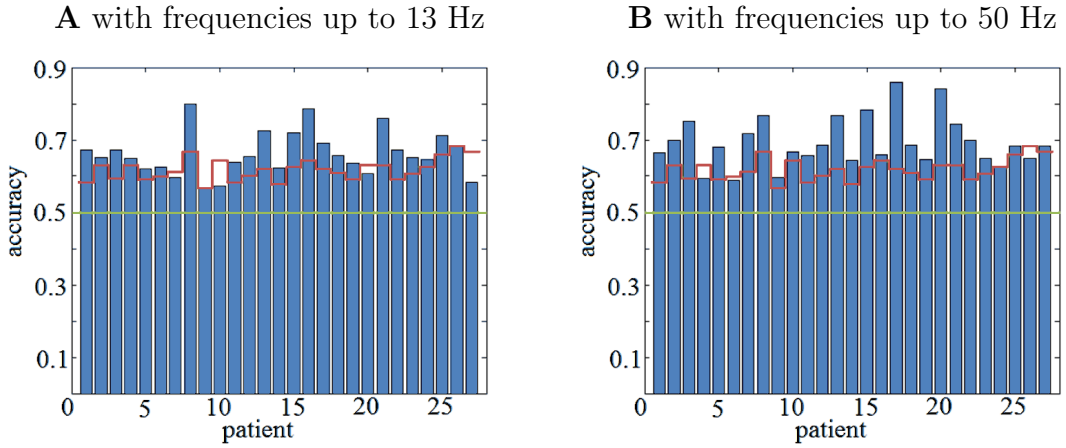

Figure 5.6: ***Individual classification accuracies for each patient.*** *Red lines mark the individual 95 % thresholds; the green line marks the 50 % accuracy.* ***(A)*** *Included frequencies up to 13 Hz.* ***(B)*** *Included frequencies up to 50 Hz.*

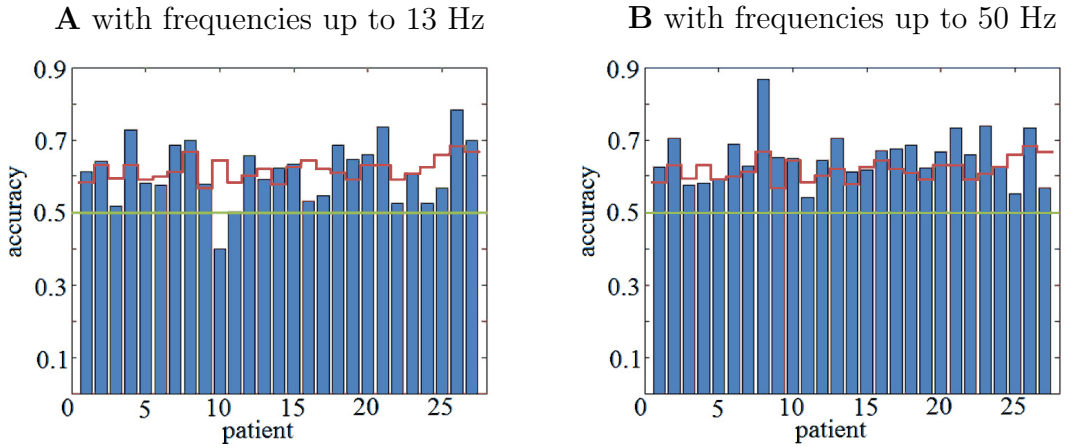

Figure 5.7: ***Individual classification accuracies for each patient for prestimulus intervals.*** *Red lines mark the individual 95 % thresholds; the green line marks the 50 % accuracy.* ***(A)*** *Included frequencies up to 13 Hz.* ***(B)*** *Included frequencies up to 50 Hz.*

For the frequency range up to 50 Hz, the ranking of classification accuracies revealed hippocampal phase as most predictive measure (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 with previous findings (e.g. Fell et al., 2001) rhinal-hippocampal phase differences were accumulated around zero (see supplementary material in Höhne et al. (2016) for data and control analyses addressing a possible influence of volume conduction). Averaged across all values selected from individual frequency-time points, the phase differences 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: $p = 0.057$).

To evaluate the predictive capabilities of single-trial power values in contrast to absolute phase values, the same procedures which were applied to phase values were independently applied to these measures. For the frequency range up to 50 Hz, average classification accuracy was 60.4 % for rhinal power and surpassed chance level in 13 subjects (vs. 63.3 % and 18 subjects for rhinal phase). For hippocampal power above chance results were achieved for 15 subjects with an overall classification accuracy of 61.5 % (vs. 64.5 % and 18 subjects for hippocampal phase). Comparing the prediction accuracies across subjects with an two-way repeated measures ANOVA revealed significantly higher accuracies for classification based on single-trial phase vs. those based on single-trial power (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).

Across subjects, prediction accuracies were not increased significantly by combining the three phase-based features (rhinal phase, hippocampal phase and rhinal-hippocampal phase difference) and the two single-trial power-based features (rhinal and hippocampal power), achieving 71.2 % overall prediction accuracy for all features vs. 69.2 % for only the phase-based features (two-sided paired t-test, $p > 0.25$).

Pat	up to 13 Hz						up to 50 Hz					
	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

Table 5.1: ***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).*

Pat	up to 13 Hz						up to 50 Hz					
	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

Table 5.2: ***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).*

5.3 Summary

Fell et al. (2008) analysed different mediotemporal EEG measures and found that measures reflecting the stability of phase values and phases differences are best to distinguish between successful and unsuccessful memory formation compared to the other examined measures. By applying the presented SVM algorithm, the predictive capability (i.e. to distinguish subsequently remembered from forgotten trials) of absolute single-trial phase values per se was examined. The described analysis revealed that such a successful above chance prediction of single-trial memory formation was possible in 85 % of patients (23 out of 27) with an average accuracy of 69.2 % across subjects. Importantly, for the prediction, only one rhinal phase value, one hippocampal phase value and one rhinal-hippocampal phase difference was used. In accordance with the findings of Fell et al. (2008) that phase-based measures outperform measures based on power values in distinguishing subsequently remembered from forgotten trials, the prediction based on single-trial phase values significantly surpassed prediction based on single-trial power.

Prediction of successful memory encoding was possible as well when reducing the time interval for feature selection to the prestimulus interval (19 out of 27 patients), although a temporal smearing of poststimulus activity into the prestimulus domain caused by the Butterworth filtering cannot be excluded. In line with this finding, there are several studies showing that successful memory performance can be related to prestimulus electrophysiological activity like ERP measures (for an overview see Cohen et al., 2015) or increased hippocampal theta activity (Fell et al., 2011; Guderian et al., 2009), as well as increased power in the 2 Hz - 4 Hz range together with increased phase synchronization especially in the temporo-parietal junction, bilateral prefrontal cortex and mediotemporal lobe (Haque et al., 2015).

Noh et al. (2014) were able to predict subsequent memory formation with an average prediction accuracy of 59.6 % across 18 subjects (with a chance level of 50 %) applying linear and SVM classifiers. They recorded high-resolution surface EEG during an object recognition paradigm with stimuli consisting of pictures of cars and birds. Compared to the attempt presented in this thesis, they used a relatively high number of features for single-trial based classification comprising pre- and peristimulus event-related potentials and EEG power in nine different frequency

bands.

These results provide evidence for the relevance of MTL EEG phases for long-term memory processes. Rhinal and hippocampal absolute phases reflect and possibly influence neural membrane potentials and thus affect the firing thresholds for neural activity via spike-field coupling (Elbert and Rockstroh, 1987). The modulation of the amount of neural excitability by field potential oscillations comparable to those measured in vivo could be found in vitro and in simulations (e.g. Anastassiou et al., 2010; Fröhlich and McCormick, 2010). In this sense, the capability of rhinal and hippocampal absolute phases to predict memory may indicate whether inhibition or facilitation of neural activity occurs within precisely the right time window required for a certain perceptual or cognitive processing sequence. This has for example been demonstrated for visual perception of stimuli close to the detection threshold, which was dependent from the phase during processing (Busch et al., 2009; Mathewson et al., 2009).

Moreover, there are other additional assumed functions of EEG phases, particularly within the MTL which can be associated to memory processes. It has been shown that the absolute phases of low-frequency hippocampal oscillations in rodents control the direction of synaptic changes (Pavlides et al., 1988; Huerta and Lisman, 1993). Additionally, rhinal-hippocampal phase synchronization is closely related to long-term memory encoding (Fell et al., 2001, 2008). This may be related to effects of rhinal-hippocampal phase differences on the communication between rhinal cortex and hippocampus and on spike-timing dependent plasticity via spike field coupling (Fries, 2005; Fell and Axmacher, 2011).

The present data confirm the relevance of absolute rhinal and hippocampal phases per se for memory formation in contrast to power. By controlling the EEG phase it could be possible to enhance memory performance. This would however only be possible in experimental settings because the oscillatory phases are changing continuously and therefore the knowledge of the exact time point of stimulus appearance is necessary. In a realistic situation, the time point at which a stimulus occurs is uncertain. Hence, the controlling of rhinal-hippocampal phase differences, which may stay relatively stable for longer time intervals, is a more practicable option.

Lately, there have been several studies exploring the effect of deep brain stimulation on memory formation (e.g. Lee et al., 2013; Suthana and Fried, 2014; Reardon,

2015). Fell et al. (2013) found that memory performance can be indeed be modulated by controlling the rhinal-hippocampal phase difference by deep brain stimulation with a frequency of 40 Hz and phase differences of 0 and 180 degree. The results presented in this chapter indicate that individually chosen stimulation frequencies and phase differences determined via classification analyses may be better suited for memory enhancement and inhibition.

By using this analysis, and applying the classification algorithm to a real iEEG dataset, it can be demonstrated that relevant information such as the importance of oscillatory phases can be detected from examining neurocognitive data. Indeed, we are able to see for certain that mediotemporal EEG phases are crucial for long-term memory processes. Absolute phases possibly regulate the timing of neural firing by interacting with neural membrane potentials, and thereby affect successful memory encoding. Based on this consideration, phase manipulation may offer the possibility for memory enhancement applications, for example through deep brain stimulation (e.g. Fell et al., 2013).

6 | Prediction of Memory Formation Based on Stimulus-Related Phase Shifts

In a further application of the algorithm to the same iEEG dataset analysed in Chapter 5 the question as to whether stimulus-related rhinal and hippocampal phase shifts are also related to successful memory formation, and if they are as well suited in predicting successful memory formation as absolute phase values, was additionally investigated. (The content of this chapter has been previously published in Derner et al. (2018b).) The analysis of these phase shifts is performed similarly to that of the absolute phases and power values, which has been outlined in detail in the previous chapter. The details of the dataset are described in Chapter 5.1.

Results summary: Differences for phase shifts between remembered and forgotten trials were found with more accumulated rhinal phase shifts and larger hippocampal phase shifts for subsequently remembered words. Prediction based on absolute single-trial phases outperforms prediction based on phase shifts. Also combining absolute phase measures with phase-shifts did not increase prediction accuracies significantly.

6.1 Material and Methods

Extraction of Phase Shifts

To compare prediction based on stimulus-related phase shifts to prediction based on absolute phases, phase shifts were calculated as the following. Absolute phases are

the phases measured at certain time points and originate from three different factors: the phase values at stimulus presentation, frequency-specific phase progressions and stimulus-related phase shifts. The phase value expected at a certain time point tx assuming no interference on the phase is defined through the phase progression. The phase progression is dependent from the chosen frequency f and can be calculated based on the time elapsed since stimulus onset as: $2\pi \cdot f \cdot tx$. The stimulus-related phase shift $\Delta\varphi$ is then given through the difference between the measured (absolute) phase and the expected phase value. It can be determined based on the relative phase φ_{rel} at the specific time point $t = tx$ (i.e. the phase at $t = tx$ relative to the phase at stimulus onset $t = 0$) which is given by the difference between the absolute phase φ_{abs} at tx and the absolute phase φ_{abs} at stimulus onset ($t = 0$): $\varphi_{rel}(tx) = \varphi_{abs}(tx) - \varphi_{abs}(0)$. With the background of phase progression and phase shift, the relative phase is also given by $\varphi_{rel}(tx) = 2\pi \cdot f \cdot tx + \Delta\varphi$. Thus, for a given time point tx and frequency f the phase shift $\Delta\varphi$ can be quantified as $\Delta\varphi = \varphi_{rel}(tx) - 2\pi \cdot f \cdot tx = \varphi_{abs}(tx) - \varphi_{abs}(0) - 2\pi \cdot f \cdot tx$.

Prediction of Subsequent Memory

To compare prediction capabilities of stimulus-related phase shift measures to absolute phase values, the presented techniques were used (see Chapter 4.2) and the same procedures which were applied to the rhinal and hippocampal phase values and phase differences (see Chapter 5.1) were independently applied to these measures (i.e. frequency and time points were selected individually for each phase and phase shift measure). Importantly, as phase shifts are a stimulus-related measure, features were only chosen from the poststimulus interval between 0 ms and 2000 ms.

Analyses of Phase Effects Related to the Dataset

Moreover, the distributions of the phase shifts selected as features for prediction were analysed. The phase-shifts on each corresponding single time-frequency point were pooled across all trials and all subjects and for each feature and condition ("remembered" and "forgotten") a Rayleigh test was performed. If a significant deviation from a uniform distribution was found, a Watson-William test (i.e. the circular analogue of the two-sample t-test) was executed to test for differences in

mean phase shifts between the two conditions.

Furthermore, several control analyses were performed for phase shifts (described in Chapter 6.3) concerning prediction accuracies in phase-scrambled surrogate data, effects related to the N400 and P600 component, dependence of prediction accuracies from signal-to-noise ratios, generalization of higher prediction accuracies for absolute phases vs. phase shifts and prediction analysis using a generalized linear model.

