

A study on the dynamical role of EEG phase for speech recognition

by

Takayuki Onojima

Submitted to the Department of Advanced Mathematical Sciences,
Graduate School of Informatics, Kyoto University,
in partial fulfillment of the requirements for the degree of

doctor of informatics

at the

KYOTO UNIVERSITY

Feb 2018

A study on the dynamical role of EEG phase for speech recognition

Abstract

Neural oscillation is attracting attention as an underlying mechanism for speech recognition. Neural oscillation is typically observed in human scalp electroencephalography (EEG) and has been discussed for its functional role and synchronization mechanisms. The research described in this thesis consists of two different study approaches applied to reveal the speech recognition mechanisms.

The first study shows that EEG phases modulating the behavioral performance are associated with motor cortical activity in speech recognition task. The EEG phase is reported to modulate the behavioral performance in a sound detection task. Moreover, speech recognition has been proposed to be enhanced by the identification of a speaker's motor signal. To verify the relationship, we measured scalp EEG and simultaneous EEG and functional magnetic resonance imaging (fMRI) during a speech recognition task. The results of recording the scalp EEG revealed that the behavioral performance depends on the theta and delta EEG phase before the input of speech. The simultaneous EEG-fMRI demonstrated the correlation of the theta and delta activity with the motor cortical activity. These results suggest that the EEG phase modulating the speech comprehension with motor activity.

The second study proposes an approach for identifying the dynamical systems from time-series data to reveal cross-frequency synchronization mechanisms. Synchronization of the neural oscillations can be observed during diverse cognitive tasks and has considered to be important in revealing brain function mechanisms. In particular, neural oscillations have been assumed to be synchronized with linguistic components while listening to a speech, thus synchronization plays a role in speech encoding and prediction of speech production. A phase oscillator model can describe the dynamics of neural synchronization by assuming that neural oscillations are limit-cycle oscillators that are weakly coupled to each other. This research provides a method for extracting the phase dynamics of cross-frequency synchronization from rhythmic time-series data. This method can estimate by which interaction the synchronization is induced. Furthermore, we have demonstrated our method's reliability by using numerical simulation time-series data and experiments with electronic circuits. Finally, we estimate the coupling function between EEG oscillations and the envelope of speech sound and discuss the validity of the estimated result.

In conclusion, we discuss the mechanisms by which the phase is synchronized between EEG oscillations and linguistic components and the functional role of the effects of EEG phase on speech recognition.

Acknowledgments

I would like to express my sincere gratitude to my advisor, Prof. Toshio Aoyagi, for continuously supporting my Ph.D study, providing motivation, and for his immense knowledge. I would like to thank Prof. Hiroaki Mizuhara for his patience and meaningful discussions. I am very grateful to have had Dr. Keiichi Kitajo as a collaborator and for his valuable comments and instructive advice. I am also grateful to Takahiro Goto for his collaboration. I was very impressed by his diligence.

I express special appreciation to Prof. Ikuko Nishikawa for her kind guidance. The experience in Nishikawa lab has given me the good motivation for research.

I am indebted to Dr. Yu Terada, Dr. Takashi Imai, and Dr. Keiichi Ota for their useful discussions. I thank to the members of the non-equilibrium physics and theoretical neuroscience group. I also express my sincere thanks to all of my friends for encouraging and inspiring me. I acknowledge careful and critical reading and useful comments on this thesis of Prof. Naoshi Nishimura, Prof. Satoshi Taguchi, and Prof. Syuji Miyazaki.

Last but not the least, I would like to thank my family: my parents, my grandmothers and my sister, for supporting me spiritually throughout writing this thesis and in my life in general.

THIS PAGE INTENTIONALLY LEFT BLANK

Contents

1	General Introduction	1
1.1	Neural Oscillation	1
1.2	EEG Phase and Brain Function	2
1.3	Synchronization of neural oscillation	3
1.4	Main purpose and structure of this thesis	3
2	Ongoing slow oscillatory phase modulates speech intelligibility in cooperation with motor cortical activity	5
2.1	Introduction	5
2.2	Material and Method	7
2.2.1	Participants	7
2.2.2	Experimental task	7
2.2.3	EEG recording	9
2.2.4	Simultaneous EEG-fMRI recording	10
2.2.5	Preprocessing	10
2.2.6	EEG analysis	12
2.2.7	Simultaneous EEG-fMRI analysis	15
2.3	Results	16
2.3.1	Changing behavioral performance based on EEG phase	16
2.3.2	Brain areas involved in PUB activity	20
2.4	Discussion	21
2.4.1	Delta and theta oscillations in speech recognition	21
2.4.2	Motor contribution in speech recognition	23

2.4.3	Prediction for sound input	26
3	Estimation of interaction of oscillators for cross-frequency phase syn-	
	chronization between EEG and speech sound	29
3.1	Introduction	29
3.2	Materials and methods	31
3.2.1	Estimation of phase coupling functions for cross-frequency syn-	
	chronization	31
3.2.2	Electronic circuit of van der Pol oscillator	36
3.2.3	Scalp-EEG experiment	39
3.3	Results	41
3.3.1	Numerical simulation of phase oscillator model	41
3.3.2	Electronic circuit experiment	43
3.3.3	Human EEG experiment	45
3.4	Discussion	51
3.4.1	Validity of estimation method for cross-frequency synchronization	52
3.4.2	Remarks on estimation method	53
4	Conclusion	55
4.1	Mechanisms of Synchronization between Speech Sound and EEG Os-	
	cillation	55
4.2	Functional Role of Phase effect and Synchronization	56

Chapter 1

General Introduction

1.1 Neural Oscillation

Neural oscillation is a rhythmic activity of the central nervous system. This rhythmic activity can be observed at various scales by different recording methods (e.g., single-unit recording, multi-unit recording, and electroencephalography [EEG]) and can be arrived at in many ways. An individual single neuron generates spike trains. Oscillatory activity is then observed in the membrane potential. Various intrinsic neuronal properties, such as voltage-gated ion channels, generate these oscillations. Neuronal network can also generate oscillatory activity. In these networks, neurons are interconnected and often synchronize their activities. These neuronal ensembles generate oscillatory patterns, e.g. in the local field potential (LFP). In general, these neural oscillations are observed using invasive methods such as electrocorticography. In a non-invasive approach, scalp EEG recordings can observe these large-scale neural oscillations. The EEG activity has been considered to reflect a summation of the oscillatory activity of neuronal ensembles. Recently, EEG oscillations have attracted attention as underlying mechanisms for various perceptual and cognitive functions.

In this thesis, we studied the functional role of EEG oscillation and the mechanisms by which neural synchronization is achieved.

1.2 EEG Phase and Brain Function

Neural oscillations reflect changes in the activity of a local neuronal ensemble varying between low and high excitability states [1, 2, 3, 4, 5]. Thus, there are “optimal” and “non-optimal” states for information processing in the phases of neural oscillations. LFP studies have reported that the spike rate is affected by the phase of spontaneous neural oscillation effect [2]. These results suggest that the neural oscillatory phase represents the sensitivity for the neuronal input. Furthermore, the EEG oscillation’s phase during speech and visual perception an effect on the intelligibility and predictability of speech [5, 6, 7, 8]. Previous studies have suggested that these neural oscillations are synchronized with periodic external input and may thus play a role in speech encoding and in the prediction of the timing of a stimulus [5, 9]. Thus, neural oscillations are attracting attention as an underlying mechanism for various brain functions.

However, revealing the relationship between EEG oscillations and brain function is difficult due to the low spatial resolution. The brain has functional localization, suggesting that different areas in the brain are specialized for different functions [10]. Therefore, it is very important to identify brain areas associated with different activities in the EEG. To solve this issue, a method for simultaneous EEG and functional magnetic resonance imaging (simultaneous EEG-fMRI) was proposed. fMRI records the hemodynamic changes in the brain by the blood-oxygen-level-dependent (BOLD) contrast imaging effect and has high spatial resolution, but low temporal one. Thus, a simultaneous EEG-fMRI allows the recording of both neuronal and hemodynamic activity, which are two important elements of the neurovascular coupling mechanism. This method can detect the fMRI activity correlated with EEG activity, thus allowing a discussion of the relationship between the EEG’s activity and brain function.

1.3 Synchronization of neural oscillation

Various cognitive tasks allow the synchronization of neural oscillations to be observed, and this activity is important for revealing the mechanisms underlying cognitive functions. Different kinds of neural synchronization can be observed: synchronization between neural oscillations and a periodic external input, same-frequency synchronization, and different-frequency synchronization. Numerous previous studies have suggested that these synchronizations contribute to various mechanisms of information processing such as neural communication, encoding, prediction, and attention [5, 9, 11]. However, it is not clear how the dynamic generation of neural synchronization takes place. Phase resetting generating phase synchronization is one possibility. Previous studies have suggested that phase resetting generates phase locking phase resetting generates phase locking across all trials at stimulus onset [12, 13]. This mechanisms allows phase synchronization by strong resetting of the phases of neural oscillations. Another possibility is that neural synchronization is induced by weak modulation of the phase velocity. In this mechanism, the neural oscillations gradually become synchronized. Both mechanisms can induce synchronization by some kind of interaction. Various estimation methods are available to identify this interaction [12, 14, 15]. These methods can detect the global interaction between neural oscillations; however, it is not clear that the estimated interactions contribute to phase synchronization.

1.4 Main purpose and structure of this thesis

This thesis consists of experimental and methodological research. In chapter 2, we showed that ongoing EEG phase modulates the behavioral performance in speech recognition task [16]. This EEG activity is associated with the motor cortical activity and may contribute to the prediction of sound inputs. In chapter 3, we provide an estimation method for identifying the dynamical systems from time-series data and show the estimated results for the coupling functions between the EEG oscilla-

tions and speech rhythms [17]. In chapter 4, we discuss the mechanisms of phase synchronization and functional role of neural oscillation for speech recognition.

Chapter 2

Ongoing slow oscillatory phase modulates speech intelligibility in cooperation with motor cortical activity

This chapter is constructed based on the published paper [16].

2.1 Introduction

Neural oscillations are attracting attention as underlying mechanisms of perceptual and cognitive abilities. The neural oscillation is a rhythmic neural activity and can be observed by EEG recordings. Recent studies reported the cortical neural excitability depends on the phase of ongoing neural oscillations in sound detection tasks [2, 18, 19]. Neural excitability can modulate perceptual performance in visual or auditory processing [6, 7, 8, 20]. Especially for auditory perception, the phase of slow oscillations (delta or theta activities) modulate the behavioral performance in scalp-EEG studies [6, 20]. The neural oscillation is important activity which mediates the process of effective auditory perception. The neural oscillatory phase can be modulated by periodic external input [2, 21]. In vocal communication, the speech rhythm entrains

the neural activity in the listener’s brain. The delta and theta oscillations show the phase synchronization to the speech envelope, inducing speech recognition [12]. The speech rhythm may shift the neural oscillatory phase in a listener’s brain to an optimal state to process the speech input. It indicates that the phase of slow oscillation also modulates speech recognition performance, as well as perceptual performance.

The neural oscillation may be associated with the enhancement of speech recognition when a listener is viewing speaker’s face [5]. When speech sound is distorted or is embedded in noise, viewing a speaker’s face enhances speech recognition [22]. The facial movement represents the timing of speech production and occurs before the speech production. Furthermore, the neural oscillatory phase before sound presentation affects the behavioral performance in a sound detection task [6]. It can be considered that the ongoing neural oscillatory phase contributes the prediction of speech production timing.