6.2 Results

Prediction performance based on phase shift values was compared to that based on absolute phase values. One rhinal and one hippocampal phase shift value and one rhinal-hippocampal phase difference shift value chosen from the poststimulus interval (from 0 ms to 2000 ms) and the whole frequency range up to 50 Hz were used for classification. The absolute phase values were accordingly chosen from the poststimulus interval for the comparison (see Table 6.1 and Fig. 6.1 for chosen frequencies and time points). The overall classification accuracy for phase shift values was 64.9 % with 19 subjects achieving results above chance level and 66.6 % with results above-chance level for 23 subjects for prediction based on absolute phase values (Fig. 6.2). Additionally, the individual predictive capability of the three different measures was assessed by performing classifications based on inclusion of only one measure. The ranking of classification accuracies revealed hippocampal phase values as most predictive measure (63.1 % for absolute phase with 18 subjects with above chance results vs. 61.2 % and 14 subjects surpassing chance level for phase shift), followed by rhinal-hippocampal phase differences (63.0 % and 18 subjects above chance results for absolute phase difference vs. 60.7 % for phase difference shift with 9 subjects surpassing chance level) and rhinal phase values (62.6 % for absolute phase with 19 subjects with above chance results compared to 60.0 % and 9 subjects surpassing chance level for phase shift). Across subjects, prediction accuracies for classification based on single-trial absolute phases were significantly higher than those based on single-trial phase shifts and again, the accuracies for the three different features were not significantly different from each other (two-way repeated measures ANOVA, main effect for MEASURE (absolute phase/phase shift), $F_{1,75} = 14.38$, $p = 0.0008$; no main effect for FEATURE (rhinal

cortex/hippocampus/phase difference) $F_{2,75} = 0.345$, $p = 0.71$; no interaction MEASURE x FEATURE, $F_{2,75} = 0.078$, $p = 0.93$). Prediction accuracies could not be significantly increased by combining the three absolute phase-based features chosen from the whole time interval (from -500 ms to 2000 ms) and the three poststimulus phase shift-based features, achieving 71.1 % overall prediction accuracy for all features vs. 69.2 % for only absolute phase-based features (no significant difference across subjects; paired t-test, $p = 0.14$).

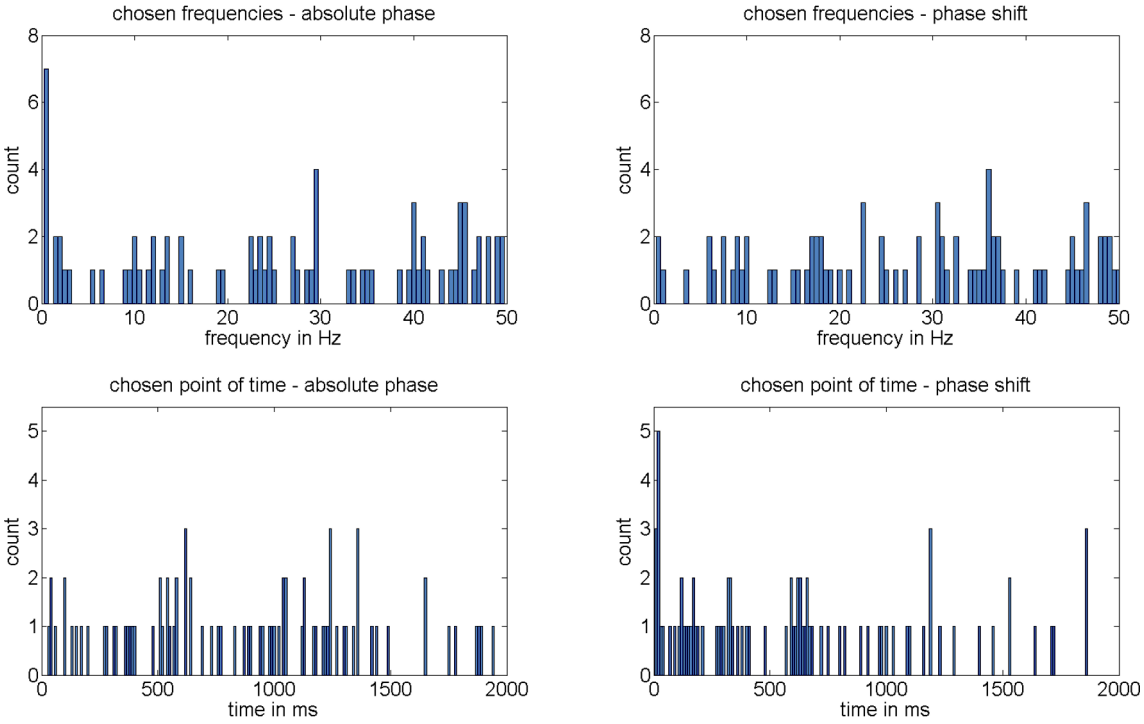

Figure 6.1: ***Frequencies and time points*** *chosen as features for classification. Distributions are shown for classification based on absolute phase values (left column) and phase shifts (right column). The depicted frequencies (top row) and time points (bottom row) are pooled over patients and all three features (i.e. rhinal phases, hippocampal phases and rhinal-hippocampal phase differences, see Table 6.1).*

Pat	Phase Shift						Absolute Phase					
	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	17.5	1400	13	1710	41.5	20	9.5	150	27.5	1300	27	1180
2	21	750	31.5	660	48.5	170	28.5	400	48	1050	19	1130
3	18	1190	9	20	36	1720	22.5	730	19.5	40	45.5	1440
4	42	1190	30.5	300	46.5	360	12	390	33.5	1050	41	270
5	50	1230	24.5	20	10	890	5.5	40	35.5	830	10	990
6	18.5	620	0.5	680	30.5	1190	29	1420	16	690	0.5	1040
7	36	140	45.5	1160	46	210	48	1210	45	940	41	170
8	3.5	590	15.5	110	9.5	600	27	1340	3	640	40	1230
9	8.5	590	37	340	45	410	49.5	1000	49.5	1360	0.5	1360
10	45	1860	46.5	400	44.5	320	40	510	24.5	380	10	130
11	0.5	330	22.5	120	22.5	70	0.5	1650	1.5	480	44.5	60
12	18	120	1	1640	49.5	1100	24	100	47	1120	45.5	1360
13	22.5	800	48.5	20	32.5	970	41.5	520	2.5	640	22.5	760
14	15	720	49	650	17	1530	35	580	0.5	1220	13.5	310
15	6	1290	36.5	610	10	280	45	620	34.5	620	29.5	550
16	7.5	130	34	1860	36.5	1460	11.5	890	0.5	1040	23.5	100
17	12.5	10	48	10	7.5	1530	2	580	45.5	980	2	1240
18	9	90	19	480	16.5	20	9	30	40.5	1270	43	1750
19	31	270	46.5	630	20	1090	47	280	24.5	200	29.5	870
20	36	620	48	820	28.5	170	0.5	900	44	360	23	320
21	37	640	17	570	41	320	38.5	1170	6.5	1130	15	1490
22	28.5	980	17.5	190	25	290	33	370	29.5	1780	13	1880
23	30.5	920	24.5	1030	34.5	30	12	1870	10.5	1240	49	1940
24	6.5	330	39	40	36	660	39.5	1650	13.5	1020	23.5	510
25	6	160	49	1860	27	380	45	570	40	1240	46.5	1890
26	32.5	630	37.5	670	35.5	10	49	540	0.5	770	29.5	620
27	35	1000	26	150	31	180	15	540	1.5	950	25	1310

Table 6.1: ***Frequencies and time points*** *based on classification features chosen in each patient. Abbreviations: RH (rhinal cortex), HI (hippocampus), diff (rhinal-hippocampal phase difference).*

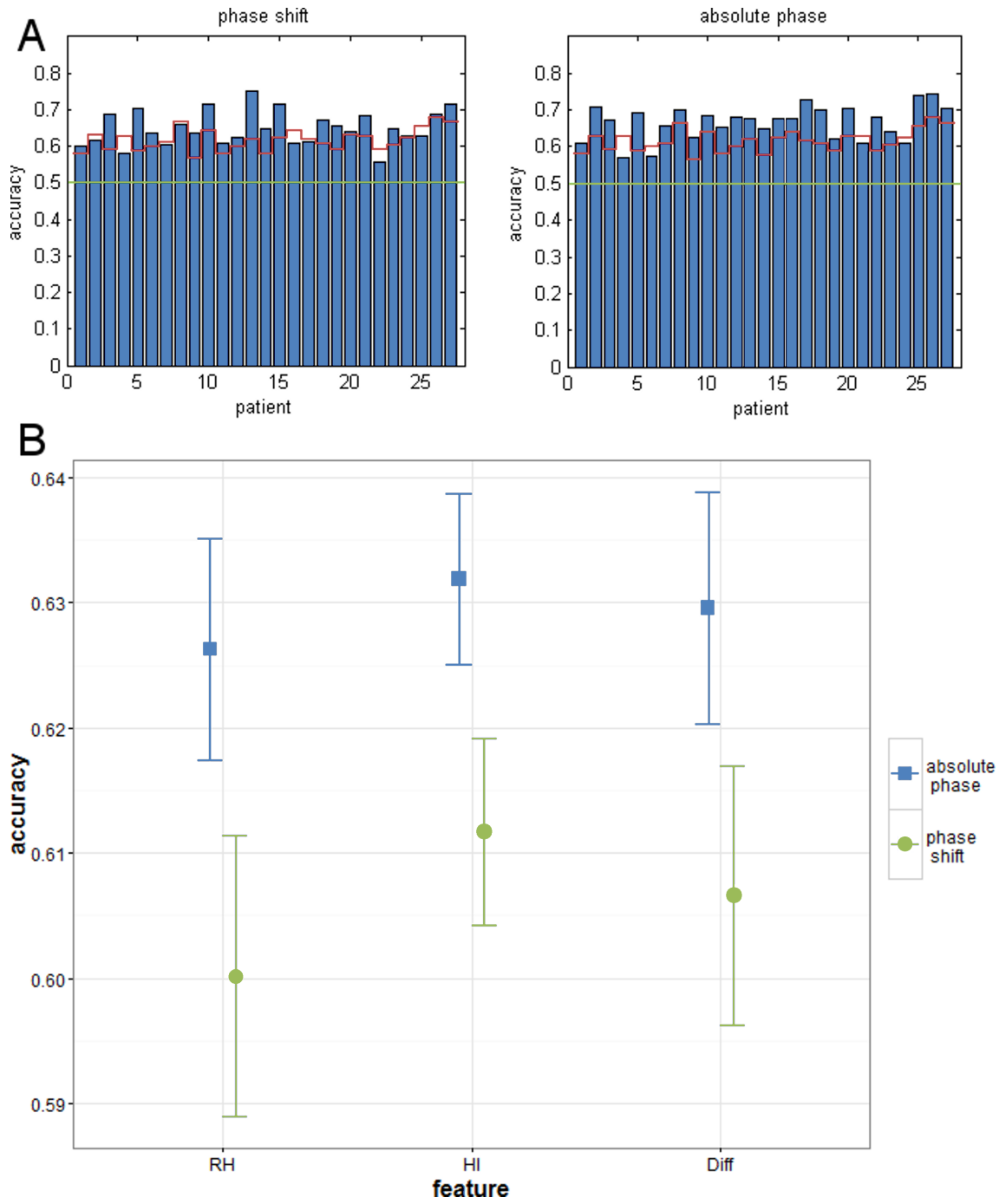

Figure 6.2: ***(A)*** *Prediction accuracies based on stimulus-related phase shifts (left side) and absolute phase values (right side) for each patient. Red lines mark the individual 95 % threshold; the green line marks the 50 % accuracy.* ***(B)*** *Mean prediction accuracies ± standard errors based on single features for absolute phase values (blue) and phase shifts (green) averaged across patients. Abbreviations: RH (rhinal cortex), HI (hippocampus), Diff (rhinal-hippocampal phase difference).*

Distributions of Phase Shifts

The phase shifts selected for prediction were pooled across all trials and subjects and their distributions were analysed for the different features and conditions. Based on Rayleigh tests, rhinal phase shifts showed a significant deviation from a uniform distribution for remembered trials ($p = 0.022$, $mean = 1.12$) but not for forgotten trials ($p = 0.44$). Additionally, significant deviations were found for shifts of rhinal-hippocampal phase differences for remembered ($p = 1.27e - 16$) and forgotten trials ($p = 2.75e - 28$), as well as trends for hippocampal phase shifts in both conditions ($p = 0.056$ for remembered and $p = 0.057$ for forgotten trials). Performing Watson-William tests for hippocampal phase shifts and shifts of rhinal-hippocampal phase differences indicated significant differences in mean phase shifts between remembered and forgotten trials for the hippocampus ($p = 0.00048$; $mean = 1.09$ vs. 0.62) with larger phase shifts for later remembered items, but no significant difference between conditions for phase shifts of phase differences ($p = 0.88$; $mean = 0.10$ vs. 0.09) (Fig. 6.3).

6.3 Control Analyses

Phase-Scrambled Surrogate Data

Phase-scrambled data was constructed to examine if the predictive capability of absolute phases and phase shifts may be influenced by data characteristics other than phase dynamics. For each patient and each channel surrogate data was generated by randomly rearranging the phases of the original data (Theiler et al., 1992), i.e. the power spectra of the surrogate data is identical to the original data. The complete feature selection and prediction procedures were repeated for these data. For absolute phases and phase shifts, prediction accuracies were not significantly higher than 50 % chance level across patients (two-tailed t-test, $p = 0.72$ for absolute phases, $p = 0.85$ for phase shifts). This shows that our results are not biased by data characteristics unrelated to phase dynamics but are specifically related to rhinal and hippocampal phase dynamics.

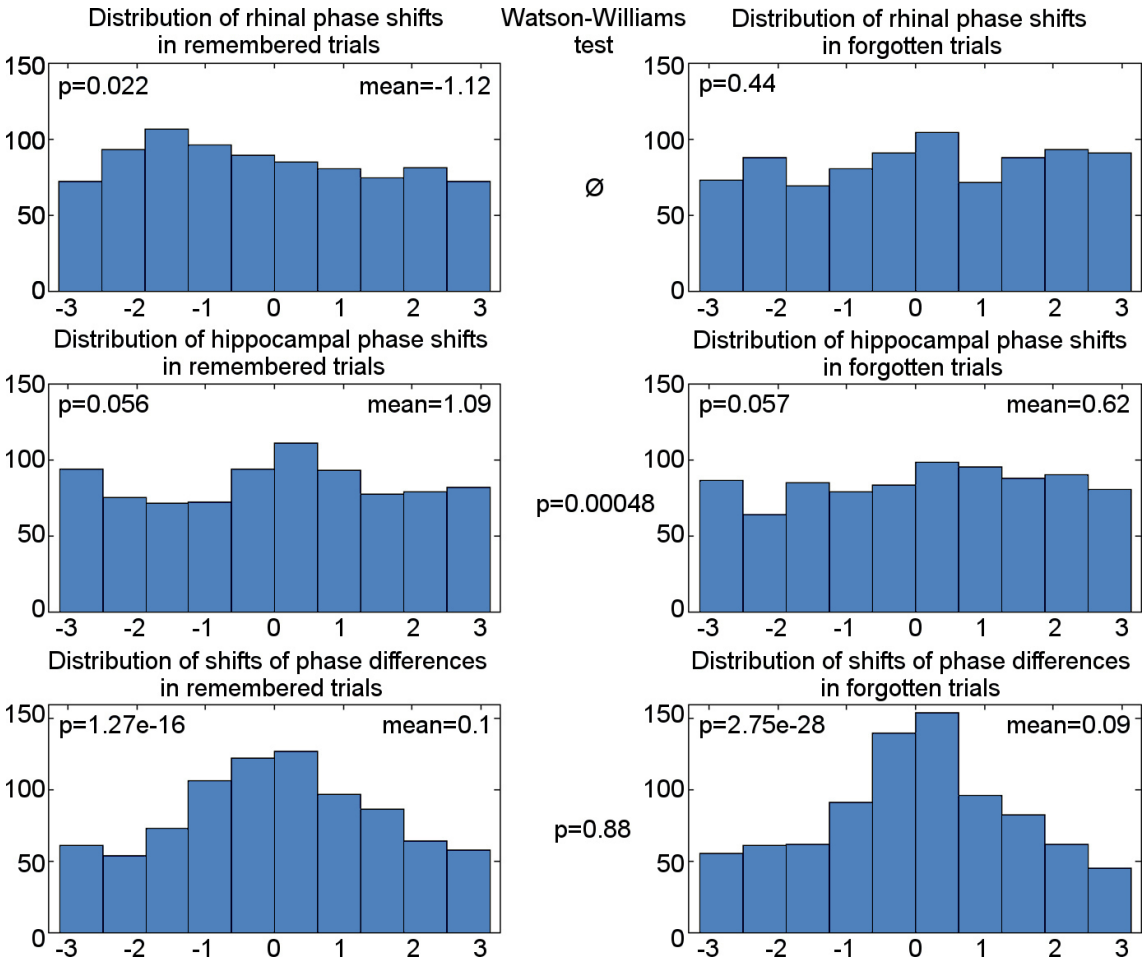

Figure 6.3: ***Distribution of phase shifts*** *for later remembered (left column) and later forgotten (right column) trials for rhinal phases, hippocampal phases and rhinal-hippocampal phase differences. The x-axis depicts the values of the phase shifts (radians* $\in [-\pi, \pi]$*, bin width* $= \pi/5$*), the y-axis depicts the number of trials exhibiting these phase shifts. P-values for the Rayleigh tests are depicted in the left upper corner of the histograms and p-values for the Watson-Williams tests can be found in the centre column.*

Effects Related to the N400 and P600 Component

Next, we tested if the prediction results reported for absolute phases and phase shifts are dominated by effects related to the rhinal N400 and hippocampal P600 component. Therefore, average ERPs were calculated separately for remembered and forgotten trials and subtracted from the individual trials for each patient. Then, based on the previously selected time-frequency points, prediction accuracies were calculated for absolute phases and phase shifts. Across patients, these accuracies did not significantly differ from the previously reported results (paired two-tailed t-test, $p = 0.96$ for absolute phases and $p = 0.65$ for phase shifts). Thus, the previously calculated accuracies cannot be mainly attributed to the N400 and P600 component.

Signal-to-Noise Ratios

The correlations between signal-to-noise ratios and prediction accuracies based on absolute phases and phase shifts were evaluated for rhinal cortex and hippocampus. The signal-to-noise ratios were calculated for each patient by dividing the absolute values of peak amplitudes of the N400 (rhinal)/P600 (hippocampal) components through the standard deviation of amplitudes across all time points and trials (in the corresponding channel). There were statistical trends for positive correlations between individual prediction accuracies based on absolute phases and rhinal signal-to-noise ratios (Pearson's cross-correlation $corr = 0.34$, $p = 0.084$) and between prediction accuracies based on phase shifts and hippocampal signal-to-noise ratios ($corr = 0.37$, $p = 0.055$) (Fig. 6.4). No significant correlation was found between prediction based on absolute phases and hippocampal signal-to-noise ratios ($p = 0.81$) or accuracies based on phase shifts and rhinal signal-to-noise ratios ($p = 0.41$).

Generalization Across Frequencies

To test if the result that prediction accuracies based on absolute phases surpass those based on phase shifts generalizes for features chosen for all patients (instead of individual feature selection), a frequency-resolved prediction analysis was performed. For this analysis, fixed time windows corresponding to the N400 and P600 component were chosen (i.e. a time window centred at 400 ms for rhinal and at 600 ms

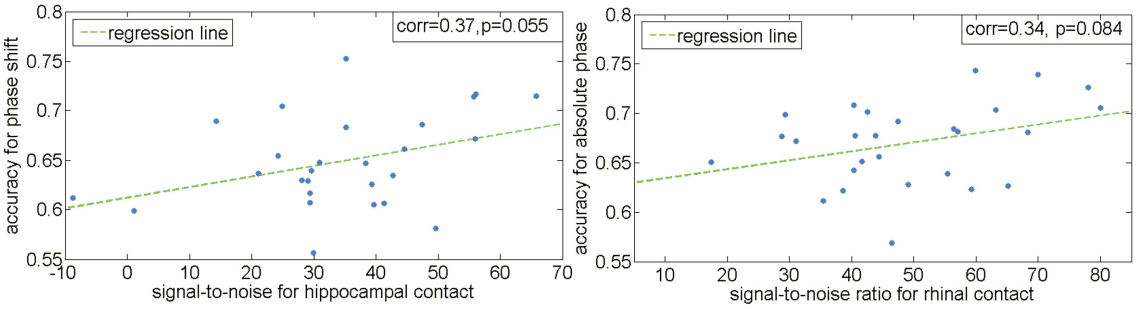

Figure 6.4: ***Prediction accuracies for each patient versus signal-to-noise-ratios (SNR)*** *for the chosen contacts. Regression lines are plotted in dashed style. Pearson's cross-correlations and related p-values are depicted in the right upper corner. Left: Prediction accuracies based on phase shifts versus SNR for hippocampal contacts. Right: Prediction accuracies based on absolute phases values versus SNR for rhinal contacts.*

for hippocampal phases/phase shifts). Then, prediction accuracies were determined separately for each frequency between 0.5 Hz and 50 Hz based on absolute phases or phase shifts (Fig. 6.5). In six frequencies, prediction accuracies across patients surpassed 50 % chance level for absolute phases (6.5 Hz, 10 Hz, 10.5 Hz, 19.5 Hz, 30 Hz, 33 Hz; two-tailed t-test, each $p < 0.05$). There were no prediction accuracies different from 50 % chance level for any of the frequencies examined for prediction based on phase shifts. Accordingly, comparison of all frequencies showed significantly higher accuracies for prediction based on absolute phases vs. phase shifts (paired two-tailed t-test, $p = 0.0064$). Thus, the result of the feature selection-based analysis generalizes for the frequency-resolved analysis.

Generalized Linear Model

Finally, the results gained using the SVM prediction algorithm were compared to prediction using a generalized linear model. Frequencies in the theta/alpha range (3 Hz - 12 Hz) were chosen from fixed time windows centred at 400 ms for rhinal and 600 ms for hippocampal phases/phase shifts for prediction (again, corresponding to the N400 and P600 components). The GLM for absolute phases and phase shifts were fitted and compared based on the Akaike information criterion (AIC, Akaike,

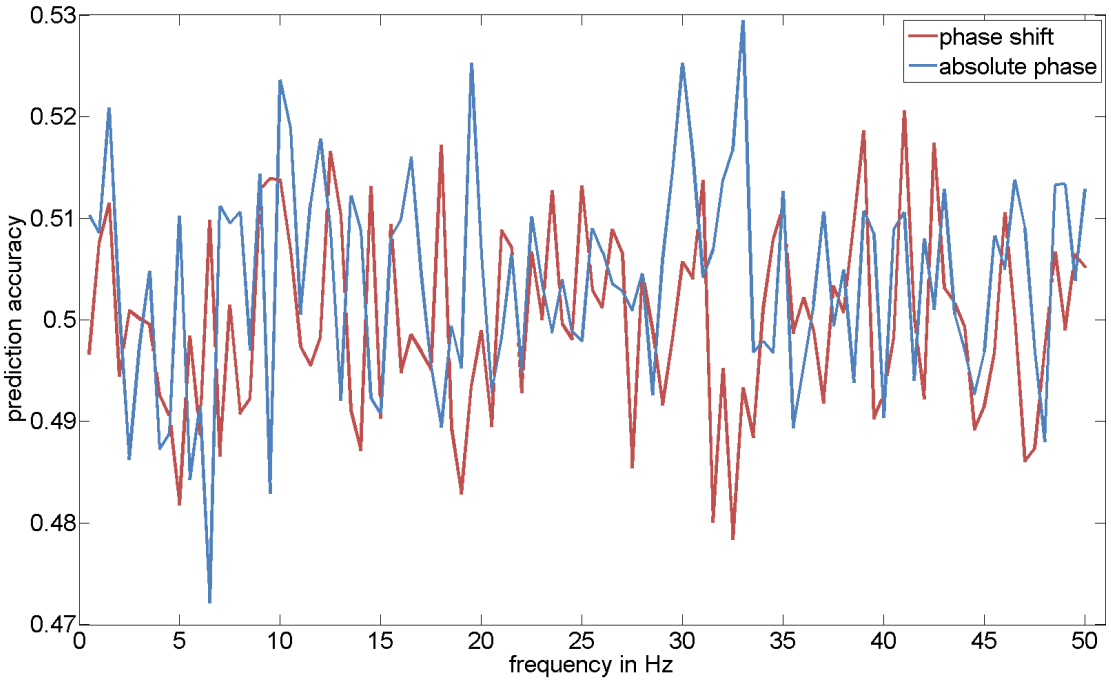

Figure 6.5: ***Frequency-resolved average prediction accuracies*** *aross patients based on rhinal and hippocampal phase shifts (red curve) and absolute phases (blue curve). Features were selected for fixed time windows centred at 400 ms (corresponding to the N400 component) for rhinal phases/phase shifts and at 600 ms (corresponding to the P600 component) for hippocampal phases/phase shifts for each frequency between 0.5 Hz - 50 Hz.*

1974) and the Vuong-test (Vuong, 1989). Evaluating the quality of predictions revealed a better model based on absolute phases with a lower AIC value of 4195 vs. the model based on phase shifts with an AIC value of 4236. The Vuong-test is suitable for comparison of non-nested GLMs and indicated a significant difference between both models ($p = 0.043$). Prediction accuracies were calculated with the GLMs fitted to the training data and applied to the test data. Average prediction accuracy for absolute phases reached 54.9 % and were above 50 % chance level across patients (two-tailed t-test, $p = 0.025$). The GLM based on phase shifts achieved an average prediction accuracy of 48.1 % and was not significantly different from 50 % chancel level ($p = 0.75$). Comparing the prediction accuracies of both models across patients yielded a trend for higher accuracies for the absolute phases v. phase shifts (paired two-tailed t-test, $p = 0.07$).

6.4 Summary

Evaluation of the predictive capability of rhinal and hippocampal phase shifts revealed that they are not as well suited in predicting memory formation as absolute phase values. Additionally, prediction accuracies were not significantly increased by combining absolute phase values and phase shifts for classification. However, phase shifts showed increased accumulation in rhinal cortex, and mean phase shifts were larger in hippocampus for subsequently remembered vs. forgotten trials, which is consistent with the suggestion that their magnitude as a marker of effective stimulus processing is important for successful memory encoding (e.g. Achuthan and Canavier, 2009). Hence, by applying the prediction algorithm, characterized in this thesis, to a real iEEG dataset, it was possible to reveal useful information about mechanisms underlying memory processes.

7 | Auditory Beat Stimulation to Control Phase Dynamics and Modulate Memory Performance

Phase synchronization underlies neural communication and facilitates synaptic plasticity, and as a result is understood to enhance memory performance (Fell et al., 2011). Applying the SVM classification method detailed in this thesis, to real data has uncovered essential information regarding the importance of EEG phases in the role of memory formation. This knowledge may be used to develop methods to alter memory performance. Becher et al. (2015) showed that monaural and binaural auditory beat stimulation with frequencies corresponding to dominant EEG bands are able to modulate EEG power and phase synchronization. The authors applied beat stimulation at 5 Hz, 10 Hz, 40 Hz and 80 Hz with low amplitudes (60 dB sound pressure level) and short durations (5 s). Several significant modulations of power and phase synchronization were found at temporo-basal, temporo-lateral, surface and also mediotemporal sites. Phase synchronization increased for 5 Hz binaural stimulation within temporo-lateral regions and decreased mediotemporal and temporo-basal for 5 Hz monaural beats. Phase synchronization has been shown to play a major role in cognitive processes, in particular in memory operations by facilitating working and long-term memory (Jutras and Buffalo, 2010; Fell et al., 2011). Hence, auditory beat stimulation (i.e. 5 Hz monaural or binaural beats) offers a non-invasive approach to interfere with iEEG characteristics. It possibly changes oscillatory phases to optimal vs. detrimental time windows for neural ac-

tivity. Hence, working and long-term memory performance may be altered through its influence on mediotemporal brain regions. The content of this chapter has been previously published in Derner et al. (2018a).

Results summary: Auditory beat stimulation is a non-invasive brain stimulation technique that is able to alter iEEG power and phase synchronization. In the presented study, data from epilepsy patients implanted with depth electrodes in the hippocampus and rhinal cortex is analysed. It is shown that 5 Hz monaural and binaural beat vs. control stimulation influenced memory performance during an associative learning task involving item and source recognition. The analysis of behavioural effects revealed a linear effect of auditory beats on memory performance: $binaural > control > monaural$ ($p = 0.036$). Additionally, increased phase-locking of 5 Hz oscillations within rhinal cortex were found for monaural and binaural stimulation but corresponding to reverse phase shifts. The data suggests that 5 Hz auditory stimulation alters long-term memory performance where opposite behavioural effects appear to be related to reverse phase shifts within rhinal cortex.

7.1 Introduction

There are a constantly growing number of persons with memory disorders like Alzheimer's disease (Wimo et al., 2003). Hence, there is a big interest in searching for new therapies that are able to improve memory. As generally known, hippocampus and rhinal cortex are central structures for long-term memory formation. Stimulating these mediotemporal regions with deep brain stimulation is one method with good prospects. So far, there are some promising studies but they show ambiguous results (for reviews see, e.g. Lee et al., 2013; Sankar et al., 2014). Contrary findings regarding memory performance are described even when performing similar experiments (e.g. Suthana et al., 2012; Jacobs et al., 2016; Hansen et al., 2018). Besides, deep brain stimulation is an invasive method requiring intracranial electrodes. Even when an application is legitimate, long-term therapy is too risky.