In addition, it is possible that viewing the facial movement enhances the speech recognition with motor cortical activity. It is well known that the motor cortex contributes to speech perception [23, 24, 25, 26, 27, 28]. Especially, phonetic identification depends on the motor cortical activity [23, 26, 27, 28]. A transcranial magnetic stimulation (TMS) on the motor cortex impairs phonemic identification in speech recognition [23, 27]. Furthermore, the activity of motor system increases when viewing the facial movement, as well as speech listening [29]. The motor cortex may contribute to the enhancement of speech recognition by viewing the articulatory action.

Taken together, we hypothesize that effect of the slow oscillatory phase is associated with the activity of motor cortex in speech recognition. We assume that the phase effect appeared before speech presentation, as well as the sound detection. In addition to the phase effect, the motor cortex contributes to the enhancement of speech recognition. The activity of motor system is increased by viewing the articulatory action or listening speech sound. The articulatory action occurs before the production of speech sound and includes the predictable information of the timing of speech production [5]. Therefore, we consider that the phase effect is associated

with motor activity. To test our hypothesis, we measured human scalp-EEG during a speech recognition task. We proposed an index to quantitatively evaluate the EEG phase effect on behavioral performance. By using the index, we show that the phase of ongoing slow oscillation prior to the speech input modulate behavioral performance in speech recognition. The simultaneous EEG-fMRI was then used to verify whether the motor cortical activity is related to EEG activities during speech recognition.

2.2 Material and Method

2.2.1 Participants

Sixteen healthy participants (all right-handed, Edinburgh handedness test score 96.4 ± 1.9 (mean $\pm$ SEM); five female; mean age = 25 years with range of 21-32 years) engaged in the EEG experiment, and ten participants (all right-handed, Edinburgh handedness test score 98.0 ± 2.0 (mean $\pm$ SEM); two female; mean age = 30 years with range of 20-39 years) engaged in the simultaneous EEG-fMRI experiment. One participant was excluded from further analysis due to the low response rate below 60 percent in the EEG experiment. One participant was excluded from further analysis due to the excessive artifact that could not be removed by EEG preprocessing in the EEG experiment. No participant was excluded in the simultaneous EEG-fMRI experiment. The experimental procedure was approved by the ethics committee of the Unit of the Integrated Studies of the Human Mind, Kyoto University (24-p-19). Participants gave their written informed consent according to the Declaration of Helsinki and were paid for their participation in the study.

2.2.2 Experimental task

The experimental task was a speech recognition task in a 2-alternative forced choice (Figure 2-1a). The participants were required to choose what they heard from one of two choices (target and distractor) as soon and accurately as possible. If the participants did not response within 2 sec, the trial was not used further analysis. The

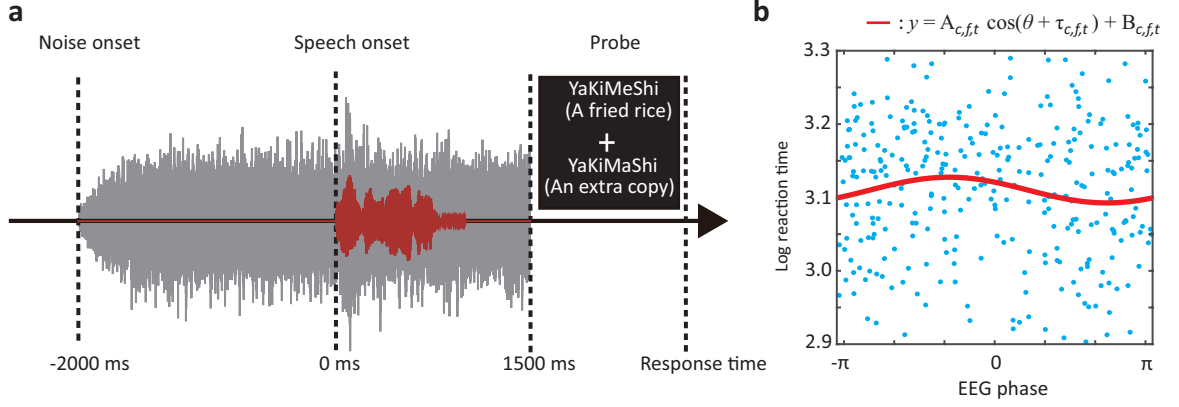


Figure 2-1: Schematic illustration of the experimental design and phase-utility on behavior (PUB) analysis. (a) Speech recognition task paradigm. Pink noise (shown as a gray wave in the figure) was presented after the trial interval, and a speech sound (shown as a red wave) was presented 2 sec after the onset of the noise. A probe was presented after the sound stimulus was presented. (b) An example of the relationship between the EEG phase and reaction time. The logarithmic reaction time was plotted as a function of the instantaneous EEG phase for each trial (blue plot). The red line was estimated by the least square method. The amplitude of the cos-curve was defined as the PUB index.

auditory stimuli were 4-letter Japanese words (e.g., “YaKiMeShi” meaning fried rice), and were spoken within approximately 1 sec. The distractors were words in which either the second or the third letter differed from the target (e.g., “YaKiMaShi” meaning extra copy). The target and distractor were in the same range of word familiarity (5-6th in a scale of 1(not familiar)-7(familiar)) . The speech spectrogram were high-pass filtered with 3 Hz cutoff frequency in the EEG experiment to rule out risks that the low-frequency component (i.e., prosodic frequency [5]) in the speech entrained the EEG activity. The speech stimulus was then masked with pink noise. Signal-to-noise (SN) ratio of the speech stimulus was set to -10 dB in the EEG experiment. The SN ratio for the EEG experiment was decided to be approximately 80 percent of accuracy in the preliminary experiment with 7 additional participants. Because MRI scan causes the big sound noise, the SN ratio of simultaneous EEG-fMRI experiment was set to +5 dB to make the accuracies of speech recognition identical between the EEG experiment and simultaneous EEG-fMRI. The accuracy was 73.9 ± 1.5 percent (mean $\pm$ SEM) for EEG experiment, and 75.5 ± 2.4 percent

for simultaneous EEG-fMRI. The reactions time was $1,190.0 \pm 44.2$ ms for EEG experiment, and $1,510.0 \pm 46.9$ ms for simultaneous EEG-fMRI. The noise volume was gradually increased over 0.5 sec after the onset to avoid the EEG phase reset by the noise onset. The speech stimulus always started 2 sec after the noise onset and lasted approximately 1 sec. The speech stimulus was randomly chosen from 120 spoken words. The trials with 21 spoken words were discarded from further analysis due to the low accuracy below the chance level in the EEG experiment. The noise sound was terminated 1.5 sec after speech onset. The probe stimulus followed the speech stimulus and disappeared when the participants pressed any of two keys. The reaction time was defined as the interval between the probe onset and the key press. The next stimulus started 1.167/1.262/1.357/1.452/1.547/1.642/1.737/1.833 sec after the termination of probe stimulus. A white fixation cross was always presented at the center of the screen during the experiment. The EEG experiment consisted of 4 sessions, and the simultaneous EEG-fMRI experiment consisted of 3 sessions. Each session consisted of 100 trials.

2.2.3 EEG recording

The EEG were recorded by a 32-channel EEG amplifier (BrainVision MR, Brain Products, Germany) with an international 10 % standard electrode cap with sintered Ag/AgCl ring electrode (Easy Cap, Falk Minow Services Germany). Four electrodes were used for the vertical and horizontal electrooculogram (VEOG and HEOG) channels. The measurement reference was linked ear-lobes, and the ground was on theinion. The sampling rate of EEG was 5 kHz with a 1-Hz high-pass software filter, a 250-Hz low-pass hardware filter, and a 60-Hz notch filter. The impedance was set to be below approximately 15 k Ω for all electrodes. All EEG experiments were done in a soundproofed room.

2.2.4 Simultaneous EEG-fMRI recording

An MR compatible amplifier (Brain Amp MR, Brain Products, Germany) was used to acquire the EEG during fMRI measurements using a 10 % standard system electrode cap with Ag/AgCl sensor electrodes (BrainCap MR, BrainProducts, Germany). The EEG cap had 27 EEG electrodes, one electrocardiogram (ECG) electrode, and four vertical and horizontal electrooculogram (VEOG and HEOG) electrodes. An FCz electrode, located between Fz and Cz, was used as the measurement reference, while an electrode on theinion was used as the measurement ground. The ECG electrode was set at the participant’s back. The sampling rate for the EEG was 5 kHz with a 1-Hz high-pass software filter, a 250-Hz low-pass hardware filter, and a 60-Hz notch filter. The impedance was set to be below approximately 15 k Ω or all electrodes. Blood oxygenation-sensitive echo-planar images (EPI) were obtained using a 3-tesla MR scanner (Magnetom Verio, Siemens, Germany) and were taken simultaneously with the EEG measurements. The participant’s head was immobilized using a vacuum pad during the measurements. The fMRI measurement settings were as follows: repetition time (TR) = 2 sec, echo time (TE) = 30 msec, flip angle (FA) = 80 deg, field of view (FoV) = 192 mm, matrix size = 64×64 , slice thickness = 5 mm, gap = 0 mm, 333 volumes, and 3 session. Prior to wearing the EEG cap, a structure image for all participants was also acquired by the magnetization prepared rapid gradient echo (MPRAGE) with the following settings: TR = 2.25 sec, TE = 3.06 msec, FA = 9 deg, FoV = 256 mm, matrix size = 256×256 , slice thickness = 1 mm, gap = 0 mm.

2.2.5 Preprocessing

Ocular artifacts in the EEG were corrected using an EEG analysis software (Brain Vision Analyzer, Brain Products, Germany), by subtracting the voltages of the VEOG and HEOG channels, multiplied by a channel-dependent correction factor calculated by linear regression among VEOG/HEOG and EEG electrodes [30]. The reference of the ocular-corrected EEG was then changed to the average of all electrodes except

the VEOG and HEOG. For the EEG data simultaneously recorded with the fMRI, MR-scan-related and ECG-related pulse artifacts were removed by the EEG analysis software (Brain Vision Analyzer, Brain Products, Germany) before the ocular artifact correction. Briefly, the waveforms for the MR-scan-related and ECG-related artifacts were identified by averaging the EEG data derived from artifact periods and were subtracted from the EEG data [31, 32]. The preprocessed EEG data were then downsampled into 500 Hz in the EEG experiment data and into 200 Hz in the simultaneous EEG-fMRI experiment data, and exported to Matlab (Mathworks Inc., USA). The EEG data were transformed into time-frequency representations of EEG power and phase from 1 to 60 Hz in 50 logarithmically-spaced steps, using the Morlet wavelet. The mother wavelet is defined by the following equation [33]:

$$w(t, f) = (\sigma_t \sqrt{\pi})^{-1/2} \exp(-t^2/(2\sigma_t^2)) \exp(i2\pi ft), \quad (2.1)$$

where t is time, f is frequency, and σ_f is the variance at the frequency f . The variance σ_f is expressed by σ_t , which is the variance at time t , as $\sigma_f = 1/(2\pi\sigma_t)$. The wavelet is characterized by the ratio of the frequency to variance f/σ_f and was set to $f/\sigma_f = 7$ in the current study. By convolving the mother wavelet (Eq. 1) to the EEG time series $s_c(t)$, the EEG power $p_c(t, f)$ and phase $\theta_c(t, f)$ at electrode c were computed by the following equation:

$$p_c(t, f) = \log_{10}(|w(t, f) * s_c(t)|^2) \quad (2.2)$$

$$e^{i\theta_c(t, f)} = w(t, f) * s_c(t) / |w(t, f) * s_c(t)|. \quad (2.3)$$

The $*$ indicates the convolution operator. The time series $s_c(t)$ was the segmented period from 2 sec before to 1.5 sec after the speech onset for the EEG analysis, while it contained the whole length of the EEG data in each session for the simultaneous EEG-fMRI analysis. Because of the computational problem, the EEG power and phase were downsampled into 100 Hz for phase-utility on behavior (PUB) computation.