There are a couple of non-invasive techniques that are able to alter signals in mediotemporal regions via brain stimulation (for an overview, see Polanía et al., 2018). Along with transcranial pulsed ultrasound stimulation (Tufail et al., 2010)

and temporal inference electric stimulation (Grossman et al., 2017), auditory beat stimulation is another non-invasive method (e.g. Chaieb et al., 2015; Hommel et al., 2016). Though its effect on memory performance has not yet been carefully studied. Auditory beats are signals that are composed of two sine waves with nearby frequencies. These amplitude modulated tones are either presented directly to one or both ears (monaural beats), or the original pure sine waves are presented separately to each ear (binaural beats) and the beat perception originates from phase-sensitive brain stem neurons (Wernick and Starr, 1968). The few studies that investigated the effect of beat stimulation on cognitive processes report inconclusive results as well (see e.g. Chaieb et al., 2015, 2017). Binaural beat was found to influence EEG phases (Schwarz and Taylor, 2005; Ross et al., 2014) and power (e.g. Gao et al., 2014; Ioannou et al., 2015) as well as interregional phase synchronization (Ioannou et al., 2015). Becher et al. (2015) showed that monaural and binaural beat stimulation in the range of typical EEG rhythms is able to alter power and phase synchronization even in hippocampus and rhinal cortex. Analysing iEEG data recorded in presurgical epilepsy patients, they detected that 5 Hz binaural beats increased temporo-lateral phase synchronization while 5 Hz monaural beats decreased mediotemporal phase synchronization. Using this modulation frequency, binaural vs. monaural beats may increase vs. decrease long-term memory formation.

To study this suggestion, 5 Hz monaural and binaural beats were presented to presurgical epilepsy patients. As control stimulation, a pure sine wave at the frequency of the beat carrier frequency was applied. During stimulation, participants performed an associative learning task (Staresina et al., 2012) that is related to brain activity in the hippocampus and rhinal cortex (Staresina et al., 2013) and involves memory for item and for associated source. Since the influence of auditory beat signals on encoding and retrieval should be studied, stimulation was applied alternately during these different memory process stages for each patient. Behavioural data as well as iEEG data within hippocampus and rhinal cortex was analysed to reveal the effect of beat stimulation on memory performance as well as iEEG power, inter-trial phase-locking and rhinal-hippocampal phase synchronization.

7.2 Material and Methods

Patients

The EEG signals were recorded in 15 presurgical epilepsy patients (8 females, mean age 36.3 ± 11.4 years) who were implanted with mediotemporal depth electrodes. Based on presurgical examinations, 8 patients suffered from seizures with focus in the hippocampus, 4 patients had temporo-lateral foci and 3 patients had frontal foci. The seizure foci could be localized in the right hemisphere in 7 patients and in the left hemisphere in 6 patients. The remaining 2 patients had bilateral foci. All patients gave written informed consent to participate in the study and for the use of the data for research purposes. The study was approved by the local Ethics Committee of the Medical Faculty at the University of Bonn.

Experimental Paradigm

An associative memory task was performed (Fig. 7.1). 50 Germans nouns (per run) together with an associated source (colour (red/blue) or a scene (office/nature) depending on the experimental block) were presented for 3.5 s during the encoding phase. During the inter-stimulus intervals of 1000 ms ± 300 ms a fixation cross was displayed. The word stimuli were presented in white uppercase letters centred on a black background 160 pixels above the 200 × 300 pixels big associated source. Patients had to decide for each trial if the association between noun and colour/scene was plausible or not. After 1 min break, the retrieval phase of the task started. Here, the 50 old nouns that had been shown during the encoding phase were presented together with 25 new words for a maximum of 5 s each. Additionally to the words, four response options were displayed underneath the noun: (i) “new”, (ii, iii) the two associated sources (red or blue/office or nature) and (iv) a question mark. Patients were asked to indicate as quickly and as accurately as possible if (i) the word has not been displayed before during the encoding phase or (ii) if it was previously shown in combination with source 1 (e.g. red/office) or (iii) with source 2 (e.g. blue/nature) or (iv) if they remember the noun but not the corresponding association. Hence, the memory not only of the noun, but of the combination of the noun and the associated source was indicated with one button press. Only one source category (colour or

scene) which was randomly chosen was used for each experimental run (consisting of 50 encoding and 75 retrieval trials). The presented nouns were randomly selected out of a list of 450 nouns without overlap between the runs.

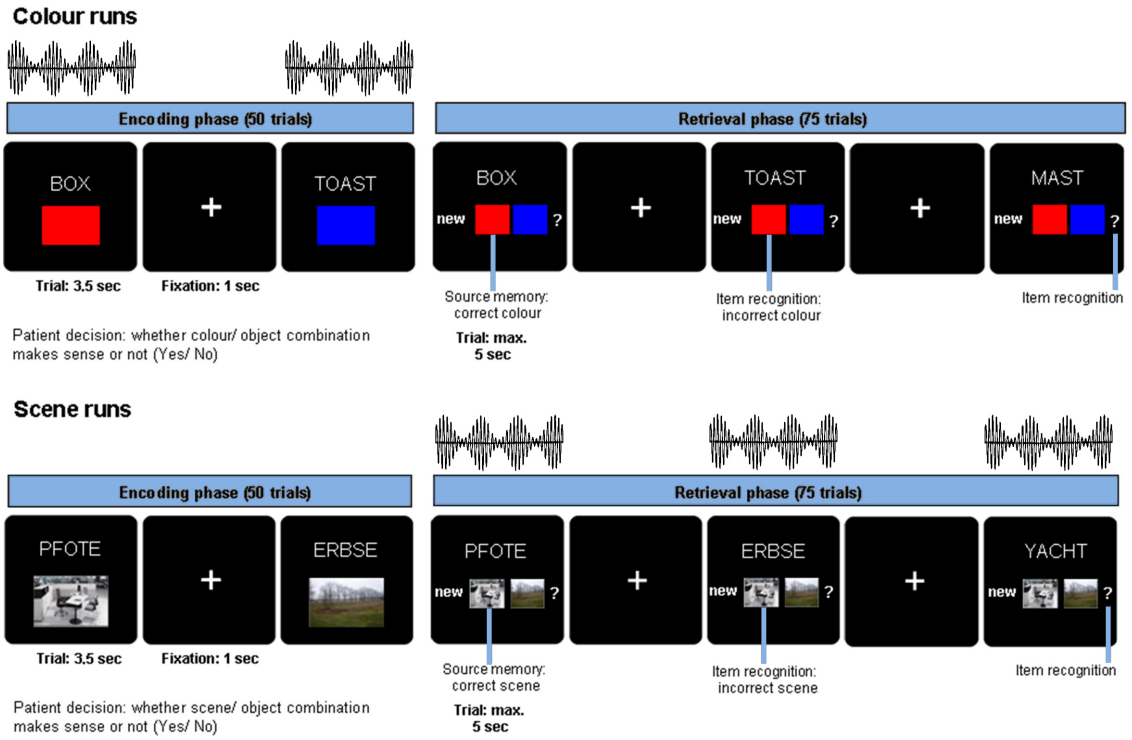

Figure 7.1: ***Associative learning paradigm and beat stimulation.*** *Nouns were presented in combination with one of two colours or one of two scenes for 3.5 seconds during encoding and patients indicated whether the combination was plausible or not. During retrieval, the previously presented nouns along with previously unstudied nouns were shown for a maximum of 5 seconds. Patients indicated their memory of the noun colour/scene association with one of four possible responses: (i) new noun, (ii) old noun associated source one, (iii) old noun associated source two, (iv) old noun but unable to remember association. Auditory beat stimulation was presented either during the encoding (colour runs, top) or during retrieval phase (scene runs, bottom).*

Auditory Beat Stimulation

The following different stimulation conditions were used: 5 Hz monaural beats, 5 Hz binaural beats and 220 Hz control tone. 5 Hz auditory beats were created by

combining two sine waves with the corresponding frequency difference. The nearby frequencies were chosen at 217.5 Hz and 222.5 Hz and physically superposed to generate monaural beats. This amplitude modulated acoustic signal (217.5 Hz sine wave plus 222.5 Hz sine wave) was presented to both ears simultaneously. For binaural beat stimulation, the sine waves were presented separately, i.e. one sine wave (e.g. 217.5 Hz) to one ear and the other sine wave (e.g. 222.5 Hz) to the other ear at the same time. The 5 Hz binaural beat percept origins from the alternating phase shifts due to the frequency mismatch between the two sine waves (with the corresponding interaural frequency difference). As control tone, a pure sine wave (without amplitude modulation) at a frequency of 220 Hz was presented to both ears simultaneously. Stimuli were produced with the NCH Tone Generator (NCH software, Canberra, ACT, Australia). And all stimuli were delivered with an average sound pressure level of 75 dB through over ear headphones via Presentation Software (Version 16.5, NeuroBehavioral Systems Inc.). The auditory stimulation was presented at stimulus onset for the duration of each trial (encoding 3.5 s, retrieval 5 s). To minimize the possibility of modulation side effects from runs with stimulation during encoding on runs with stimulation during retrieval, the auditory stimulation (monaural beats, binaural beats or control tone) was applied either for the encoding phase (in colour source runs only) or the retrieval phase (in scene source runs only) across six experimental runs (see Fig. 7.1). In this way, the possible spill-over of beat stimulation influence on associative word-colour-networks during encoding to runs with stimulation during retrieval should be reduced by engaging word-scene-networks via scene associations and vice versa. The sequence of stimulation conditions was randomized and counterbalanced across patients.

Behavioural Data Analysis

Two different memory effects were analysed in behavioural data: (i) an adjusted item memory effect given by "the probability of a hit minus the probability of a false alarm", i.e. (correct old responses to studied items / all responses to studied items) - (incorrect old responses to new items / all responses to new items) and (ii) an adjusted source memory effect given by "the probability of a hit minus the probability of a failure" (excluding "unsure" responses), i.e. (correct source decisions - incorrect source decisions) / correct old responses to studied items. One

patient performed only three colour runs, and one patient chose in the case of hits only the question mark response during one of the runs and source memory could not be reliably accessed. Hence, due to the missing data, these two patients were excluded from the behavioural analysis. The data from the remaining 13 patients were statistically analysed with repeated-measures ANOVA (with Huynh-Feldt correction) and subsequent paired two-tailed t-tests.

IEEG Recordings and Artefact Rejection

Patients were implanted with depth electrodes with 8 cylindrical platinum contacts (diameter: 1.3 mm; length: 1.6 mm) in the medial temporal lobe. Electrode contact placement was ascertained based on the individual magnetic resonance images and comparison with standardized anatomical atlases (Duvernoy, 1988). IEEG was recorded using a sampling rate of 2048 Hz and referenced to linked mastoids. Only recordings from contacts from the non-pathological hemisphere were included in the analysis. For this reason, data from two patients with no unilateral seizure onset zones but bilateral foci were excluded from iEEG analyses (but were included in behavioural analysis). Via visual inspection with BrainVision Analyzer (Version 2.0, Brain Products), channels were selected and artefacts deriving from movement, epileptic activity or technical interference were rejected. After artefact rejection, 77.6 % of all trials were included in the analysis. For each patient one rhinal and one hippocampal electrode contact was chosen based on structural MR information and average ERPs during recognition phase. Trials defined from 1000 ms before to 3000 ms after stimulus onset were baseline corrected by subtracting the averaged baseline interval spanning -200 ms to 0 ms. ERPs were determined as average across correctly classified old and new items. The rhinal contact was defined as anatomically located within the rhinal cortex with the largest mean amplitude of the negative component between 250 ms and 750 ms (memory-related anterior medial temporal lobe N400 component, e.g. Guillem et al., 1995; Nobre and McCarthy, 1995; Grunwald et al., 1999; Fell et al., 2008; Staresina et al., 2012, 2013). The hippocampal contact was defined as located within anterior or middle hippocampus and showing the largest mean amplitude of the positive component between 350 ms and 850 ms (memory-related hippocampal P600 component, e.g. Guillem et al., 1995; Ludowig et al., 2008; Fell et al., 2008; Staresina et al., 2012, 2013). After discard-

ing datasets without contacts located in RH or HI or without pronounced ERPs, datasets from seven patients could be used for further iEEG analysis. One of these patients performed only colour source runs and for one patient only a hippocampal channel was available. This resulted in the following number of contacts for the different experimental conditions: (i) colour source, auditory stimulation during encoding phase: six rhinal contacts and seven hippocampal contacts; (ii) scene source, auditory stimulation during recognition phase: five rhinal contacts and six hippocampal contacts.

Quantification of Phase Synchronization, Phase-Locking and Power Values

Phase and power values were determined for trials with auditory stimulation which was applied during encoding phases of the colour runs and retrieval phases of the scene runs. All trials in the encoding phases, and old/previously studied trials in retrieval phases were segmented from -1000 ms to 2800 ms with regard to stimulus onset and sorted according to the different stimulation conditions. This resulted in an average number of 39.5 (23 - 48) trials per patient in the encoding phase and 38.9 (28 - 47) old/previously studied trials per patient in the retrieval phase. IEEG signals were filtered at 5 Hz using continuous wavelet transforms with Morlet wavelets of five cycle length. To avoid edge effect, 800 ms were cut from the resulting signals ω at both sides, leaving the interval from -200 ms to 2000 ms. The time dependent phase values, recorded from within rhinal cortex and hippocampus ($\phi_j = arctan(Im(\omega_j)/Re(\omega_j))$), power values ($Pow_j = abs(\omega_j)^2 = Re(\omega_j)^2 + Im(\omega_j)^2$) as well as phase differences between rhinal cortex and hippocampus ($\Delta_j = \phi_j(RH) - \phi_j(HI)$) were extracted for each time point of each trial. Inter-trial phase-locking and phase synchronization are measures for directionality and are determined based on circular phase variance. They are given by the length of the mean complex phase vector (for phase-locking) and mean complex phase different vector (for phase synchronization) across all trials for each condition (Lachaux et al., 1999). To account for the expected dependency of the circular variance on the number of values entering the calculation, the number of trials was adjusted. For each patient, a random subsection was chosen from each condition by selecting the same number of trials from each condition (to match the conditions

with more trials to conditions with fewest trials; range: 23 - 43 trials). Baseline normalization of power, phase-locking and phase synchronization values was executed by dividing them by the average value from the baseline interval from -200 ms to -100 ms across trials separately for each condition and subject (i.e. the baseline level corresponds to the value 1). Since auditory beat stimulation and stimuli relevant for memory encoding and retrieval (i.e. nouns plus associative stimuli) are delivered at the same onset time, changes in phase-locking values can reflect changes in inter-trial phase concentration with regard to both, beat and memory stimuli.