2.2.6 EEG analysis

Phase-utility on behavior

We proposed the PUB index to evaluate how much behavioral performance depends on the ongoing EEG phase. Behavioral performance (accuracy or reaction time) were sorted by instantaneous phase and fitted by the cosine-curve with the least square method (Figure 2-1b). The accuracy was derived from the segmentation with eight bins of EEG phase ($y_i, i = 1, 2, \dots, 8$), and the reaction time was transformed into logarithmic scale ($y_i = \log_{10}[RT(ms)], i = 1, 2, \dots, \#trials$). The independent variable was the instantaneous EEG phase ($\theta_i, i = 1, 2, \dots, \#trials$) in the least square method. Thus, the computation of the PUB is defined by the following equation:

$$\mathbf{y} = A_c(t, f) \cos(\boldsymbol{\theta}_c(t, f) + \tau_c(t, f)) + B_c(t, f) + \boldsymbol{\xi}_c(t, f), \quad (2.4)$$

where c is the electrode, t is time, f is frequency, A is the regression coefficient, B is the intercept, and τ is the parameter of shifting the cosine function. The constant $\tau (0 \leq \tau \leq \pi)$ was decided to minimize the sum of residuals ξ . The amplitude $|A|$ is defined as the PUB at the optimal τ . A larger PUB means a larger dependence of behavioral performance on the EEG phase and vice versa. Based on the constant τ , we can compute the optimal and non-optimal phase showing the best and worst behavioral performance. When A is a positive value, the optimal phase is $\pi - \tau$. When A is negative value, the optimal phase is $-\tau$. Note that the index can be computed appropriately only when the phase is uniformly distributed among the trials, since the dependent variable must be explained by the phases ranging from 0 to 2π in the least square method. If the EEG phase is locked among the trials, it decreases the accuracy of the PUB computation. We computed the PUB by using all trials except those in which the participants failed to press a key within 2 sec. The number of remaining trials was approximately 305.9 ± 7.6 (mean $\pm$ s.e.m. among fourteen participants). When we computed the PUB using only the correct trials, the results were similar to the phase-dependent behavioral modulation that was computed using

all the trials. Therefore, we decided not to exclude the incorrect trials in order to get better statistical power for the PUB computation, which required larger numbers of trials for the analysis.

Statistical test

For the statistical analysis, the t-test between the PUB and surrogate data was used to test whether the behavior performance depended on the EEG phase. The surrogate data for the PUB analysis were computed using the random sampling data of the behavioral performance instead of the performance vector $\mathbf{y}$ in Eq. (2.4) (Figures 2-2a and 2-2b). The random sampling was repeatedly computed 200 times, and the median values of these data were used as the surrogate data for each participant. The distribution of these data were non-normal and asymmetrical in form, since the PUB was computed as the absolute value, leading us to use the median of the PUB distribution for the surrogate data. The t-values were calculated by two sample t-tests between the original PUB and surrogate PUB. To avoid the multiple comparison problem, the statistical threshold was decided by a nonparametric statistical test derived from a cluster-based Monte Carlo p-value [34]. The PUB clusters were decided by the temporal, spectral, and spatial adjacency of the samples that exceeded a student's t-value ($p < 0.001$) (Figure 2-2c). The cluster-level statistics were the summation of the t-values within a cluster. To decide the cluster-level threshold, the original PUB and surrogate PUB were pooled into one group and randomly divided into two groups. Then, the cluster-level statistics were decided in the same manner as above (Figure 2-2d). The random shuffling of the original PUB and surrogate PUB was repeated 10,000 times, and the maximum cluster-level statistic was computed to obtain the empirical distribution. The statistical threshold was decided by the empirical distribution of the maximum cluster-level statistics ($p < 0.05$). We excluded clusters showing no spatial expansion, and the extent of the threshold was < 2 electrodes.

In order to confirm the effect of the EEG phase, we used an ANOVA analysis to compare the reaction times between the trials of the optimal and non-optimal phases. There is a risk of double dipping when the data point of the EEG phase is

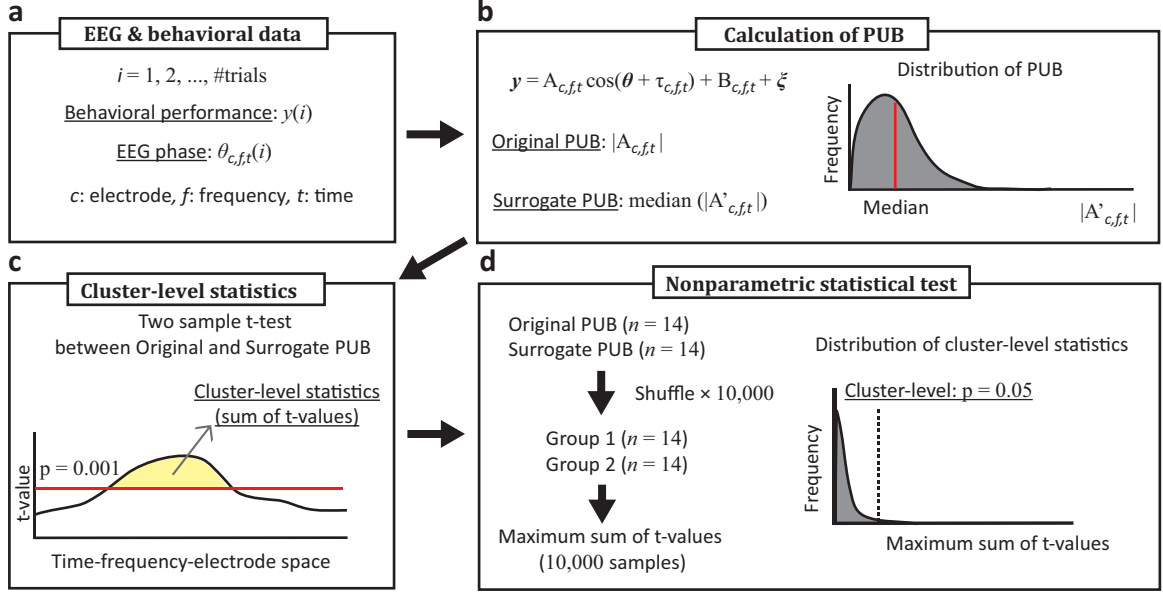


Figure 2-2: Schematic illustration of the cluster-level statistical test. (a) The EEG was decomposed into instantaneous phases. (b) The original PUB and surrogate PUB were calculated from the EEG phases and behavioral performance data. (c) Clusters in the time-frequency-electrode space were decided based on the threshold of the t-value, and the cluster-level statistics were computed. (d) The cluster-level statistics were tested by the empirical distribution of the maximum cluster-level statistics.

simply decided by the t-value, which is calculated with all the participants. To avoid the double dipping problem, we used the leave-one-participant-out cross-validation procedure, which gave a data point for each participant that could be used in the ANOVA [35]. The t-values were calculated by a two-sample t-test among $N-1$ participants, where N was the number of participants. The data point of the EEG phase for one participant was, thus, decided by the t-value from the remaining participants. The t-values calculated with other participants were independent from the data point for one participant. The clusters were then decided by the temporal, spectral, and spatial adjacency of the samples that exceeded the t-value ($p < 0.01$). The clusters corresponding to the clusters in Figure 2-3 were decided by eye inspection for each participant. If the clusters derived from the leave-one-participant-out procedure are not isolated from the other clusters, the clusters are manually separated based on the time-frequency-electrode space between the clusters in Figure 2-3. We used the data points that had the maximum t-value in each cluster for the individual participant.

The trials were classified into six groups based on the estimated optimal phase by the PUB analysis. The averaged logarithmic reaction times from the trials in the EEG phase groups were compared each other using an ANOVA analysis.

2.2.7 Simultaneous EEG-fMRI analysis

SPM12 software (Wellcome Department of Cognitive Neurology, London, UK, URL: www.fil.ion.ucl.ac.uk/spm) was used for the image preprocessing and voxel-based statistical analysis of the fMRI data. The first three volumes were discarded, and slice timing was corrected on the middle slice to remove the time delay from scanning the entire brain. The echo-planar images (EPIs; 330 volumes for each session $\times$ 3 sessions) were transferred into the first image volume for each participant to correct for head motion. The individual EPIs were normalized to a standard brain by applying the parameters estimated by matching the MPRAGE anatomical image to the stereotactic coordinate image from the Montreal Neurological Institute. The EPIs were then smoothed with a 10-mm full-width half-maximum Gaussian kernel. The following general linear model was used to identify the cortical areas that correlated with the EEG activity:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\eta}, \quad (2.5)$$

where $\mathbf{X}$ is the design matrix consisting of the EEG power, task-related model, and the six head-motion parameters, $\mathbf{Y}$ is the vector of fMRI BOLD signals, $\boldsymbol{\beta}$ is the regression coefficient vector, and $\boldsymbol{\eta}$ is the residual vector. The design matrix for the EEG activity was computed by the convolution of the canonical hemodynamic response function and the EEG power time series to bridge the temporal gap between the fMRI and EEG data, and then subsampled at the time of scanning the fMRI measurements [36, 37, 38]. The EEG power time series was z-transformed before the convolution of it with the hemodynamic response function. The task-related model was computed by convolving the hemodynamics response function with the probe presentation duration to reduce the effect of the visual presentation. The regression coefficients for the EEG powers were computed for each participant using the fixed

effect model, and then taken into the group analysis using a random-effect model of a one-sample t-test.

2.3 Results

2.3.1 Changing behavioral performance based on EEG phase

To test whether behavioral performance in speech recognition changed based on the ongoing EEG phase, we evaluated the relationship between the accuracy, reaction time and EEG phase by using the proposed PUB index. The PUB analysis was applied to the EEG experiment but not to the simultaneous EEG-fMRI analysis because the EEG data from the simultaneous recording was insufficient to analyze the details of the EEG phase due to the MRI-related artifacts. The delta and theta activities showed significant PUB modulation in the reaction time analysis (Figures 2-3 and 2-4). The comparison of the grand averaged PUB and surrogate showed theta activity (4–5 Hz) just before the speech presentation and delta activity (1–2 Hz) around the onset of speech (Figure 2-3a). We confirmed that these activities were statistically significant using a cluster-level test that was based on temporal, spectral, and electrode adjacency (cluster level $p < 0.05$; Figure 2-3b). The topographies showed that theta activity appeared around the left parietal electrodes (P3 and CP1) (Figure 2-3c-1). The delta activity was spatially divided into two clusters: right occipito-parietal electrodes (P4 and O2) (Figure 2-3c-2), and fronto-central electrodes (Fz, Cz, FC1, FC2, C4, CP2 and FCz) (Figure 2-3c-3). To confirm the consistency of the optimal phase among the participants, a selected point showed maximum PUB value within each cluster, and the phase consistency at this point was verified by the Rayleigh test (Figure 2-3d). The results showed significant phase consistency on the fronto-central delta oscillation ($p < 0.05$) (Figure 2-3d-3). The left parietal theta and occipito-parietal delta activities were not consistency among all participants ($p > 0.05$) (Figure 2-3d-1 and 2-3d-2). The accuracy analysis showed no statistically significant results (cluster level $p > 0.05$). Therefore, further analyses focused on the

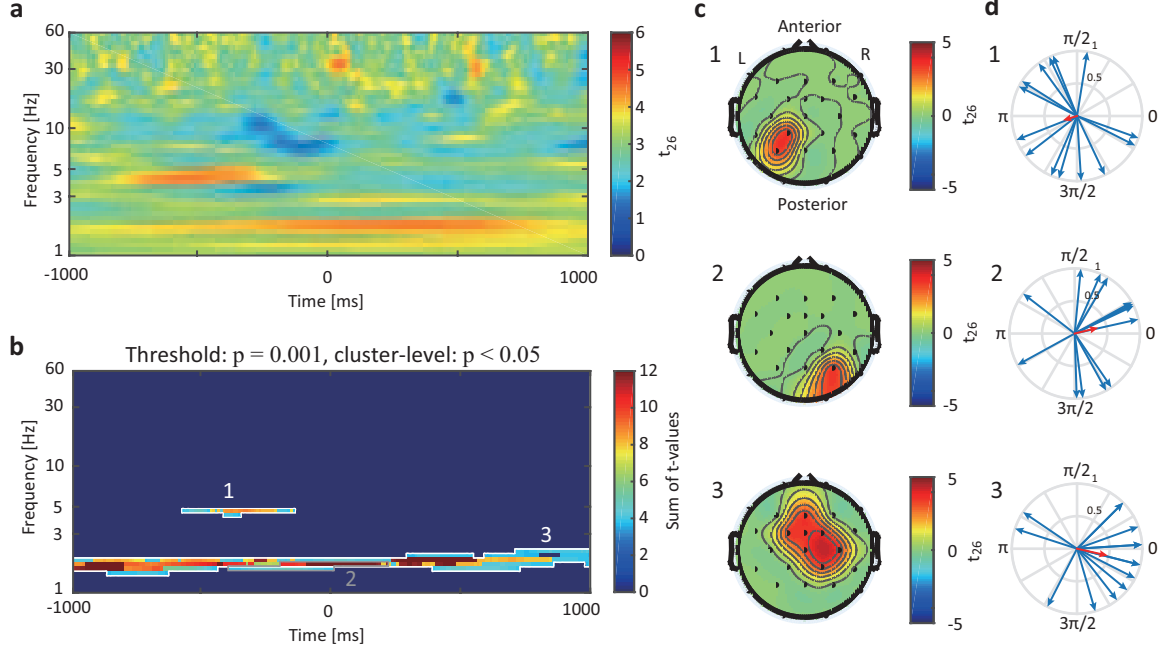


Figure 2-3: Phase-utility on behavior (PUB) result. (a) Statistical t-map of the comparison between averaged PUB across all electrodes and surrogate data. The color bar represents the t-value of the two-sample t-test ($n = 14$). Larger t-values mean the PUB increased relative to the surrogate. (b) Cluster-level analysis of the PUB and surrogate comparison based on the temporal, spectral, and electrode adjacency. The three areas indicated by the numbers 1–3 in the figure are the significant clusters (cluster level $p < 0.05$). For presentation purposes, the color bar represents the summation of t-values among the electrodes, while the significant clusters were decided through a summation of the temporal, spectral, and electrodes spaces. (c) The scalp topographies show the average t-values among the temporal and spectral spaces of each cluster. The numbers 1–3 corresponded to the cluster numbers in (b). (d) Individual optimal phase (blue arrows) and circular mean among all participants (red arrow; top: CP1 theta; middle: P4 delta; bottom: FC2 delta). The position of time-frequency electrode was chosen to show maximal PUB value within each cluster for this circular mean analysis. The Rayleigh test for uniformity was applied in each of the optimal phases (CP1 theta, $Z(14) = 0.7207$, $p > 0.05$; P4 delta, $Z(14) = 1.9491$, $p > 0.05$; FC2 delta, $Z(14) = 3.2771$, $p < 0.05$).

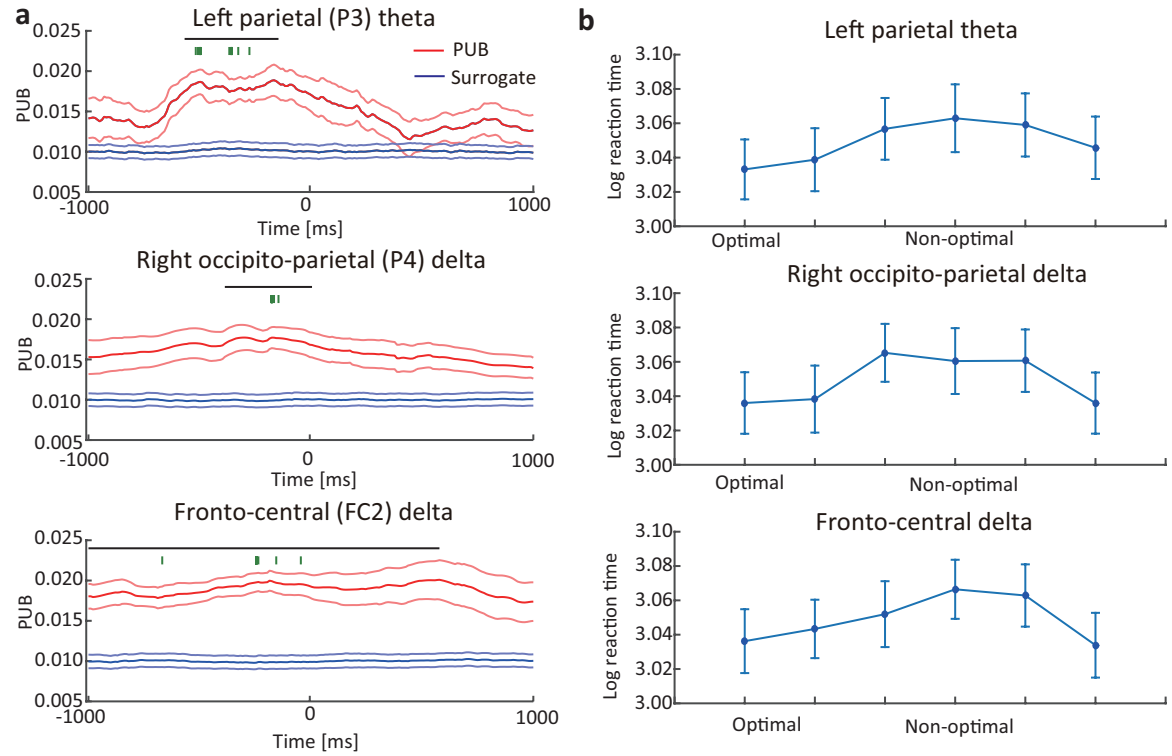


Figure 2-4: The effect of the phase on reaction time. (a) The PUB (red line with $\pm$ s.e.m.) and surrogate value (blue line $\pm$ s.e.m.) were averaged among the participants for the P3 (top), P4 (middle), and FC2 (bottom) electrodes. The black horizontal bars on each panel represent the period when a significant difference occurred (cluster level $p < 0.05$). Green vertical bars indicate the time point for the ANOVA analysis selected by the leave-one-participant-out cross-validation statistics. (b) Logarithmic reaction time as a function of phase bins derived from the PUB analysis. The trials were classified into six groups based on the optimal phase estimated by the PUB analysis. The blue plots and vertical bars represent the average log-reaction times and their s.e.m.

reaction time.

To confirm whether the optimal/non-optimal phase derived from the PUB analysis showed a behavioral difference, we compared the reaction times between the optimal and non-optimal phases. To avoid double dipping, the data point of the time-frequency-electrode space for the comparison was individually decided using the leave-one-participant-out cross-validation procedure [35]. The data points were the maximum t-value in the space of each cluster in each leave-one-participant-out group for each participant (Figure 2-4a). Based on the optimal/non-optimal phase, we classified all trials into the six bins. The ANOVA analysis showed the main effect of the reaction time among these bins (left parietal theta, $F(5,65) = 6.31$, $p < 0.001$; right occipito-parietal delta, $F(5,65) = 12.198$, $p < 0.001$; fronto-central delta, $F(5,65) = 10.538$, $p < 0.001$) (Figure 2-4b). The results indicate that our method can successfully identify the EEG activity whose phase modulated the reaction times.

2.3.2 Brain areas involved in PUB activity

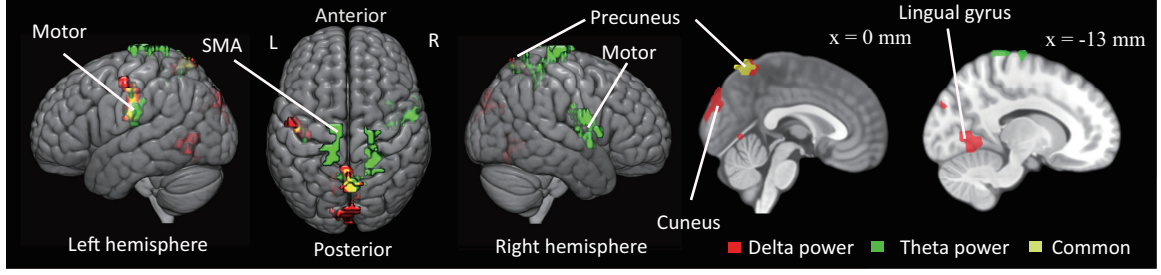


Figure 2-5: Brain areas for EEG activity. The brain areas showing correlations with the left parietal theta and fronto-central delta power fluctuation. The red areas indicate a significant correlation between the delta power and the BOLD time series (cluster-level FWE-corrected $p < 0.05$). The green areas indicate a significant correlation with theta power. The yellow areas indicate common areas that showed correlations with delta and theta power.

We performed a simultaneous EEG-fMRI analysis to identify the brain areas associated with the EEG that also showed PUB modulation. The EEG and fMRI were combined through a multiple regression analysis. The EEG regressor, which was simultaneously recorded with the fMRI, was composed of a power of theta or delta oscillations whose electrode and frequency were decided based on the PUB results (see Figure 2-3c; the left parietal theta regressor was the mean theta power between the P3 and CP1 electrodes; the right occipito-parietal delta regressor was the mean delta power between the P4 and O2 electrodes; the fronto-central delta regressor was the mean delta power among the Cz, FC1, FC2, C4, CP2, and FCz electrodes). We performed a voxel-based statistical analysis of the fMRI BOLD images to identify the cortical areas of the theta and delta bands activities (cluster-level FWE-corrected $p < 0.05$). The left parietal theta regressor was computed by averaging the EEG powers of the P3 and CP1 electrodes at a frequency of $\sim 4.5\text{Hz}$ (Figure 2-3c-1). The theta power was positively correlated with the activities in the bilateral motor cortices, supplementary motor area (SMA), and precuneus (Table 2.1; Figure 2-5). The fronto-central delta regressor was computed by averaging the EEG power of the Fz,

Cz, FC1, FC2, C4, and CP2 electrodes at a frequency of $\sim 1.7\text{Hz}$ (Figure 2-3c-3). The delta power was positively correlated with the activities in the left motor cortex, cuneus, precuneus, and lingual gyrus (Table 2.1; Figure 2-5). The left-parietal delta regressor was derived from the P4 and O2 electrodes at a frequency of $\sim 1.5\text{ Hz}$ (Figure 2-3c-2) and showed no significant response.

Table 2.1: Brain areas showing correlations with EEG activities.