Statistical Analyses

Auditory beat stimulation conditions were compared to the control condition by conducting paired t-tests across all patients for each time point (-200 ms to 2000 ms). To correct for multiple comparisons non-parametric label permutation cluster statistics were performed (Maris and Oostenveld, 2007). Based on the paired t-tests, neighbouring significant time points ($p < 0.05$) were clustered and the sum of t-values within the cluster was calculated to determine the cluster-value. Then, condition labels (beat vs. control) were permuted (31/63/127 possible permutations, corresponding to 5/6/7 contacts) and cluster-values were again calculated for the label-shuffled data. For each permutation the maximum cluster-value was determined and each cluster value for the original data was ranked among these maximum cluster-values to get the final p-value. Phase distributions of significant intervals in the rhinal phase-locking data were further analysed to compare changes during monaural and binaural beat stimulation. For this purpose, trials were merged across all patients. To ensure equal weights, the following analyses were performed 10 times, each with a new randomly chosen subset with the same number of trials for each patient (23 per condition). Results were averaged across all 10 calculations. To test if phases show significant phase accumulations (in contrast to uniform distribution), Rayleigh tests (function circ_rtest) were conducted across all patients for each time point of the selected intervals. If Rayleigh tests indicated significant phase accumulations, the differences in phase distribution between binaural and monaural beat stimulation for these time points was tested with non-parametric multi-sample tests for equal circular medians (function circ_cmtest, similar to Kruskal-Wallis tests for linear data). All circular statistics were calculated using the free CircStat toolbox

for MATLAB (Version 8.2, MathWorks Inc.; Berens, 2009). The distribution of differences between mean phases of binaural vs. monaural beat trials was additionally analysed for time points with significant Rayleigh and Kruskal-Wallis test. The phase values were averaged across colour trials (encoding) for each patient and each selected time point. The differences between both stimulation conditions were calculated and Rayleigh tests were conducted across all patients (6) and all time points (58). To test for significant differences between mean phases of binaural vs. monaural beat trials circular one-sample tests (function circ_mtest) similar to one sample t-tests were performed, testing whether the mean directions of phase differences were different from zero. Differences in mean phase values between binaural vs. control and monaural vs. control condition were analysed in the same manner.

7.3 Results

Behavioural Responses

Probability of hits minus false alarms (correct minus incorrect old decisions) indicated significant above chance recognition memory (colour: 57 % ± 25 %, $t_{12} = 8.23$, $p < 0.001$, scene: 51 % ± 24 %, $t_{12} = 7.78$, $p < 0.001$). Reaction times at the time of retrieval were significantly faster for remembered vs. forgotten words (remembered: 1.85 s ± 0.42 s; forgotten: 1.97 s ± 0.53 s; paired t-test $t_{12} = -3.03$, $p = 0.0105$). Probability for correct minus incorrect source recognition was also significantly above chance (colour: 39 % ± 22 %, $t_{12} = 6.24$, $p < 0.001$, scene: 32 % ± 25 %, $t_{12} = 4.61$, $p < 0.001$). A one-way repeated measures ANOVA revealed significantly different reaction times for the three types of source responses ($F_{2,24} = 14.82$, $p < 0.001$; correct: 1.76 s ± 0.42 s; incorrect: 1.90 s ± 0.42 s, unsure: 2.18 s ± 0.57 s) and showed a linear effect ($correct < incorrect < unsure$, $F_{1,12} = 17.18$, $p = 0.001$). Pairwise comparisons of reaction times between conditions (paired two-tailed t-tests) yielded: correct vs. incorrect: $p = 0.043$; correct vs. unsure: $p < 0.001$; incorrect vs. unsure: $p = 0.022$. Significant main effects were found for stimulation ($F_{2,24} = 4.45$; $p = 0.03$; Huynh-Feldt corrected), for association ($F_{1,12} = 6.82$, $p = 0.023$, $colour > scene$), and as expected, for memory ($F_{1,12} = 16.17$, $p = 0.002$, $item > source$) in a 3-way repeated-measures ANOVA (memory: item/source; association: colour/scene; stim-

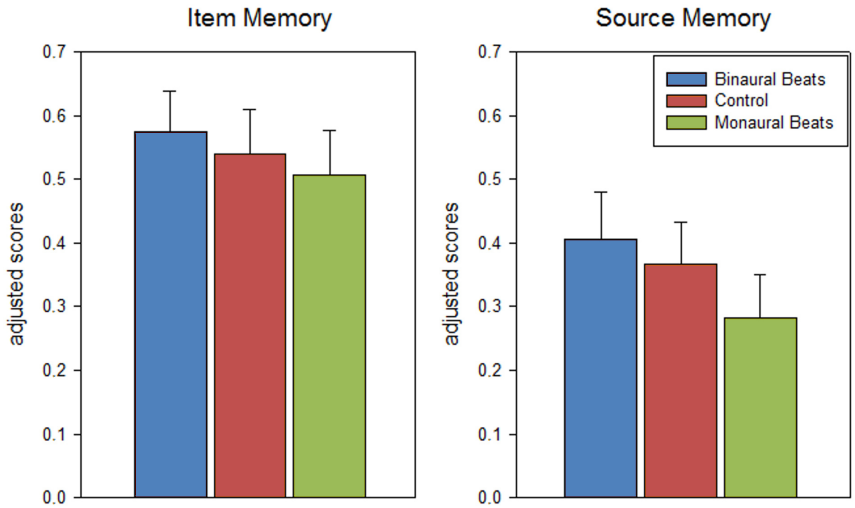

Figure 7.2: ***Behavioral results:*** *dependence of memory scores on beat stimulation conditions. Bars show mean ± S.E.M. (standard error of the mean) of adjusted scores for item (hits minus false alarms) and source (correct minus incorrect source association) memory across colour and scene runs and across patients ($n = 13$).*

ulation: binaural beat/monaural beat/control). None of the interactions between any of the three factors were significant. As hypothesized, a significant linear effect was found for the factor stimulation, revealing more adjusted hits and higher adjusted source memory under binaural beat vs. control vs. monaural beat stimulation (*binaural* > *control* > *monaural*; $F_{1,12} = 5.59$, $p = 0.036$; see Fig. 7.2; adjusted hits (mean ± S.E.M.): binaural: 0.57 ± 0.07; control: 0.54 ± 0.07; monaural: 0.51 ± 0.07; adjusted source memory: binaural: 0.41 ± 0.07; control: 0.37 ± 0.07; monaural: 0.28 ± 0.07). Pairwise comparisons between the stimulation conditions (paired one-tailed t-tests, effect size Hedges' g corrected for sample size (Hedges and Olkin, 1985) yielded for adjusted hits: *binaural* > *monaural*: $p = 0.056$, $g = 0.27$; *binaural* > *control*: $p = 0.085$, $g = 0.13$; *control* > *monaural*: $p = 0.161$, $g = 0.13$; and for adjusted source memory: *binaural* > *monaural*: $p = 0.024$, $g = 0.47$; *binaural* > *control*: $p = 0.162$, $g = 0.15$; *control* > *monaural*: $p = 0.015$, $g = 0.34$.

Phase Synchronization, Phase-Locking and Power Values

IEEG phase synchronization, phase-locking and power values of different stimulation conditions were compared based on non-parametric label-permutation cluster

statistics. All time points within beat stimulation and control trials were analysed at the beat stimulation frequency of 5 Hz. Clusters showing significant differences in phase-locking between beat and control conditions were found within rhinal cortex. Phase-locking values were higher for binaural beats vs. control during scene trials between 622 ms and 762 ms ($p = 0.031$), for binaural beats vs. control during colour trials between 409 ms and 611 ms ($p = 0.016$), and for monaural beats vs. control during colour trials between 595 ms and 738 ms ($p = 0.047$; see Fig. 7.3). No significant clusters were found for phase synchronization or power values. Additionally, differences of mean phase-locking values between stimulation and control trials averaged across the complete stimulation interval (0 s - 2 s) were analysed (paired t-tests). Mean phase-locking values in rhinal cortex were higher for binaural beats vs. control condition during colour trials ($p = 0.025$) and scene trials ($p = 0.032$), and for monaural beats vs. control condition during colour trials ($p = 0.043$). There were no significant differences for monaural beats vs. control during scene trials ($p = 0.45$), or for any of these contrasts in the hippocampus (each $p > 0.10$).

Rhinal Phase Values

Time intervals with significant phase-locking effects in rhinal cortex were merged (409 ms - 762 ms) for further analysis. The distributions of phase values for different stimulation conditions were compared with Rayleigh tests for each time point within the merged interval. As expected based on the phase-locking results, significant phase accumulations were found for the studied time points (colour binaural: 37.62 %, colour monaural: 49.41 %, colour control: 2.46 %, scene binaural: 15.21 %, scene monaural: 0.03 %, scene control: 2.78 %). Next, the proportion of time points with significant phase accumulation that overlapped for both beat stimulation conditions were determined. Time points overlapped in 14.66 % of the interval during colour trials and no time points overlapped for scene trials. Time points for colour trials with significant Rayleigh tests in both, monaural and binaural beat condition, were tested for significant phase differences between both conditions based on Kruskal-Wallis tests. Significant differences in average phase values for monaural vs. binaural stimulation were found in 60.72 % of the time points with significant Rayleigh tests (see Fig. 7.4 A for an exemplary time point and Fig. 7.4 B for all time points with significant Rayleigh and Kruskal-Wallis tests). Finally, the directions of

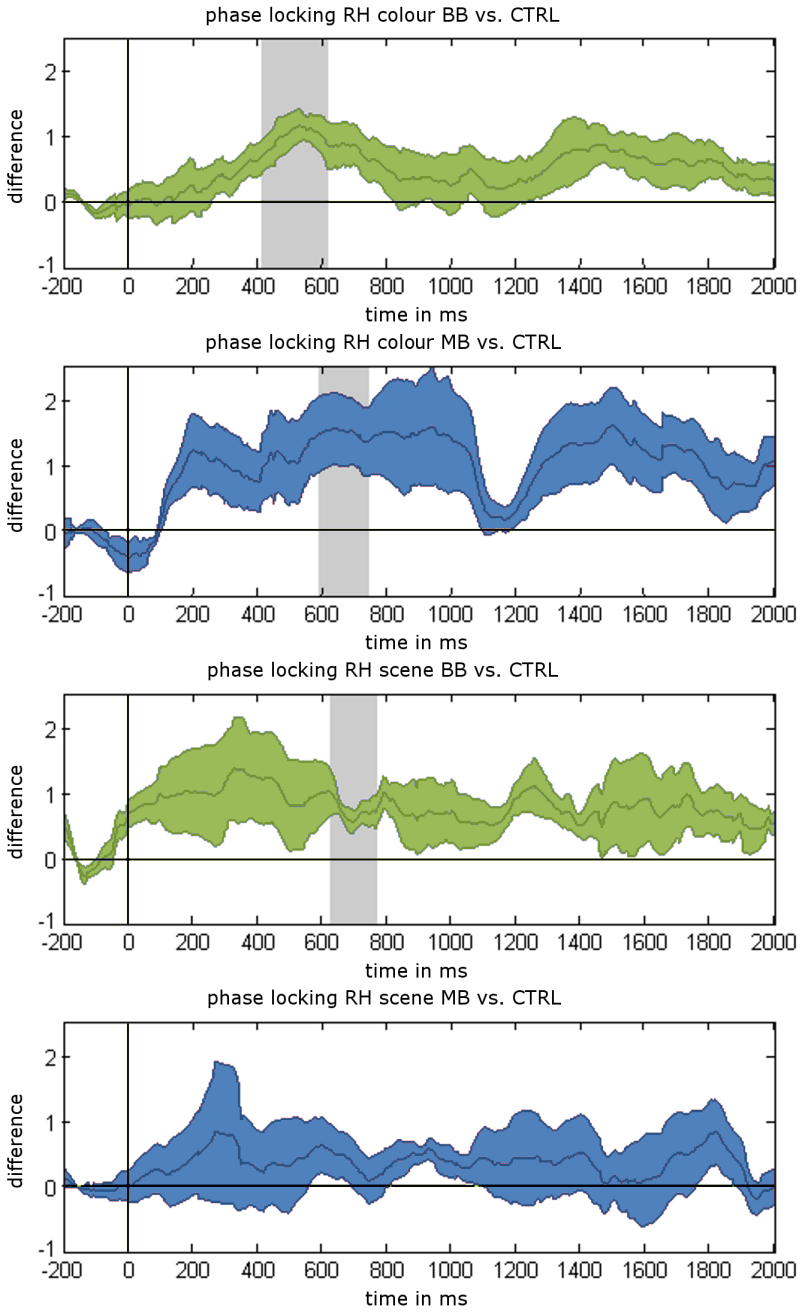

Figure 7.3: ***Phase-locking results for 5 Hz beat stimulation versus control condition.*** *Mean phase-locking differences in the rhinal cortex are shown for binaural beats (BB) vs. control (CTRL) and monaural beats (MB) vs. control stimulation in colour and source runs averaged across patients. Shaded areas indicate the standard error of the mean. Grey areas show significant time intervals in a label-shuffled cluster statistic for stimulation versus control in each condition ($p < 0.05$).*

rhinal phase data for binaural/monaural beats vs. control condition were compared. As can be seen in Fig. 7.4 A, B phase values may be significantly shifted towards opposite directions. Opposite phases can be optimal vs. detrimental for memory formation. For the time points of colour trials with significant Rayleigh and Kruskal-Wallis tests, mean phase differences were calculated for binaural vs. control, monaural vs. control and binaural vs. monaural. The differences were tested for significant phase directions with Rayleigh tests. All tested differences were significant (each $p < 0.02$) and circular one-sample tests revealed significant deviations from zero for the phase differences between binaural vs. monaural ($p < 0.001$, mean phase difference 2.35), and binaural vs. control ($p < 0.001$, mean phase difference 1.84), as well as a trend for monaural vs. control condition ($p = 0.085$, mean phase difference -0.67, see Fig. 7.4 C). These findings indicate that binaural beat stimulation shifts rhinal phases towards opposite directions compared to monaural beat stimulation.