EEG activity	Anatomical region	MNI coordinates (mm)			t-value
		x	y	z	
Left parietal theta	l-Motor (BA4)	-39	-13	25	8.09
	r-Motor (BA4)	57	-1	20	9.15
	r-Precuneus (BA7)	6	-69	65	8.12
	r-SMA (BA5)	18	-49	65	7.34
	l-SMA (BA6)	-15	-16	80	7.07
Right occipito-parietal delta	None				
Fronto-central delta	l-Motor (BA4)	-45	-10	50	9.86
	Lingual gyrus (BA19)	-12	-70	5	9.77
	Cuneus (BA10/9)	0	-97	20	7.67
	r-Precuneus (BA7)	3	-67	65	5.89

(SMA: supplementary motor area; BA: Brodmann area; MNI: Montreal Neurological Institute; l-: left; r-: right)

2.4 Discussion

2.4.1 Delta and theta oscillations in speech recognition

We evaluated the correlation between the phase of ongoing EEG activity and behavioral performance by PUB analysis. The ongoing EEG phase modulates neural excitability and perceptual performance during the visual/auditory perception tasks, even if it appeared prior to the stimulus onset [2, 6, 7]. In line with previous studies, we successfully found the EEG phase accompanying the behavioral modulation dur-

ing a speech recognition task. The PUB index can continuously evaluate the state of phase-dependency of behavioral performance, while previous method evaluates discretely the states [7]. In addition to the previous method, the circular mean can compute the similar value with PUB, where the reaction time and the EEG phase are defined as the length and angle of vectors, respectively. When distribution of EEG phase is uniform, the length of circular mean vector indicates the phase-dependency of the reaction time. However, the length of circular mean vector increases irrespective of behavioral performance when the EEG phase is locked. In contrast, the PUB value is close to zero on the phase locking condition. However, the PUB also has a problem that the mean value depends on the numbers of trials as well as the circular mean. Therefore, it is necessary to compare the original PUB to surrogate PUB to avoid the problem of numbers of trials. The number of trials of surrogate PUB was same as original PUB. The comparison can thus identify the phase-dependency of the behavior performance without the problem of number of trials. It might enable us more sensitive identification of the phase and behavior relationship than the previous method and the circular mean. Our PUB result demonstrates that the reaction time is modulated depending on the theta phase prior to the speech onset. It is also reported that the theta oscillatory phase effects the behavioral performance in a sound detection task [6]. Our PUB method clearly found that the ongoing theta oscillation is correlated with the performance of speech recognition. It implies that the theta phase influence appears not only in perceptual performance, but also in recognition performance.

In addition to the theta oscillation, we found the delta oscillation modulating the reaction time. The front-central delta activity was the only oscillation to show the consistency of optimal phase among all trials (Figure 2-3d-3). It is known that the delta and theta oscillations are crucial components in vocal communication [5, 9, 39]. The speech sound has rhythms of the prosody and syllable. The prosody is the stress and intonation patterns of an utterance. The envelope of prosodic rhythms appears in the 1-3Hz frequency range. The syllable is a unit of speech to separate a word into sound chunks. For example, the word “oscillation” is composed of four

syllables: “os-cil-la-tion” and its sound envelope appears in the 4-8 Hz frequency range in vocal communication. The frequency of the delta oscillation is identical with the frequency of the prosodic rhythms, and the theta oscillation corresponds to the syllabic rhythms. These speech rhythms are synchronized with the listener’s EEG oscillations in the delta or theta frequency [12, 39]. The synchronization is stronger when the speech is recognized, in contrast to when the speech is presented backward. Together, our finding of delta and theta activities is consistent with previous studies on how it modulates behavioral performance with these phases in speech recognition.

2.4.2 Motor contribution in speech recognition

We showed that the delta and theta EEG phases contribute not only to sound detection performance, but also to speech recognition performance. We then asked how the EEG phase effects behavior performance. The fronto-central delta and left parietal theta powers were correlated with the BOLD activities in the motor cortices in the simultaneous EEG-fMRI experiment. However, the scalp distribution of EEG oscillations is seemingly unrelated with cortical areas. In the simultaneous EEG-fMRI analysis, we computed correlation of EEG power fluctuation and BOLD fluctuation. If a cortical activity shows high temporal correlation to the EEG power, it is detected as a part of related cortices. In this case, the results include not only source (generator) of the scalp-EEG activity, but also activities that have temporal correlation to the real source. Therefore, the seemingly unrelated cortical activity may be detected as showing the temporal correlation to the scalp-EEG power. Only from this simultaneous EEG-fMRI result, we are unable to conclude which cortex is real “source” of the scalp-EEG activity. Here we are just able to consider that the cortical activities, those appeared in the simultaneous EEG-fMRI results, were “associated” with the scalp EEG-activity. Our result showed that the motor activity were clearly found in the simultaneous EEG-fMRI result, and were consistent with a previous study [15]. The results provide important information on the functional relevance of fronto-central delta and left parietal theta oscillation in cooperating with

the motor cortex.

Motor cortical activity contributes to speech perception (e.g., motor theory of speech perception) [24, 25, 40]. Speech intelligibility is enhanced when the listener can identify the speaker’s motor signal for speech production. While viewing a speaker’s face or listening to a speech, the muscle activity for lip movement increases [29]. The neural activities in the motor cortex for the lip and tongue movement enhances speech comprehension in a TMS study [23]. Especially, the motor cortex activity effects phonemic identification during listening speech sound with noise [23, 26, 27, 28]. The motor cortex is not necessary for speech recognition in general but may help compensating a distorted speech sound [41]. In our EEG and simultaneous EEG-fMRI experiments, speech stimuli were also embedded in noise sound. In these experimental conditions, the degraded speech sound must be compensated by viewing the speaker’s facial movement. Together with the previous study, our simultaneous EEG-fMRI result supported the contribution of motor cortex activity in speech recognition. The spatial location of the motor activity in our simultaneous EEG-fMRI results was consistent with the spatial coordinate for the lip movement [42]. It suggested that the delta and theta activities corresponds with the activity in the motor cortex for the lip movement.

The delta and theta oscillation in the motor cortex plays a role in top-down control for auditory activity in continuous speech recognition [15]. The auditory delta and theta oscillations are phase-synchronized with the rhythms of the speech envelope during vocal communication. The top-down signal increases the synchronization between the neural oscillation and the speech envelope to enhance speech intelligibility. It implies that the delta and theta oscillations in the motor cortex are a crucial component for enhancing speech recognition. The finding of the delta and theta phases in the current study is in line with the previous study. The delta and theta phases in the PUB analysis may be associated with the top-down signal in speech recognition.

Note that one may be concerned with the possibility of motor activity for finger movements in association with the participant response. In fact, the theta activity was related to the activity in the SMA during a motor executive task [43]. We

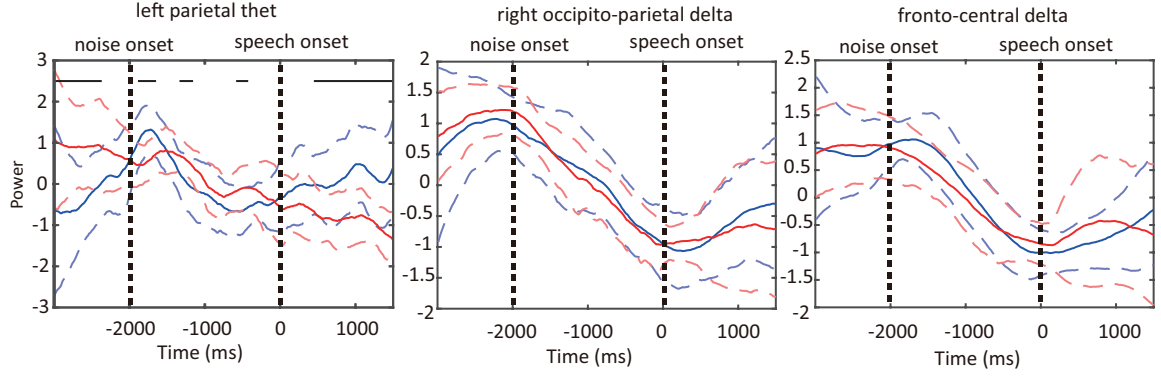


Figure 2-6: EEG power modulation in association with the noise and speech onsets. The EEG power for the EEG experiment (blue line) and for simultaneous EEG-fMRI (red line). The power was averaged among participants (mean $\pm$ s.e.m.) for left parietal theta (left), right occipito-parietal delta (middle) and fronto-central delta (right). The black horizontal bars on each panel represent the period showing the significant difference ($p < 0.05$).

found the theta power modulation in association with the participant response in the current study (data not shown). The motor activity enhances the movement speed in cooperation with distributed brain activities via delta synchronization [44]. The delta oscillation was, in fact, accompanied by the activity in the motor cortex together with the activities in the cuneus, precuneus, and lingual gyrus in our result. These results may raise the risk that the enhancement of reaction time was merely induced by speeding up of the participant response. However, the spatial location of the motor activity was found in the region for lip movement, but not for finger movement. In addition to the location of motor activity, the paramidline theta activity was related to speech recognition[15]. Taken together, our results supports the idea that motor activity is not only related to the listener's finger response but also with speech recognition, while the possibility of contamination from the finger response cannot be completely denied in the activity in the motor cortex and SMA.

Another concern is whether the EEG activities were identical between the EEG experiment and the simultaneous EEG-fMRI. The electrode position and frequency for the simultaneous EEG-fMRI analysis was decided by the result of EEG experiment without fMRI. Here we assumed that the EEG activities, those were used for the simultaneous EEG-fMRI analysis, were identical to the activities showing the phase-

dependency of behavioral modulation in the EEG experiment. The accuracies were identical, while sound noise level was different between these two experiments. The temporal profile of EEG power modulation was also identical in association with the fronto-central and occipito-parietal delta activities (Figures 2-6b and 2-6c). Note that the left parietal theta activity showed different temporal profile of EEG power modulation between the EEG experiment and simultaneous EEG-fMRI (Figure 2-6a)($p < 0.05$). We assume that this difference of EEG power modulation may be caused by the different experimental environment. In the MRI scanner, sound noise always occurred during the MRI recording, while there was no sound input during the inter-trial interval in the EEG experiment. The theta activity is an important component to sound input, and showed the EEG power modulation by the MRI sound noise [45]. MRI sound noise (before the noise onset in Fig 2-6a) might modulate the theta EEG power, resulting in the different temporal profile of the power at the left parietal electrode between the EEG experiment and simultaneous EEG-fMRI. Therefore, the different temporal modulation of the theta EEG power dose not denied our assumption that the activity in the simultaneous EEG-fMRI was identical to the EEG which was found in the EEG experiment. These results strongly supported our assumption that the results of cortical activities in the simultaneous EEG-fMRI analysis tightly correlated with the delta activities that showed the phase dependent behavior modulation, while the conclusion for theta activity is still remained for further study.

2.4.3 Prediction for sound input

Our results suggeste that speech recognition is enhanced with the identification of motor activity for speaker’s speech production by the delta and theta oscillation. On the other hand, the delta and theta phases effect not only appears in speech recognition, but also in the simple tone detection [6, 20]. It raises the question of whether the phase of delta and theta oscillations identifies the speaker’s speech production in speech recognition, or purely predicts the auditory input timing in sound detection. One possible interpretation of our results is that the selective attention in time

for sensory inputs enhances performance of speech recognition with low frequency oscillations [5].

In sensory perception, low-frequency oscillations were associated with selective attention or prediction for external sensory inputs [6, 20, 46, 47]. Furthermore, a previous study reported that the prediction was realized by the neural synchronization in the motor cortex [48]. Also in speech recognition, the delta and theta oscillations are considered a predictive function for speech input. Speech recognition can be enhanced by a prediction for speech input timing through viewing a speaker’s face [5]. The speaker’s facial movements coincide with the prosodic rhythm in the speech. The frequency of prosody is known to appear in the delta range (1-3Hz). In addition to the delta oscillation effect in speech recognition, the theta oscillation encodes the visual feature particularly on mouth shape [49]. The theta oscillation is also an important rhythm in speech processing (i.e., syllabic rhythm) [9]. These delta and theta oscillations may be synchronized with the head/mouth movements, resulting in the prediction to the auditory input to enhance speech recognition [5]. The delta and theta oscillations in the current study, therefore, may be associated with the prediction to the auditory input timing, resulting in the enhancement of speech intelligibility.

Another possible interpretation of the behavioral modulation by delta/theta oscillatory phase is articulatory prediction by the motor cortex. Neural excitability can be enhanced at an optimal phase, and suppressed at a non-optimal phase of the slow neural oscillation. Our simultaneous EEG-fMRI showed that the slow neural oscillation is related with the activity in the motor cortex. The motor activity is related to the recognition at the phonemic level, and generates the articulatory prediction during the phoneme discrimination task [23, 26, 27, 28, 41]. In the current study, the target and distractor words differed only at one phonemic level. The phonemic discrimination can be accelerated by enhancing the neural activity in the motor cortex for tongue or lip motion when the neural oscillation comes at an optimal phase. The prediction of auditory input would be the neural mechanism of behavioral modulation by the slow neural oscillation in any way, while it is still remained for future works whether the prediction acts on the timing or articulation of the speech sounds.

Our study provides a new index to quantitatively evaluate the EEG phase effect on behavior performance, and indicates that the delta and theta phases modulated reaction times in speech recognition. The simultaneous EEG-fMRI analysis shows the motor cortical activity is correlated with the delta and theta oscillations. These results suggests that the phase-dependency of speech intelligibility is associated with motor contribution for speech recognition.

Chapter 3

Estimation of interaction of oscillators for cross-frequency phase synchronization between EEG and speech sound

This chapter is constructed based on the published paper [17].

3.1 Introduction

Synchronization of neural oscillations is considered an important activity that can help reveal the mechanisms underlying various cognitive functions. Neural oscillation is a rhythmic neural activity and is usually observed by EEG. Neural oscillations are classified into a few frequency bands (e.g. delta, theta and alpha frequency bands) and are synchronized within the same-frequency band between different brain areas during various cognitive tasks [50, 51, 52, 53]. Synchronization of oscillations of the same frequency is considered to integrate distributed brain activities [54] and regulate communication between distant neural groups [11, 55].

Synchronization between slow and fast oscillations (cross-frequency synchronization) also appears during a few cognitive tasks [56, 57, 58, 59]. In particular, $1 : p$

phase synchronizations (p is an integer) can be observed in the resting state, mental arithmetic tasks, and working memory tasks [60, 61, 62, 63, 64, 65], and may integrate activities over different time scales [66]. $1 : p$ phase synchronization refers to phase locking of a single cycle of one oscillation to p cycles of the other oscillation. Although $1 : p$ phase synchronization is considered important from the perspective of brain function, to the best of our knowledge, there is no effective and practical method to analyze the $1 : p$ phase synchronization mechanism.

Various methods to identify this synchronization have been used in EEG studies. For example, the phase locking index is used frequently to identify phase synchronization. This index measures the temporal consistency or intertrial variability of the phase difference between different brain areas or cross-frequency oscillations [67, 50, 68, 69, 70]. In addition, the directional connectivity between neural oscillations has been evaluated in terms of transfer entropy [14, 15]. Transfer entropy evaluates the directed transfer of information between two random processes. Many previous studies have examined the roles of neural oscillation using these methods. However, these methods could not reveal how neural synchronization is achieved by a dynamical system. Therefore, we have developed a method to identify a dynamical system for synchronization.

It is widely believed that the dynamical system of EEG activity can be described by the neurophysiological model of a cortical column [71, 72]. If this dynamical system can be explained by a weakly coupled oscillator, the corresponding neurophysiological model can be described using the phase oscillator model in which each oscillator is described by only one variable, i.e., the phase [73]. Some previous studies have provided estimation methods to derive the phase oscillator model directly from time-series data without detailed modeling [74, 75, 76, 77, 78, 79, 80]. However, such methods cannot be applied to cross-frequency synchronization data. Therefore, we extend previous methods to make them applicable to $1 : p$ phase synchronization.

In this paper, we describe an extended method to explain $1 : p$ phase synchronization based on the phase oscillator model and verify the reliability of the estimation method through numerical simulation and an electronic circuit experiment. Then, we

apply the proposed method to EEG oscillation and speech sound. Speech rhythms are synchronized with neural activity in a listener’s brain [5], and speech rhythm consists of a few important linguistic components (e.g., syllable and prosody). It is believed that synchronization between neural oscillation and linguistic rhythm contributes to parsing continuous speech [45] and predicting the timing of linguistic component production [5, 16]. Furthermore, the causality between EEG activity and speech sound is clear, whereas the causality among neural activities is generally unknown in advance. Therefore, we estimate interactions between EEG activity and speech sound to confirm the validity of the estimation results and demonstrate that the proposed method can successfully estimate the dynamical system based on EEG data.

3.2 Materials and methods

3.2.1 Estimation of phase coupling functions for cross-frequency synchronization

Neural oscillations can be observed easily from EEG data, and many EEG studies have reported various types of synchronization [81, 57], which can be roughly divided into same- and cross-frequency synchronization. A few experimental results suggest that same-frequency phase-phase synchronization plays a role in modulating neuronal interaction [11, 55]. In contrast, cross-frequency synchronization is considered to play a role in the integration of activities over different time scales [66]. However, it is unclear how these synchronizations, particularly $1 : p$ phase synchronization, are achieved by a dynamical system. Therefore, we develop an effective method to identify the dynamical system that performs these synchronizations.

In general, synchronization of neural oscillation is thought to be described by a network of limit-cycle oscillators, which can be described generally by the multidimensional differential equation $\frac{d\mathbf{X}_i}{dt} = \mathbf{F}_i(\mathbf{X}_i) + \sum_{j \neq i}^N \mathbf{G}_{ij}(\mathbf{X}_i, \mathbf{X}_j)$, where $\mathbf{X}_i$ denotes the multidimensional state of the i -th oscillator, such as membrane voltages and gate variables of ionic channel. We assume that a system $\mathbf{X}_i$ can generate a limit-cycle

oscillation by itself without external interaction. An EEG signal is thought to be generated by some neuronal system consisting of many interacting neurons. In this context, it is plausible that the neuronal system of an EEG signal can be represented by the $\mathbf{X}_i$ system. According to the phase reduction method, the limit-cycle oscillator can be characterized theoretically by a phase ϕ as a simple dynamical system with one degree of freedom. If the oscillators are weakly coupled, the dynamics of the networks among N oscillatory systems can be described by [82, 73]:

$$\frac{d\phi_i}{dt} = \omega_i + \sum_{j \neq i}^N \Gamma_{i,j}(\phi_j - \phi_i), \quad (3.1)$$

where ω_i is the natural frequency of the oscillator and Γ_{ij} is a phase coupling function representing the influence from the j -th oscillator to the i -th oscillator. It is known theoretically that the phase coupling function depends only on the phase difference $\phi_j - \phi_i$. When the phase difference is constant over time, these oscillators are said to be synchronized. Specifically, the synchronization of same-frequency oscillators is referred to as 1 : 1 phase locking. Equation (3.1) can describe the 1 : 1 phase-locking state between rhythms in real systems.

Various synchronizations between slow and fast oscillators, e.g., theta (4–8 Hz) and gamma (>30 Hz) EEG activities, have been observed ubiquitously, and they appear to play an important role in brain function [56, 57, 58, 59]. In fact, 1 : p phase locking has been observed ubiquitously in human EEG experimental studies during the resting state, mental arithmetic tasks, and working memory tasks [60, 61, 62, 63]. However, 1 : p phase synchronizations cannot be described by the model expressed by Eq. (3.1). Therefore, we consider the 1 : p phase-locking state among slow and fast oscillators. If 1 : p phase locking occurs, the value of the phase difference $\phi_2 - p\phi_1$ is constant over time, where ϕ_1 and ϕ_2 are slow and fast phases, respectively. Using the phase reduction theory, we found that 1 : p phase locking can be described as follows [82, 83, 84]:

$$\omega_1 : \omega_2 \cong 1 : p, \quad (3.2)$$

$$\frac{d\phi_1}{dt} = \omega_1 + \Gamma_{1,2}(\phi_2 - p\phi_1), \quad (3.3)$$

$$\frac{d\phi_2}{dt} = \omega_2 + \Gamma_{2,1}(p\phi_1 - \phi_2). \quad (3.4)$$

Here, we explain a simple case of two coupled oscillators described by Eqs. (3.2)–(3.4). Note that many real rhythmic systems generally consist of many oscillators. We assume that the ratio of the natural frequencies of the two oscillators is close to some integer p . Note that, in this situation, the coupling function $\Gamma_{1,2}$ depends on only the phase difference $\phi_2 - p\phi_1$. In our approach, to investigate the nature of interactions between neuronal rhythms, we directly estimate both the natural frequencies ω_i and the phase coupling functions $\Gamma_{i,j}$ from experimental time-series data. In addition, considering unavoidable sources of uncertainty (e.g., observational error or additional unknown disturbance to the system), we introduce independent Gaussian white noise $\eta_i(t)$ into the phase oscillator model as follows:

$$\omega_i : \omega_j \cong p_i : p_j, \quad (3.5)$$

$$\frac{d\phi_1}{dt} = \omega_1 + \sum_{j \neq i}^N \Gamma_{i,j}(p_i\phi_j - p_j\phi_i) + \eta_i. \quad (3.6)$$

Here, we assume that the independent Gaussian white noise $\eta_i(t)$ satisfies $\langle \eta_i(t) \rangle = 0$, $\langle \eta_i(t)\eta_j(t') \rangle = 2D_i\delta_{ij}\delta(t-t')$, where δ_{ij} and $\delta(t)$ are the Kronecker delta and the Dirac delta functions, respectively. D_i indicates the noise strength and $p_i\phi_j - p_j\phi_i$ notes the phase difference, where the p values are integers. Note that this phase oscillator model can explain $p_i : p_j$ synchronization (e.g., $2 : 3$ phase synchronization and $2 : 7$ phase synchronization). We estimate the phase oscillator model (Eq. (3.6)) using almost periodic time-series data. In the following, we employ a straightforward extended version of a previously proposed method [77] and explain the outline of our method.

First, we transform the experimentally-recorded signal $s(t)$ into the phase time-

series $\theta(t)$ by computing the analytic signal as follows:

$$A(t)e^{i\theta_i(t)} = s_i(t) + is_i^H(t), \quad (3.7)$$

where $s_i^H(t)$ denotes the Hilbert transformation of the recorded signal $s_i(t)$ [85], and $\theta(t)$ is the phase of the analytic signal. However, the variable θ is generally different from the phase ϕ in Eq. (3.1) because, according to phase reduction theory, ϕ evolves linearly over time without interaction and noise. It is therefore necessary to transform θ into ϕ , as follows [74, 75]:

$$\phi(\theta) = 2\pi \int_0^\theta f(\theta') d\theta', \quad (3.8)$$

where $f(\theta)$ denotes the probability density distribution of θ .

Second, Bayesian linear regression [86, 87] is applied to estimate the parameters of the phase oscillator model given by Eq. (3.6). Because the coupling function is periodic, we consider Fourier series expansion of the coupling function as:

$$\Gamma_{ij} = a_{ij}^{(0)} + \sum_{m=1}^{M_{ij}} \{a_{ij}^{(m)} \cos(m\psi_{ij}(t_\tau)) + b_{ij}^{(m)} \sin(m\psi_{ij}(t_\tau))\}, \quad (3.9)$$

where $\psi_{i,j}(t_\tau)$ is the extended version of the phase difference $p_i\phi_j(t_\tau) - p_j\phi_i(t_\tau)$ at time t_τ . The times t_τ are discrete points, $t_\tau = t_1 + (\tau - 1)\Delta t$ for $\tau = 1, 2, \dots, T$, and Δt is the sampling interval. In this expansion, M_{ij} denotes the Fourier series order for each coupling function, and the parameters M_{ij} control the complexity of the coupling function. The parameters M_{ij} can be determined by a model selection method based on logarithmic evidence, as explained later.