7.4 Summary

This study investigated the influence of monaural and binaural beat stimulation on long-term memory. It found that 5 Hz binaural stimulation enhanced and 5 Hz monaural stimulation decreased memory performance in an associative learning task. The observed effect of auditory beat stimulation was similar on memory encoding and retrieval as well as on item and source memory. Caused by the small sample size the statistical power may not be sufficient to detect all interactions. In the rhinal cortex phase-locking of stimulus-related theta oscillations were enhanced for binaural and monaural stimulation compared to the control condition. These increases in phase-locking are likely due to entrainment of rhinal EEG oscillations with the beat stimuli. However, almost opposite phase values resulting from reverse phase shifts for binaural and monaural stimulation were found. Possibly, the different phase values are due to differences in sensory processing. For instance, steady-state responses to monaural beats may be triggered by signal peaks, whereas responses to binaural beats may be triggered when the intracranial sound image jumps from one ear to the other. The data suggests that rhinal phases are shifted towards optimal vs. detrimental values for memory processes by binaural vs. monaural stimulation

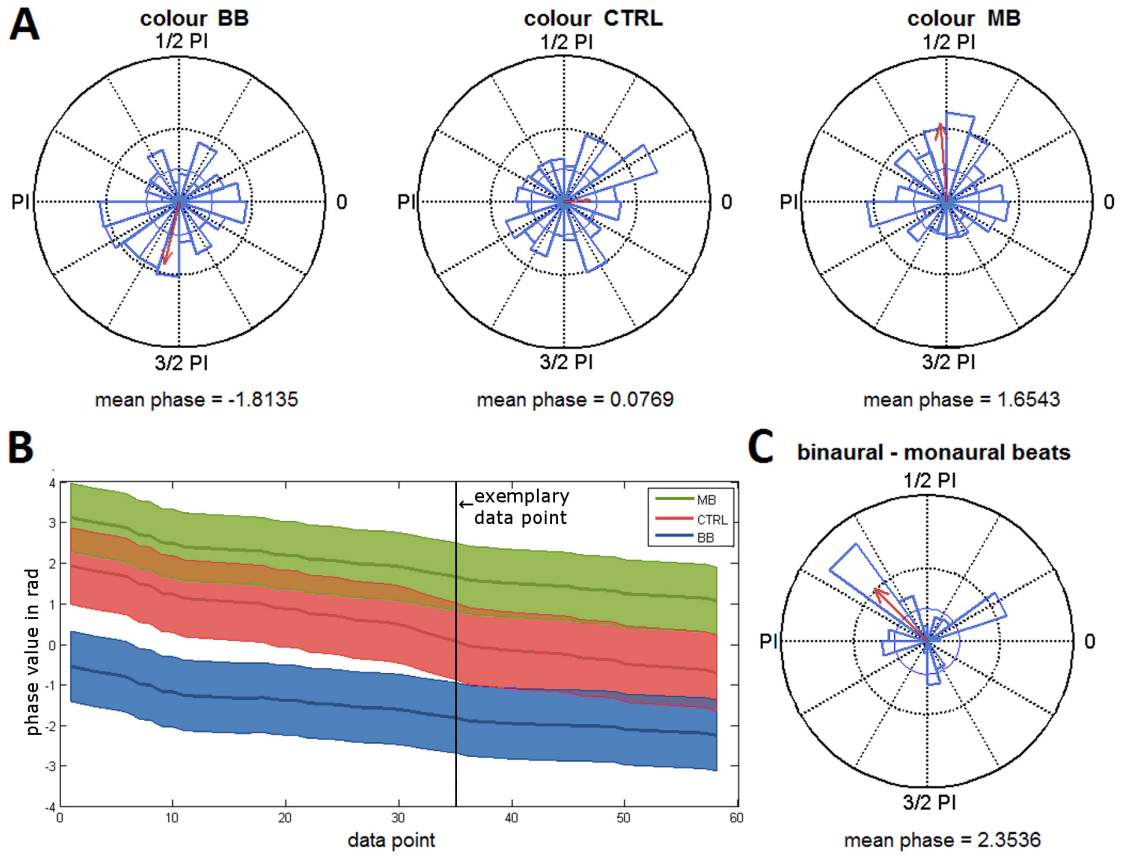

Figure 7.4: ***Phase distributions for the different beat stimulation conditions.*** ***(A)*** *Phase distributions from a time point with significant Rayleigh and Kruskal-Wallis tests (663 ms). Red arrows give the mean resultant vector with a scaled length (the inner circle indicates the vector length corresponding to a p-value of 0.05).* ***(B)*** *Mean phase values of all time points with significant Rayleigh and Kruskal-Wallis tests for binaural (BB) and monaural (MB) beat as well as for control (CTRL) condition (encoding) are shown. Shaded areas indicate circular variance. The exemplary data point marked shows the time point chosen for the plot of phase distributions in (A).* ***(C)*** *Phase differences between binaural and monaural beats for significant time points (Rayleigh and Kruskal-Wallis tests) during encoding.*

which is in line with the reported influence of phase values on neural communication (e.g. Womelsdorf et al., 2007). Furthermore, the result is similar to previous findings indicating that rhinal engagement predicts both item and associative memory during encoding (Staresina and Davachi, 2008) and retrieval (Staresina et al., 2012).

Rhinal phases most likely affect memory processes by influencing neural membrane potentials and firing threshold through spike-field coupling (e.g. Elbert and Rockstroh, 1987). That neural activity can be modulated by field potential oscillations have been showed in vitro and in simulations (e.g. Anastassiou et al., 2010; Fröhlich and McCormick, 2010). Rhinal memory operations such as semantic processing and novelty detection may be determined through an optimal vs. detrimental time window for neural activity, reflected by rhinal phase. In line with these results, it has been demonstrated that phases of human local field potentials in medial temporal regions code correct vs. incorrect matches in a card-matching task (Lopour et al., 2013). Furthermore, it has been shown that successful verbal memory encoding can be predicted based on rhinal and hippocampal phase values (Höhne et al., 2016; Derner et al., 2018b). Several studies reported similar results for scalp EEG for auditory and visual perception of stimuli close to the detection threshold (e.g. Busch et al., 2009; Mathewson et al., 2009; Neuling et al., 2012).

In a study based on intracranial EEG data from 31 epilepsy patients, stimulus-related rhinal phase-locking measures were found superior predictors of long-term memory compared to other mediotemporal EEG measures, i.e. power and phase-synchronization (Fell et al., 2008). This data is in line with the findings that stimulus-related phase-locking of EEG activity within the rhinal cortex influences long-term memory. In the studies described in the previous chapters it has been shown that prediction of successful memory encoding is possible with SVM based on rhinal and hippocampal single-trial phase values, in particular, based on single-channel phase values from rhinal cortex (Höhne et al., 2016; Derner et al., 2018b). It has to be clarified that the predictive optimal phase value at a certain time point lies within a larger range of different phase values that are predictive for remembering (e.g. $1/2\pi$ at 400 ms and $3/2\pi$ at 500 ms and other values at time points in between). This is reflected by memory-related phase-locking effects with durations of several hundreds of milliseconds (Fell et al., 2008), that have similarly been found with regard to a card matching task (Lopour et al., 2013). In conclusion, increased

rhinal phase-locking indicated by phase accumulation around a certain optimal vs. detrimental mean value can affect memory formation.

In the presented data and a previous study (Fell et al., 2008), phase-locking of theta oscillations indicating optimal phases were found in extended periods of the temporal order of a theta cycle and above. The fact that optimal phases can be found at a wider section across the theta cycle is supported by the idea of memory coding through phases across a full theta cycle. For instance, it has been shown in rodents that information is "chunked" by location selective neurons firing at specific theta phases distributed across the entire theta cycle (e.g. Colgin, 2013). The general coding mechanism underlying memory formation, in particular associative memory, may be based on this so-called theta-phase precession that is suggested to be generated in the entorhinal cortex (Yamaguchi et al., 2007). Thus, memory performance may be modulated by binaural/monaural beat stimulation through phase shifts of the rhinal theta cycle to different phases during arrival of visual information at the entorhinal-hippocampal system. This leaves the question if a shifted stimulation signal would reverse the detrimental effect of monaural beat stimulation (i.e. a shift by radian π which is a half-cycle/180 degrees). This is to be studied in the future.

Although stronger percepts caused by monaural beats compared to binaural beats were found (e.g. Grose et al., 2012), this does not necessarily mean that electrophysiological effects are stronger, but otherwise similar for monaural vs. binaural beats. Monaural beats are amplitude modulated tones presented directly to one or both ears and then decoded via the auditory cortex. The beat perception of binaural beats (i.e. the pure sine signals presented separately to each ear) however, originates from phase-sensitive brain stem neurons (Chaieb and Fell, 2017; Wernick and Starr, 1968). Becher et al. (2015) reported that mediotemporal power decreased for 5 Hz monaural beats with a highly significant effect but there was no significant decrease for 5 Hz binaural beats. On the contrary, a highly significant increase of temporo-lateral phase synchronization was found for 5 Hz binaural beats, but there was no increase for 5 Hz monaural beats. These results support the qualitatively different EEG effects of binaural beats compared to monaural beats (i.e. different effects in different brain regions). The observed linear behavioural effect for memory performance ($binaural > control > monaural$) indicates that it is not due to cap-

turing of attention by the beat stimulation (compared to control stimulation), which may divert attention from the processing of memory-related stimuli. In this case, binaural and monaural beat would both lead to decreased memory performance, with a stronger effect for monaural beats based on the stronger percept.

Different studies on binaural stimulation reported opposite effects. In line with the presented study, Ortiz et al. (2008) found increased verbal memory performance after 5 Hz binaural stimulation during the encoding phase of the memory paradigm. An opposite effect was reported by Garcia-Argibay et al. (2019a) who detected a decreased verbal memory effect related to 5 Hz binaural beats as well as by Wahbeh et al. (2007) who found a detrimental effect of 7 Hz binaural beat stimulation on immediate verbal memory recall, as measured by the Rey Auditory Verbal Learning Test. In these studies, the binaural beat stimulation was applied before and not during the memory task, which is a major difference to the study of Ortiz et al. (2008) and the presented study. Hence, the timing of beat stimulation may be crucial for the direction of memory effects, which is in agreement with the results of a recent meta-analysis (Garcia-Argibay et al., 2019b). In conclusion, the presented study suggests opposite effects of binaural and monaural beats on long-term memory based on reverse rhinal phase adjustments, probably dependent on instantaneous beat stimulation.

It has to be mentioned that it is difficult to prove that iEEG results obtained from epilepsy patients generalize to healthy subjects. A direct comparison between epilepsy patients and healthy control is not feasible since iEEG data is not recorded in healthy subjects. However, a study supporting the validity of the presented approach found that iEEG data obtained from the non-pathological MTL in patients with unilateral seizure origins during auditory and visual oddball experiment are qualitatively similar to the detected effects in iEEG data recorded in healthy monkeys (Paller et al., 1992). An influence of epilepsy pathophysiology cannot be completely excluded since there are studies reporting that the performance of healthy controls was superior to patients with temporal-lobe epilepsy in a sound-lateralization task (Tezer et al., 2012). Hence, one cannot be absolutely sure that binaural beat stimulation is processed identically in patients with temporal lobe epilepsy and in healthy subjects.

In this chapter, the results of Chapters 5 and 6 are turned into assumptions

for further research. By applying the presented SVM algorithm, it has been shown that successful memory formation can be predicted based on absolute rhinal and hippocampal phase values. Hence, mediotemporal EEG phases are crucial for long-term memory processes. The informative value of these results is confirmed in the present analysis. Here a non-invasive approach (i.e. auditory beat stimulation) to modulate EEG characteristics was applied. We can see that 5 Hz binaural and monaural beat stimulation both increase, rather than decrease phase-locking within rhinal cortex. And by changing this mechanism which underlies memory processes, long-term memory performance is altered. The reported effects of auditory beat stimulation on memory performance are in line with the results of the previous chapters that show the importance of oscillatory phases for memory formation. This means that the manipulation of phases with regard to optimal vs. detrimental time windows for neural activity offers a possible way for memory enhancement applications. In addition, we can see that key information derived from neurocognitive data can be exposed by applying the SVM algorithm detailed in the previous chapters.

8 | Conclusion

When applying pattern recognition and prediction algorithms to neurocognitive data, particular challenges have to be considered in order to develop sufficient methods for the understanding how the human brain is working. This thesis presents a machine learning approach that is able to handle these complex datasets and cope with their specific characteristics (i.e. small sample sizes, complex data structures, high-dimensionality). Regression models, which are a classical method, do not offer the necessary flexibility for complex data structures of neurocognitive data. While deep learning is a trending algorithm, it needs huge datasets for automatic pattern recognition, and in addition, the underlying mechanisms are hard to interpret. SVM is another promising prediction algorithm. It has the ability to use complex, high-dimensional training data and generalize its results to unseen samples (see e.g. Press et al., 2007). Additionally, it provides the possibility to use small datasets. Therefore, it was chosen for further development in this thesis for intended use with neurocognitive data. In particular, the typically small and unbalanced sample sizes are considered through downsampling of the observations of the larger group. By repeating the procedure several times (i.e. cross-validation) the number of trials used to assess performance is increased and the bias from random sampling is reduced. Thus, the obtained accuracy gains reliability and generalization to new data is feasible. Because of interindividual differences, it would be hard to find general mechanisms based on features across all subjects. Hence, features are chosen individually for each patient. The high-dimensionality of the data in the time and frequency domains (i.e. high sampling rate and large number of biologically relevant frequencies) offers a huge number of possible features where typically only few training trials are available. In combination with the limited electrode coverage, features have to be carefully selected. In the described algorithm, computing time

and the risk of overfitting are reduced via averaging of time points to cut the number of time windows and via feature preselection, which is realized by validation sets and circular statistics.

The aim of the presented approach was to provide as much biological information as possible in neurocognitive datasets. The results obtained by training the SVM algorithm with (i)EEG data provide valuable insight into basic mechanisms of brain functions, which was the intention of the designed method. Application of this method revealed that it is possible to successfully predict memory formation based on few phase values in significantly more patients than expected by chance and that absolute phase values are more predictive than phase shifts and power values (see Chapters 5 and 6). The informative value of these results could be confirmed when using the outcomes as research background for further studies on memory enhancement methods. For instance, it could be demonstrated that auditory beat stimulation affects memory performance ($binaural > control > monaural$) and that it is associated with increased phase locking, particularly with phase shifts of opposite directions for binaural vs. monaural beats (see Chapter 7). Hence the presented algorithm is able to expose key information derived from neurocognitive data and offers new insights into the functionality of the brain.

When applying a prediction algorithm, it is important to make sure that training (and validation) and test datasets are clearly separated for model building to achieve accurate predictions (i.e. data that is used to validate and test the model must not be included in the data to train the SVM). Only in this way is the prediction accuracy reliable, and the model is able to generalize to new data. In the presented SVM algorithm, this is ensured via partitioning of the data into training, validation and test datasets. In particular, the SVM is fitted with the preselected features of the training dataset and prediction accuracies are calculated based on the validation dataset. Hence, the evaluation of the final model is executed via the test dataset that was consistently held out before. In clinical research there are many studies that use the complete data for the fitting of the model and internal validation like bootstrapping (the same considerations apply for regression models; e.g. Molfese et al., 2001; Whalen et al., 2008; Nitschke et al., 2009; Aharoni et al., 2013). As a consequence, these models are not able to reliably estimate prediction results and have to be judged with caution. In the worst cases, the accuracies that can be

achieved with new (previously unseen) data are not better than by chance and the fitted model is useless.

SVMs typically surpass classical statistical methods like regression models in performance and generalizability. Here, a traditional GLM model was applied to the same dataset as the presented SVM algorithm (see Chapter 4.4.4). The GLM couldn't achieve successful prediction with above chance results for phase shifts and prediction accuracies were clearly outperformed by the SVM for both analysed features (i.e. absolute phases (SVM 66.6 % vs. GLM 54.9 %) and phase shifts (SVM 64.9 % vs. GLM 48.1 %)). This might be most likely caused by the stiffness of parameters that have to be chosen prior to the fitting of the model and that do not enable the necessary flexibility for complex datasets (i.e. iEEG data). Still, SVM has its limitations. With regard to medical applications, a SVM might not be the method of choice. Even if it outperforms traditional analysis methods, the objective of these kinds of applications in medical fields have to be considered. With a focus on diagnostic outcome, a particularly high prediction performance might be less important than having a functional and easy model. Furthermore, the cost-benefit ratio is an important factor. It has to be asked whether demands like longer computational time are worth the extra effort (e.g. training a SVM vs. fitting a regression model). Especially for models built on individuals, the training for each patient separately can be very time-consuming compared to simply applying an already fitted regression model (i.e. evaluating measured values by solving a specific function of an existing model).

Contrary to the practice of using simpler models, deep learning is a trending topic in many research fields. This also includes medical applications such as mind reading approaches (e.g. Spampinato et al., 2017; Wen et al., 2018; Shen et al., 2019a,b). When considering the analysis of basic brain functions, these models are hard to interpret. In turn, a reduced number of features as used in the presented SVM algorithm (in contrast to previous approaches; see 2.3) makes it easier to interpret the importance of single features and a lower prediction performance is acceptable in favour of a better understanding of the underlying mechanisms. Furthermore, deep learning approaches typically require large datasets that are often not available when working with patient data. Thus, the usually small (iEEG) datasets do not allow for the use of these models. Collecting data across different studies to get bigger samples

will probably not become a realistic option in the future. This is not only because of the small number of potential patients forming coherent groups, but because of ethical issues and data protection regulations (Boyd and Crawford, 2012) that may lead to several concerns when it comes to data sharing and collapsing.

Concerning future work, using the presented SVM algorithm for online systems is an interesting line of research. One can think about using an online system for memory enhancement applications or in further research on basic brain functions. When thinking about possible online systems with practical use, a learning assistant would be a conceivable example. In particular, using the example of learning of vocabulary, this would work as the following: The system would provide feedback after each word if it is successfully encoded or not. Based on the phase information (see Chapters 5 and 6), successful memory would be predicted. In case of ineffective memory formation, the corresponding word would be shown again later in the learning run. This way, learning can be optimized since successfully learned words are only shown once, while forgotten words will be shown repeatedly until finally learned. In other words, unnecessary repetitions are avoided while at the same time learning of each single world is assured. Naturally, recording methods have to be non-invasive. The most time consuming part of the presented method comprises individual feature selection. A huge number of different time-frequency points have to be scanned for possible predictors for each subject separately. Although feature can be extracted in parallel, due to interindividual differences, it is not possible to train the model for some patients and use it for other patients. A convenient option to train the SVM for an online application might be to use some pilot trials. Based on these individual trials a suitable time-frequency point for a specific subject can be identified prior to the actual task. However, it has to be taken into consideration that the computational time of the online system is also dependent from the selected frequency, i.e. for a low frequency, the time window for computation needs to be accordingly long (e.g. 0.5 Hz frequency needs 2 s long signals).

In conclusion, although deep learning is a trending machine learning algorithm and classical regression models might be preferred in medical applications, the presented SVM algorithm is suitable for intended use with neurocognitive datasets and can help to gain valuable insight into basic brain mechanisms. Additionally, it might offer an option for the development of clinical online applications.

Bibliography

Achuthan, S. and Canavier, C. C. (2009). Phase-resetting curves determine synchronization, phase locking and clustering in networks of neural oscillators. *J. Neurosci.*, 29:5218–5233. doi: 10.1523/JNEUROSCI.0426-09.2009.

Aharoni, E., Mallett, J., Vincent, G. M., Harenski, C. L., Calhoun, V. D., Sinnott-Armstrong, W., Gazzaniga, M. S., and Kiehl, K. A. (2014). Predictive accuracy in the neuroprediction of rearrest. *Soc. Neurosci.*, 9:332–336. doi: 10.1080/17470919.2014.907201.

Aharoni, E., Vincent, G. M., Harenski, C. L., Calhoun, V. D., Sinnott-Armstrong, W., Gazzaniga, M. S., and Kiehl, K. A. (2013). Neuroprediction of future rearrest. *PNAS*, 110:6223–6228. doi: 10.1073/pnas.1219302110.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Autom. Control*, 19:716–723.

Akbari, H., Khalighinejad, B., Herrero, J. L., Mehta, A. D., and Mesgarani, N. (2019). Towards reconstructing intelligible speech from the human auditory cortex. *Sci. Rep.*, 9:874. doi: 10.1038/s41598-018-37359-z.

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

Angrick, M., Herff, C., Mugler, E., Tate, M. C., Slutzky, M. W., Krusienski, D. J., and Schultz, T. (2019). Speech synthesis from ECoG using densely connected 3D convolutional neural networks. *J. Neural Eng.*, 16:036019. doi: 10.1088/1741-2552/ab0c59.

Anumanchipalli, G. K., Chartier, J., and Chang, E. F. (2019). Speech synthesis from neural decoding of spoken sentences. *Nature*, 568:493–498. doi: 10.1038/s41586-019-1119-1.

Becher, A. K., Höhne, M., Axmacher, N., Chaieb, L., Elger, C. E., and Fell, J. (2015). Intracranial electroencephalography power and phase synchronization changes during monaural and binaural beat stimulation. *Eur. J. Neurosci.*, 41:254–263. doi: 10.1111/ejn.12760.

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

Berns, G. S. and Moore, S. E. (2011). A neural predictor of cultural popularity. *J. Consum. Psychol.*, 22:154–160. doi: 10.1016/j.jcps.2011.05.001.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning.*, volume 1. Springer, New York.

Boyd, D. and Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Inf. Commun. Soc.*, 15:662–679. doi: 10.1080/1369118X.2012.678878.

Bruns, A. (2004). Fourier-, hilbert- and wavelet-based signal analysis: are they really different approaches? *J. Neurosci. Methods*, 137:321–332. doi: 10.1016/j.jneumeth.2004.03.002.

Bryant, R. A., Felmingham, K., Kemp, A., Das, P., Hughes, G., Peduto, A., and Williams, L. (2008). Amygdala and ventral anterior cingulate activation predicts treatment response to cognitive behaviour therapy for post-traumatic stress disorder. *Psychol. Med.*, 38:555–561. doi: 10.1017/S003329170700223.

Burke, J. F., Long, N. M., Zaghloul, K. A., Sharan, A. D., Sperling, M. R., and Kahana, M. J. (2014). Human intracranial high-frequency activity maps episodic memory formation in space and time. *NeuroImage*, 85:834–843. doi: 10.1016/j.neuroimage.2013.06.067.

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

Canli, T., Cooney, R. E., Goldin, P., Shah, M., Sivers, H., Thomason, M. E., Whitfield-Gabrieli, S., Gabrieli, J. D., and Gotlib, I. H. (2005). Amygdala reactivity to emotional faces predicts improvement in major depression. *Neuroreport*, 16:1267–1270. doi: 10.1097/01.wnr.0000174407.09515.cc.

Chaieb, L. and Fell, J. (2017). Binaural beat stimulation. In *Colzato, L. S. (eds) Theory-Driven Approaches to Cognitive Enhancement.*, volume 1, page 167–181. Springer, Cham. doi: 10.1007/978-3-319-57505-6_12.

Chaieb, L., Wilpert, E. C., Hoppe, C., Axmacher, N., and Fell, J. (2017). The impact of monaural beat stimulation on anxiety and cognition. *Front. Hum. Neurosci.*, 11:251. doi: 10.3389/fnhum.2017.00251.

Chaieb, L., Wilpert, E. C., Reber, T. P., and Fell, J. (2015). Auditory beat stimulation and its effects on cognition and mood states. *Front. Psychiatry*, 6:70. doi: 10.3389/fpsyt.2015.00070.

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

Colgin, L. L. (2013). Mechanisms and functions of theta rhythms. *Annu. Rev.Neurosci.*, 36:295–312. doi: 10.1146/annurev-neuro-062012-170330.

Costafreda, S., Khanna, A., Mourao-Miranda, J., and Fu, C. H. (2009). Neural correlates of sad faces predict clinical remission to cognitive behavioural therapy in depression. *Neuroreport*, 20:637–641. doi: 10.1097/WNR.0b013e3283294159.

Derner, M., Chaieb, L., Surges, R., Staresina, B., and Fell, J. (2018a). Modulation of item and source memory by auditory beat stimulation: A pilot study with intracranial EEG. *Front. Hum. Neurosci.*, 12:500. doi: 10.3389/fnhum.2018.00500.

Derner, M., Jahanbekam, A., Bauckhage, C., Axmacher, N., and Fell, J. (2018b). Prediction of memory formation based on absolute electroencephalographic phases

in rhinal cortex and hippocampus outperforms prediction based on stimulus-related phase shifts. *Eur. J. Neurosci.*, 47:824–831. doi: 10.1111/ejn.13878.

Doehrmann, O., Ghosh, S. S., Polli, F. E., Reynolds, G. O., Horn, F., Keshavan, A., Triantafyllou, C., Saygin, Z. M., Whitfield-Gabrieli, S., Hofmann, S. G., Pollack, M., and Gabrieli, J. D. (2013). Predicting treatment response in social anxiety disorder from functional magnetic resonance imaging. *JAMA Psychiatry*, 70:87–97. doi: 10.1001/2013.jamapsychiatry.5.

Duvernoy, H. M. (1988). *The human hippocampus. An atlas of applied anatomy.* J. F. Bergmann Verlag, München. doi: 10.1007/978-3-642-54195-7.

Elbert, T. and 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. and Axmacher, N. (2011). The role of phase synchronization in memory processes. *Nat. Rev. Neurosci.*, 12:105–118. doi: 10.1038/nrn2979.

Fell, J., Klaver, P., Lehnertz, K., Grunwald, T., Schaller, C., Elger, C. E., and Fernández, G. (2001). Human memory formation is accompanied by rhinal-hippocampal coupling and decoupling. *Nat. Neurosci.*, 4:1259–1264. doi: 10.1038/nn759.

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

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

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

Fernández, G., Effern, A., Grunwald, T., Pezer, N., Lehnertz, K., Dümpelmann, M., Van Roost, D., and Elger, C. E. (1999). Real-time tracking of memory formation in the human rhinal cortex and hippocampus. *Science*, 285:1582–1585. doi: 10.1126/science.285.5433.1582.

Föllinger, O. (2003). *Laplace-, Fourier- und z-Transformation.*, volume 1. Hüthig, Heidelberg.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.*, 9:474–480. doi: 10.1016/j.tics.2005.08.011.

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

Gao, X., Cao, H., Ming, D., Qi, H., Wang, X., Wang, X., Chen, R., and Zhou, P. (2014). Analysis of EEG activity in response to binaural beats with different frequencies. *Int. J. Psychophysiol.*, 94:399–406. doi: 10.1016/j.ijpsycho.2014.10.010.

Garcia-Argibay, M., Santed, M. A., and Reales, J. M. (2019a). Binaural auditory beats affect long-term memory. *Psychol. Res.*, 83:1124–1136. doi: 10.1007/s00426-017-0959-2.

Garcia-Argibay, M., Santed, M. A., and Reales, J. M. (2019b). Efficacy of binaural auditory beats in cognition, anxiety, and pain perception: a meta-analysis. *Psychol. Res.*, 83:357–372. doi: 10.1007/s00426-018-1066-8.

Grose, J. H., Buss, E., and Hall, J. W. I. (2012). Binaural beat salience. *Hear. Res.*, 285:40–45. doi: 10.1016/j.heares.2012.01.012.

Grossman, N., Bono, D., Dedic, N., Kodandaramaiah, S. B., Rudenko, A., Suk, H. J., Cassara, A. M., Neufel, E., Kuster, N., Tsai, L.-H., Pascual-Leone, A., and Boyden, E. S. (2017). Noninvasive deep brain stimulation via temporally interfering electric fields. *Cell*, 169:1029–1041. doi: 10.1016/j.cell.2017.05.024.

Grunwald, T., Beck, H., Lehnertz, K., Blümcke, I., Pezer, N., Kurthen, M., Fernández, G., Van Roost, D., Heinze, H. J., Kutas, M., and Elger, C. E.

(1999). Evidence relating human verbal memory to hippocampal N-methyl-D-aspartate receptors. *Proc. Natl. Acad. Sci. U.S.A.*, 96:12085–12089. doi: 10.1073/pnas.96.21.12085.

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

Guillem, F., N'aoua, B., Rougier, A., and Claverie, B. (1995). Intracranial topography of event-related potentials (N400/P600) elicited during a continuous recognition memory task. *Psychophysiology*, 32:382–392. doi: 10.1111/j.1469-8986.1995.tb01221.x.

Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., and Lounasmaa, O. V. (1993). Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.*, 65:413–497. doi: 10.1103/RevModPhys.65.413.

Hansen, N., Chaieb, L., Derner, M., Hampel, K. G., Elger, C. E., Surges, R., Staresina, B., Axmacher, N., and Fell, J. (2018). Memory encoding-related anterior hippocampal potentials are modulated by deep brain stimulation of the entorhinal area. *Hippocampus*, 28:12–17. doi: 10.1002/hipo.22808.

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

Hedges, L. V. and Olkin, I. (1985). *Statistical Methods for Meta-Analysis.* Academic Press., Cambridge, MA.

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

Helmstaedter, C., Fritz, N. E., González Pérez, P. A., Elger, C. E., and 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.*, 70:257–262. doi: 10.1016/j.eplepsyres.2006.03.005.

Hoeft, F., McCandliss, B. D., Black, J. M., Gantman, A., Zakerani, N., Hulme, C., Lyytinen, H., Whitfield-Gabrieli, S., Glover, G. H., Reiss, A. L., and Gabrieli, J. (2011). Neural systems predicting long-term outcome in dyslexia. *Proc. Natl. Acad. Sci. U.S.A.*, 108:361–366. doi: 10.1073/pnas.1008950108.

Höhne, M., Jahanbekam, A., Bauckhage, C., Axmacher, N., and Fell, J. (2016). Prediction of successful memory encoding based on single-trial rhinal and hippocampal phase information. *NeuroImage*, 139:127–135. doi: 10.1016/j.neuroimage.2016.06.021.

Hommel, B., Sellaro, R., Fischer, R., Borg, S., and Colzato, L. S. (2016). High-frequency binaural beats increase cognitive flexibility: evidence from dual-task crosstalk. *Front. Psychol.*, 7:1287. doi: 10.3389/fpsyg.2016.01287.