Finally, the proposed method estimates the model as follows:

$$\frac{d\phi_i}{dt} = \hat{\omega}_i + \sum_{j \neq i}^N \sum_{m=1}^{M_{ij}} \{a_{ij}^{(m)} \cos(m\psi_{ij}(t_\tau)) + b_{ij}^{(m)} \sin(m\psi_{ij}(t_\tau))\} + \eta_i(t_\tau), \quad (3.10)$$

where $\hat{\omega}_i = \omega_i + \sum_{j \neq i}^N a_{ij}^{(0)}$. Thus, the unknown model parameters are $\{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)}\}_i$

and D_i . Here, $\{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)}\}_i$ denotes $\{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)} | j = 1, 2, \dots, N, m = 1, 2, \dots, M_{ij}\}$. The phase velocity $\frac{d\phi_i}{dt}$ is a dependent variable in a standard linear regression problem that is computed from phase time-series data as $\{(\phi_i(t_{\tau+1}) - \phi_i(t_\tau))/\Delta t\}$. Furthermore, an independent variable is computed by the phase difference as $\{\cos(m\psi_{ij}(t_\tau)), \sin(m\psi_{ij}(t_\tau))\}_i$. Here, η_i is an independent and identically distributed random variable. This linear regression problem corresponds to maximizing the following likelihood function:

$$L(\{\phi_i(t_\tau)\} | \{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)}\}_i, D_i) = \prod_{\tau=1}^T \mathcal{N}(\dot{\phi}_i(t_\tau) | \hat{\omega}_i + \sum_{j \neq i}^N \hat{\Gamma}_{i,j}(\psi_{ij}(t_\tau)), \frac{2D_i}{\Delta t}), \quad (3.11)$$

where $\hat{\Gamma}_{i,j}$ equals $\Gamma_{ij} - a_{ij}^{(0)}$, and $\mathcal{N}(x | \mu, \sigma^2)$ denotes Gaussian distribution, where μ and σ^2 are the mean and variance of x , respectively. Using Bayesian theory, the posterior distribution $p(\{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)}\}_i, D_i | \{\phi_i(t_\tau)\})$ is proportional to the product of the likelihood function and the prior distribution $p(\{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)}\}_i, D_i)$:

$$p(\{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)}\}_i, D_i | \{\phi_i(t_\tau)\}) \propto L(\{\phi_i(t_\tau)\} | \{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)}\}_i, D_i) p(\{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)}\}_i, D_i), \quad (3.12)$$

Here, we adopt a Gaussian inverse gamma distribution for the prior distribution $p(\{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)}\}_i, D_i)$. This prior distribution is a conjugate to the likelihood function (the so-called conjugate prior):

$$p(\{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)}\}_i, D_i) = \mathcal{N}(\{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)}\}_i | \chi_i^{old}, \frac{2D_i}{\Delta t} \Sigma_i^{old}) IG(\frac{2D_i}{\Delta t} | \alpha_i^{old}, \beta_i^{old}), \quad (3.13)$$

where $IG(x | \alpha, \beta)$ denotes the inverse gamma distribution with shape parameter α and scale parameter β . χ_i^{old} and $\frac{2D_i}{\Delta t} \Sigma_i^{old}$ are the mean and covariance of model parameters $\hat{\omega}$, $a^{(m)}$, and $b^{(m)}$, respectively. Note that the prior parameters χ_i^{old} , Σ_i^{old} , α_i^{old} , and β_i^{old} are referred to as hyperparameters. We can easily calculate the posterior distribution parameters from the conjugate prior distribution and the likelihood function in Eq. (3.12). The posterior distribution with the updated parameters takes

the form of a Gaussian inverse gamma distribution: χ_i^{new} , Σ_i^{new} , α_i^{new} , and β_i^{new} . The estimated model parameters are the mean of $\mathcal{N}(\{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)}\}_i | \chi_i^{new}, \frac{2D_i}{\Delta t} \Sigma_i^{new})$. The estimated noise level is the mean of $IG(\frac{2D_i}{\Delta t} | \alpha_i^{new}, \beta_i^{new})$. Using the posterior, prior, and likelihood functions, we can compute the logarithmic evidence $\log p(\{\phi_i(t_\tau)\})$:

$$\log p(\{\phi_i(t_\tau)\}) = \log \frac{L(\{\phi_i(t_\tau)\} | \{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)}\}_i, D_i) p(\{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)}\}_i, D_i)}{p(\{\hat{\omega}_i, a_{ij}^{(m)}, b_{ij}^{(m)}\}_i, D_i | \{\phi_i(t_\tau)\})}. \quad (3.14)$$

Thus, we find M_{ij} with the largest logarithmic evidence among all models.

3.2.2 Electronic circuit of van der Pol oscillator

To test the proposed method, we conducted an experiment in which an electronic circuit was used to implement two coupled van der Pol oscillators [88]. The coupling function $\Gamma_{i,j}$ between the oscillators can be obtained theoretically from the corresponding differential equations.

In this experiment, we recorded the rhythmic signals of the electronic circuit. Each oscillator consisted of two multipliers U_1 (AD633, Low Cost Analog Multiplier) and three operational amplifiers U_2 (TL082, $\frac{1}{2}$ Dual BiFET Op Amp) (Figure 3-1a). We conducted two experiments under different conditions. In the first experiment, two same-frequency oscillators were coupled directly. In this experiment, 1:1 phase locking was expected to occur. The parameters of the electronic component were set to $R_1 = 100k\Omega$, $R_2 = 1k\Omega$, $R_{coupling} = 1M\Omega$, $C_1 = C_2 = 0.01\mu F$, $V_1 = 0.115V$, and $V_2 = 0.12V$. R_k and C_i are the parameters of the resistor and capacitor, respectively. Voltages V_i were monitored using a digital voltmeter. In the second experiment, a slow oscillator was coupled to a fast oscillator. In this experiment, 1 : 2 phase locking was expected to occur. The natural frequencies of the slow and fast oscillators were set to satisfy a nearly 1 : 2 ratio. The parameters of the electronic components were the same as those in the first experiment, except that C_1 was changed to $C_1 = 0.02\mu F$ to reduce the natural frequency by one-half. In both cases, the electric potentials x_i and y_i were recorded using an I/O device (NI SCB-68, National Instruments, US). The sampling rate of the electric potential was 15,000 Hz, and the data size was 180

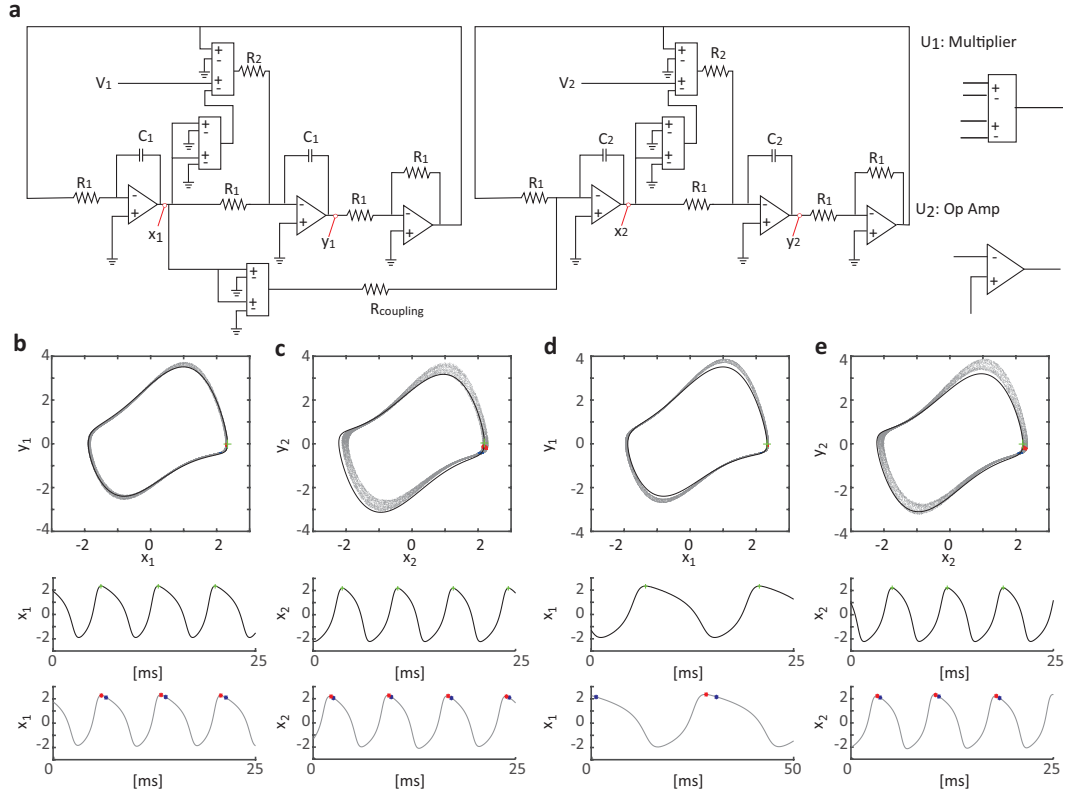


Figure 3-1: Electronic circuit of a pair of van der Pol oscillators and recorded electric potential. (a) Schematic of electronic circuit of two coupled van der Pol oscillators, where x_i and y_i are positions for recording electric potential, R_k denotes resistors, and C_i denotes condensers. Electronic units U_1 and U_2 represent the multiplier and operational amplifiers, respectively. $R_{coupling}$ is a resistor whose resistance is the parameter of the strength of connectivity. (b) Experimental data of electric potentials x_1 and y_1 show the limit-cycle oscillator under the same-frequency (129.1 Hz) coupling condition (gray dots and line). The black trajectory shows the theoretical value computed by the van der Pol oscillator Eqs. (3.15)–(3.18). Here, the frequency is 142.1 Hz. Blue dots represent the zero-phase reference points on the experimental data, which were determined automatically via Hilbert transformation. Green crosses represent the theoretical zero-phase reference points defined as the peak points of x_i . Red dots denote the adjusted zero-phase reference points. (c) x_2 and y_2 show the oscillators under the same-frequency oscillator condition. The frequency of the experimental data is 132.5 Hz and that of the theoretical trajectory is 146.4 Hz. (d) Recorded electric potentials show the slow limit-cycle oscillator under cross-frequency coupling conditions (experimental frequency, 64.1 Hz; theoretical frequency, 71.1 Hz). (e) x_2 and y_2 denote the fast oscillator (experimental frequency, 131.1 Hz; theoretical frequency, 146.4 Hz).

s.