Huerta, P. T. and Lisman, J. E. (1993). Heightened synaptic plasticity of hippocampal CA1 neurons during a cholinergically induced rhythmic state. *Nature*, 364:723–725. doi: 10.1038/364723a0.

Ioannou, C. I., Pereda, E., Lindsen, J. P., and Bhattacharya, J. (2015). Electrical brain responses to an auditory illusion and the impact of musical expertise. *PLoS One*, 10:e0129486. doi: 10.1371/journal.pone.0129486.

Jacobs, J., Miller, J., Lee, S. A., Coffey, T., Watrous, A. J., Sperling, M. R., Sharan, A., Worrell, G., Berry, B., Lega, B., Jobst, B. C., Davis, K., Gross, R. E., Sheth, S. A., Ezzyat, Y., Das, S. R., Stein, J., Gorniak, R., Kahana, M. J., and Rizzuto, D. S. (2016). Direct electrical stimulation of the human entorhinal region and hippocampus impairs memory. *Neuron*, 92:983–990. doi: 10.1016/j.neuron.2016.10.062.

Jutras, M. J. and Buffalo, E. A. (2010). Synchronous neural activity and memory formation. *Curr. Opin. Neurobiol.*, 20:150–155. doi: 10.1016/j.conb.2010.02.006.

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., and Hudspeth, A. J. (2012). *Principles of Neural Science, Fifth Edition.* McGraw-Hill Education, New York.

Kiencke, U. (1998). *Signale und Systeme*, volume 1. Oldenbourg Verlag, München.

Koutsouleris, N., Meisenzahl, E. M., Davatzikos, C., Bottlender, R., Frodl, T., Scheuerecker, J., Schmitt, G., Zetzsche, T., Decker, P., Reiser, M., Möller, H. J., and Gaser, C. (2009). Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. *Arch. Gen. Psychiatry*, 66:700–712. doi: 10.1001/archgenpsychiatry.2009.62.

Lachaux, J. P., Rodriguez, E., Martinerie, J., and Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Hum. Brain Mapp*, 8:194–208. doi: 10.1002/(SICI)1097-0193(1999)8:4<194::AID-HBM4>3.0.CO;2-C.

Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., and Lance, B. J. (2018). EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces. *J. Neural Eng.*, 15:056013. doi: 10.1088/1741-2552/aace8c.

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

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

Ludowig, E., Trautner, P., Kurthen, M., Schaller, C., Bien, C. G., Elger, C. E., and Rosburg, T. (2008). Intracranially recorded memory-related potentials reveal higher posterior than anterior hippocampal involvement in verbal encoding and retrieval. *Cogn. Neurosci.*, 20:841–851. doi: 10.1162/jocn.2008.20507.

Manovich, L. (2012). *Trending: The promises and the challenges of big social data. Chapter in Debates in the digital humanities.*, volume 2, pages 460–475. The University of Minnesota Press. doi: 10.5749/minnesota/9780816677948.003.0047.

Maris, E. and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods*, 164:177–190. doi:10.1016/j.jneumeth.2007.03.024.

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

Maurer, U., Bucher, K., Brem, S., Benz, R., Kranz, F., Schulz, E., van der Mark, S., Steinhausen, H.-C., and Brandeis, D. (2009). Neurophysiology in preschool improves behavioral prediction of reading ability throughout primary school. *Bio. Psychiatry*, 66:341–348. doi: 10.1016/j.biopsych.2009.02.031.

Meyer, M. (2006). *Signalverarbeitung: Analoge und digitale Signale, Systeme und Filter.*, volume 4. Springer Vieweg, Wiesbaden.

Molfese, V. J., Molfese, D. L., and Modgline, A. A. (2001). Newborn and preschool predictors of second-grade reading scores: an evaluation of categorical and continuous scores. *J. Learn. Disabil.*, 34:545–554. doi: 10.1177/002221940103400607.

Mormann, F., Fell, J., Axmacher, N., Weber, B., Lehnertz, K., Elger, C. E., and Fernández, G. (2005). Phase/amplitude reset and theta-gamma interaction in the human medial temporal lobe during a continuous word recognition memory task. *Hippocampus*, 15:890–900. doi: 10.1002/hipo.20117.

Neuhäuser, M. (2011). *Nonparametric statistical tests: a computational approach.* CRC Press, Boca Raton.

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

Niedermeyer, E. and da Silva, F. L. (2004). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields.*, volume 1. Lippincott Williams & Wilkins, Philadelphia.

Nitschke, J. B., Sarinopoulos, I., Oathes, D. J., Johnstone, T., Whalen, P. J., Davidson, R. J., and Kalin, N. H. (2009). Anticipatory activation in the amygdala and anterior cingulate in generalized anxiety disorder and prediction of treatment response. *Am. J. Psychiatry*, 166:302–310. doi: 10.1176/appi.ajp.2008.07101682.

Nobre, A. C. and McCarthy, G. (1995). Language-related field potentials in the anterior-medial temporal lobe: II. Effects of word type and semantic priming. *J. Neurosci.*, 15:1090–1098. doi: 10.1523/JNEUROSCI.15-02-01090.1995.

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

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

Ortiz, T., Martínez, A. M., Fernández, A., Maestu, F., Campo, P., Hornero, R., Escudero, J., and Poch, J. (2008). Impact of auditory stimulation at a frequency of 5 Hz in verbal memory. *Actas Esp. Psiquiatr.*, 36:307–313.

Paller, K. A., McCarthy, G., Roessler, E., Allison, T., and Wood, C. C. (1992). Potentials evoked in human and monkey medial temporal lobe during auditory and visual oddball paradigms. *Electroencephalogr. Clin. Neurophysiol.*, 84:269–279.

Pavlides, C., Greenstein, Y. J., Grudman, M., and Winson, J. (1988). Long-term potentiation in the dentate gyrus is induced preferentially on the positive phase of theta-rhythm. *Brain Res.*, 439:383–387. doi: 10.1016/0006-8993(88)91499-0.

Pikovsky, A., Rosenblum, M., and Kurths, J. (2001). *Synchronization, a universal concept in nonlinear sciences.*, volume 1. Cambridge University Press, Cambridge.

Polanía, R., Nitsche, M. A., and Ruff, C. C. (2018). Studying and modifying brain function with non-invasive brain stimulation. *Nat. Neurosci.*, 21:174–187. doi: 10.1038/s41593-017-0054-4.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). *Support Vector Machines. Chapter in Numerical Recipes: The Art of Scientific Computing.*, volume 3, chapter 16.5. Cambridge University Press, New York.

Ramsundar, B., Kearnes, S. M., Riley, P., Webster, D., Konerding, D. E., and Pande, V. S. (2015). Massively multitask networks for drug discovery. *arXiv:1502.02072*.

Ramsundar, B., Liu, B., Wu, Z., Verras, A., Tudor, M., Sheridan, R. P., and Pande, V. (2017). Is multitask deep learning practical for pharma? *J. Chem. Inf. Model.*, 57:2068–2076. doi: 10.1021/acs.jcim.7b00146.

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

Rizzuto, D. S., Madsen, J. R., Bromfield, E. B., Schulze-Bonhage, A., Seeling, D., Aschenbrenner-Scheibe, R., and Kahana, M. J. (2003). Reset of human neocortical oscillations during a working memory task. *Proc. Natl. Acad. Sci. U.S.A.*, 100:7931–7936. doi: 10.1073/pnas.0732061100.

Ross, B., Miyazaki, T., Thompson, J., Jamali, S., and Fujioka, T. (2014). Human cortical responses to slow and fast binaural beats reveal multiple mechanisms of binaural hearing. *J. Neurophysiol.*, 112:1871–1884. doi: 10.1152/jn.00224.2014.

Sankar, T., Lipsman, N., and Lozano, A. M. (2014). Deep brain stimulation for disorders of memory and cognition. *Neurotherapeutics*, 11:527–534. doi: 10.1007/s13311-014-0275-0.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61:85–117. doi: 10.1016/j.neunet.2014.09.003.

Schwarz, D. W. F. and Taylor, P. (2005). Human auditory steady state responses to binaural and monaural beats. *Clin. Neurophysiol.*, 116:658–668. doi: 10.1016/j.clinph.2004.09.014.

Sederberg, P. B., Schulze-Bonhage, A., Madsen, J. R., Bromfield, E. B., McCarthy, D. C., Brandt, A., Tully, M. S., and Kahana, M. J. (2007). Hippocampal and neocortical gamma oscillations predict memory formation in humans. *Cereb. Cortex*, 17:1190–1196. doi: 10.1093/cercor/bhl030.

Shannon, C. E. (1949). Communication in the presence of noise. *Proc. IRE*, 37:10–21. doi: 10.1109/jrproc.1949.232969.

Shen, G., Dwivedi, K., Majima, K., Horikawa, T., and Kamitani, Y. (2019a). End-to-end deep image reconstruction from human brain activity. *Front. Comput. Neurosci.*, 13:21. doi: 10.3389/fncom.2019.00021.

Shen, G., Horikawa, T., Majima, K., and Kamitani, Y. (2019b). Deep image reconstruction from human brain activity. *PLoS Comput. Biol.*, 15:e1006633. doi: 10.1371/journal.pcbi.1006633.

Soon, C. S., Brass, M., Heinze, H.-J., and Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nat. Neurosci.*, 11:543–545. doi: 10.1038/nn.2112.

Spampinato, C., Palazzo, S., Kavasidis, I., Giordano, D., Souly, N., and Shah, M. (2017). Deep learning human mind for automated visual classification. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honululu, HI:4503–4511. doi: 10.1109/CVPR.2017.479.

Staresina, B. P. and Davachi, L. (2008). Selective and shared contributions of the hippocampus and perirhinal cortex to episodic item and associative encoding. *J. Cogn. Neurosci.*, 20:1478–1489. doi: 10.1162/jocn.2008.20104.

Staresina, B. P., Fell, J., Dunn, J. C., Axmacher, N., and Henson, R. N. (2013). Using state-trace analysis to dissociate the functions of the human hippocampus and perirhinal cortex in recognition memory. *Proc. Natl. Acad. Sci. U.S.A.*, 110:3119–3124. doi: 10.1073/pnas.1215710110.

Staresina, B. P., Fell, J., Lam, D., T., A., Axmacher, N., and Henson, R. N. (2012). Memory signals are temporally dissociated in and across human hippocampus and perirhinal cortex. *Nat. Neurosci.*, 15:1167–1173. doi: 10.1038/nn.3154.

Steyerberg, E. W. (2009). *Clinical Prediction Models - A Practical Approach to Development Validation, and Updating.*, volume 1. Springer, New York. doi: 10.1007/978-0-387-77244-8.

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

Suthana, N., Haneef, Z., Stern, J., Mukamel, R., Behnke, E., Knowlton, B., and Fried, I. (2012). Memory enhancement and deep-brain stimulation of the entorhinal area. *N. Engl. J. Med.*, 366:502–510. doi: 10.1056/NEJMoa1107212.

Tezer, F. I., Ilhan, B., Erbil, N., Saygi, S., Akalan, N., and Ungan, P. (2012). Lateralisation of sound in temporal-lobe epilepsy: comparison between pre-and postoperative performances and ERPs. *Clin. Neurophysiol.*, 123:2362–2369. doi: 10.1016/j.clinph.2012.06.015.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Farmer, J. D. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Physica D*, 58:77–94. doi: 10.1016/0167-2789(92)90102-S.

Thompson, R. F. (2001). *Das Gehirn - Von der Nervenzelle zur Verhaltenssteuerung.*, volume 3. Spektrum akademischer Verlag, Heidelberg.

Tufail, Y., Matyushov, A., Baldwin, N., Tauchmann, M. L., Georges, J., Yoshihiro, A., Helms Tillery, S. I., and Tyler, W. J. (2010). Transcranial pulsed ultrasound stimulates intact brain circuits. *Neuron*, 66:681–694. doi: 10.1016/j.neuron.2010.05.008.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57:307–333. doi: 10.2307/1912557.

Wahbeh, H., Calabrese, C., Zwickey, H., and Zajdel, D. (2007). Binaural beat technology in humans: a pilot study to assess neuropsychologic, physiologic, and electroencephalographic effects. *J. Altern. Complement. Med.*, 13:199–206. doi: 10.1089/acm.2006.6201.

Walter, H. (2005). *Funktionelle Bildgebung in Psychiatrie und Psychotherapie - Methodische Grundlagen und klinische Anwendungen.*, volume 1. Schattauer, Stuttgart.

Wen, H., Shi, J., Zhang, Y., Lu, K. H., Cao, J., and Liu, Z. (2018). Neural encoding and decoding with deep learning for dynamic natural vision. *Cereb. Cortex*, 28:4136–4160. doi: 10.1093/cercor/bhx268.

Wernick, J. S. and Starr, A. (1968). Binaural interaction in the superior olivary complex of the cat: an analysis of field potentials evoked by binaural-beat stimuli. *J. Neurophysiol.*, 31:428–441. doi: 10.1152/jn.1968.31.3.428.

Whalen, P. J., Johnstone, T., Somerville, L. H., Nitschke, J. B., Polis, S., Alexander, A. L., Davidson, R. J., and Kalin, N. H. (2008). A functional magnetic resonance imaging predictor of treatment response to venlafaxine in generalized anxiety disorder. *Biol. Psychiatry*, 63:858–863. doi: 10.1016/j.biopsych.2007.08.019.

Whitaker, J. C. (2000). *Signal Measurement, Analysis, and Testing.*, volume 1. CRC Press, Boca Raton.

Wimo, A., Winblad, B., Aguero-Torres, H., and von Strauss, E. (2003). The magnitude of dementia occurrence in the world. *Alzheimer Dis. Assoc. Disord.*, 17:63–67. doi: 10.1097/00002093-200304000-00002.

Womelsdorf, T., Schoffelen, J. M., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., and Fries, P. (2007). Modulation of neuronal interactions through neuronal synchronization. *Science*, 316:1609–1612. doi: 10.1126/science.1139597.

Yamaguchi, Y., Sato, N., Wagatsuma, H., Wu, Z., Molter, C., and Aota, Y. (2007). A unified view of theta-phase coding in the entorhinal-hippocampal system. *Curr. Opin. Neurobiol.*, 17:197–204. doi: 10.1016/j.conb.2007.03.007.

Zhang, Z., Duan, F., Solé-Casals, J., Dinarès-Ferran, J., Cichocki, A., Yang, Z., and Sun, Z. (2019). A novel deep learning approach with data augmentation to classify motor imagery signals. *IEEE Access*, 7:15945–15954. doi: 10.1109/ACCESS.2019.2895133.