The corresponding theoretical equation of the electronic circuit is given as:

$$\frac{dx_1}{dt} = \frac{1}{R_1 C_1} y_1, \quad (3.15)$$

$$\frac{dy_1}{dt} = \frac{1}{100 R_2 C_1} (10 V_1 - x_1^2) y_1 + \frac{1}{R_1 C_1} x_1, \quad (3.16)$$

$$\frac{dx_2}{dt} = \frac{1}{R_1 C_2} y_2 + \frac{1}{10 R_{coupling} C_2} x_1^2, \quad (3.17)$$

$$\frac{dy_2}{dt} = \frac{1}{100 R_2 C_2} (10 V_2 - x_2^2) y_2 + \frac{1}{R_1 C_2} x_2 + \frac{1}{R_1 C_2} b, \quad (3.18)$$

where x_i and y_i are the corresponding theoretical electric potentials of the i -th oscillator. The trajectories of the van der Pol oscillator are shown by the x_i and y_i signals in Figures 3-1b–e. Note that only x_i was used to estimate the coupling function. The experimental parameters (R_k and C_i) and those in Eqs. (3.15)–(3.18) were the same as the parameters of the electronic components. Note that the term b does not exist in the original van der Pol oscillator. In the case of $b = 0$, the theoretical trajectory and coupling function do not agree with the experimental data and the estimated coupling functions, respectively. Note that the original van der Pol oscillator generates a symmetrical trajectory. However, in the electronic circuit experiment, the trajectory was not exactly symmetrical due to small additional disturbances in the system or the uneven quality of the electronic components. By introducing parameter b ($b = 0.4$), the simulated trajectory can be adjusted to the experimental data (Figures 3-1b–e).

The corresponding theoretical coupling functions were calculated based on Eqs. (3.15)–(3.18) using the adjoint method [89]. In this method, the zero phase is defined as the peak point of x_i (green crosses in Figures 3-1b–e represent the zero-phase reference points on the theoretical orbits). From the experimental data, the zero-phase reference points were determined based on electric potentials x_i using Hilbert transformation (blue dots denote the point of zero-phase reference on the experimental data). As shown in Figures 3-1b–e, a small gap exists between the zero-phase reference points of the theoretical and experimental orbits. In principle, an arbitrary point on the limit-cycle orbit can be defined as the zero-phase point. However, to compare

theoretical and estimated results, the zero-phase reference point of the theoretical model should be consistent with that of the experimental data. Therefore, the zero-phase reference points of the experimental orbits were adjusted to coincide with those of the theoretical model by shifting the experimental reference points to the theoretical points as $\phi_1 = \phi_1 + \pi/6, \phi_2 = \phi_2 + \pi/10$ (red dots denote the revised zero-phase reference points), where ϕ_1 is the experimental phase based on the electric potential x_1 (Figures 3-1b and 3-1d) and ϕ_2 is the experimental phase based on the electric potential x_2 (Figures 3-1c and 3-1e). We estimated the coupling function from the shifted phase data and compared the theoretical coupling function with the estimated function.

3.2.3 Scalp-EEG experiment

We applied the proposed method to scalp EEG data. The method can estimate the coupling functions for same-frequency and cross-frequency synchronization assuming that EEG activities can be considered weakly coupled oscillators. However, it is unclear whether EEG activity can be considered a weakly coupled oscillator system. Thus, we must confirm that the proposed method can estimate the dynamical system from the EEG data successfully.

We used the EEG data recorded during a speech recognition task, and detailed information is described in Chapter 2 [16]. The participants were 16 healthy adults (five females; 11 males; 21–32 years; mean age, 25 years). One participant was excluded due to a low response rate, and another participant was excluded due to an excessive artifact that could not be removed during preprocessing. Note that these participants were also excluded in chapter 2.

We estimated the phase oscillator model between the 3–6 Hz EEG (theta oscillation) and the speech envelope. The theta oscillation is synchronized with the envelope of speech sound, and it plays an important role in speech processing [39, 9, 5]. Speech rhythm consists of linguistic components, e.g., syllabic and prosodic rhythms. A syllable is a unit of speech that separates a word into sound chunks. For example, the Japanese word “KaKuShiKi” (“formality” in English) is composed of four syllable-

bles “/Ka/Ku/Shi/Ki/,” and its sound envelope appears in the 4–5 Hz frequency range in the current speech stimulus. Prosody is the stress and intonation patterns of an utterance. The envelope of prosodic rhythms appears in the 1–3 Hz frequency range. We estimated the coupling function between the EEG oscillation and these speech rhythms. To obtain the theta oscillations, the preprocessed EEG signals were bandpass filtered within 3–6 Hz. The syllabic and prosodic rhythms were computed from the sound stimulus consisting of the noise and speech sounds (Figure 3-2a). The sound envelope was computed as the absolute value of the Hilbert-transformed sound data (Figure 3-2b). To compute the syllabic and prosodic rhythms, the envelope signal was bandpass filtered within 3–6 Hz and 1–3 Hz, respectively (Figures 3-2c and 3-2d). The syllabic and prosodic signals were downsampled to 500 Hz. The instantaneous phases of these rhythms were computed using both Hilbert transformation and correction (Eq. (3.8)). We estimated the coupling functions of 1:1 phase locking (syllable and theta oscillation) and 1:2 phase locking (prosody and theta oscillation).

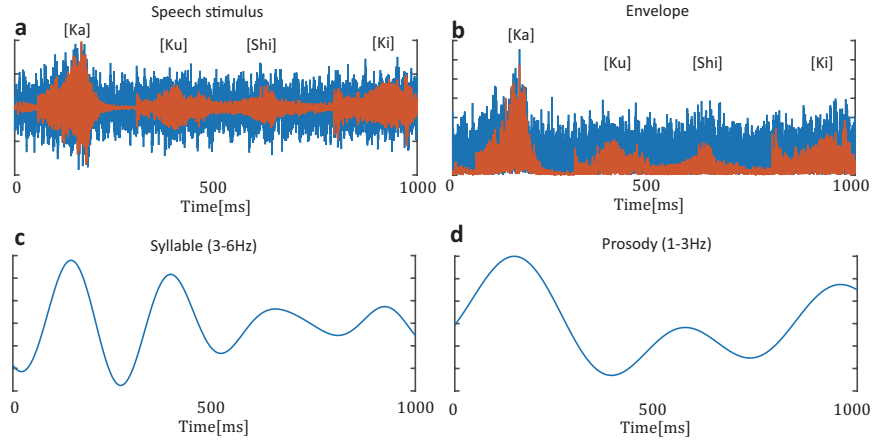


Figure 3-2: Syllable and prosody rhythms in speech sound. (a) Example of speech stimulus. The stimulus consisted of noise and a four-syllable Japanese word. The red line represents a speech wave. The blue line represents the presented sound wave, which consists of speech plus noise sounds. (b) Speech envelope was computed as the absolute value of Hilbert-transformed speech sound. (c) Syllabic rhythms were computed from the speech envelope through the bandpass filter within 3–6 Hz. (d) Prosodic rhythms were computed from the speech envelope through the bandpass filter within 1–3 Hz.

3.3 Results

3.3.1 Numerical simulation of phase oscillator model

First, we applied our Bayesian method to numerical simulation data which was generated from three cross-frequency oscillators with somewhat complicated connections, to evaluate the validity of the proposed method. Simulation data were generated from a network comprising one fast oscillator and two slow oscillators (Figure 3-3) based on the Euler–Maruyama method [90] using the following differential equations:

$$\frac{d\phi_1}{dt} = \omega_1 + 0.1 \sin(\phi_3 - \phi_1) + \eta_1(t), \quad (3.19)$$

$$\begin{aligned} \frac{d\phi_2}{dt} &= \omega_2 + 0.1 \sin(2\phi_1 - \phi_2) \\ &\quad + 0.05\{\sin(2\phi_3 - \phi_2) + \sin(2(2\phi_3 - \phi_2))\} + \eta_2(t), \end{aligned} \quad (3.20)$$

$$\frac{d\phi_3}{dt} = \omega_3 + 0.05 \cos(\phi_2 - 2\phi_3) + \eta_3(t). \quad (3.21)$$

We set the natural frequencies as $\omega_1 = 0.9$, $\omega_2 = 2.1$, and $\omega_3 = 1.1$. ϕ_2 is the phase of the fast oscillator, and ϕ_1 and ϕ_3 are the phases of the slow oscillators. η is independent Gaussian white noise with zero mean and standard deviation of 0.1. Using the method proposed in this paper, we estimated the coupling functions and the natural frequencies from the phase time-series data generated in the experiment.

Figure 3-3 shows the estimated coupling functions and the correct coupling functions. In this case, the correct coupling functions were defined explicitly by Eqs. (3.19)–(3.21). Despite the complicated connections, the results indicate that the estimated and correct coupling functions agree fairly well. Furthermore, the complexity parameter of the coupling function was selected correctly by optimizing the logarithmic evidence. Therefore, the proposed method works quite well at estimating a nontrivial network of phase oscillators comprising oscillators with different natural frequencies.

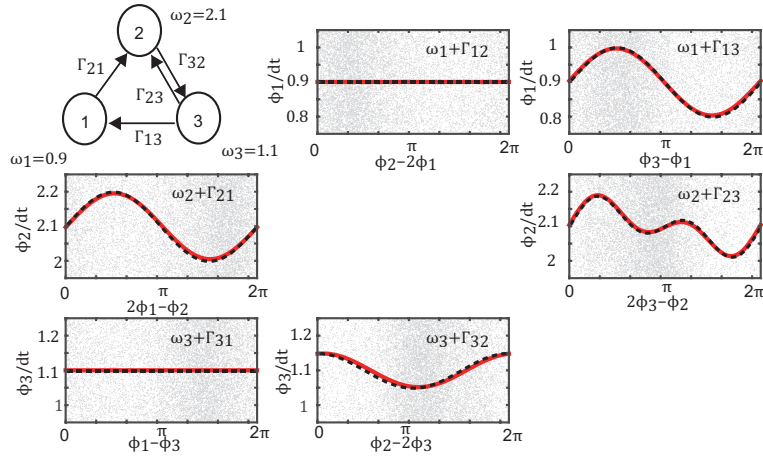


Figure 3-3: Estimated coupling function for numerical simulation data. Upper-left diagram shows the network structure. The estimated coupling functions (red lines) were nearly identical to the correct functions (dashed black line). The gray dots represent the phase time-series data. When the interaction did not exist, the estimated coupling function was identically zero. The proposed method estimated all coupling functions correctly for the simulation data.

3.3.2 Electronic circuit experiment

Before applying the proposed method to EEG data, we recorded the electric potential of the van der Pol electronic circuit and tested the ability of the estimation method using the experimental data. Since the electronic circuit can be explained by the corresponding theoretical differential equations (Eqs. (3.15)–(3.18)), we can derive the correct coupling function theoretically using the adjoint method. We conducted two experiments. One involved coupling oscillators of the same frequency (Figure 3-4a), and the other involved coupling slow and fast oscillators (Figure 3-4e). We transformed the x_1 and x_2 electric potentials of the first and second oscillators, respectively, to phase time-series data and estimated the coupling functions from these data.

The estimated coupling functions with no interaction from the second to first oscillator were identically zero in the experiments involving oscillators of the same frequency (Figure 3-4b) and cross frequency (Figure 3-4f). When coupling existed, the estimated coupling function was the same as the theoretical function under the same-frequency (Figure 3-4c) and cross-frequency conditions (Figure 3-4g). Furthermore, to confirm whether the estimated phase oscillator model can explain the experimental data, we compared the phase difference histograms for the experimental phase data and the two types of simulated phase data. The experimental phase data were computed from electric potentials by both Hilbert transformation and Eq. (3.8). One set of simulated phase data was computed in the phase oscillator model based on the estimated parameters using the Euler–Maruyama method. The other was computed based on the theoretical coupling functions and natural frequency, which were determined by Eqs. (3.15)–(3.18). We computed the experimental, estimated, and theoretical histograms from these phase time-series data. The experimental and estimated histograms of the phase difference were nearly the same under each condition (Figures 3-4d and 3-4h). However, the theoretical histograms differed. The difference among these histograms was caused by the difference of the natural frequency between the theoretical and experimental oscillators (Figures 3-1b–e) due to uncon-

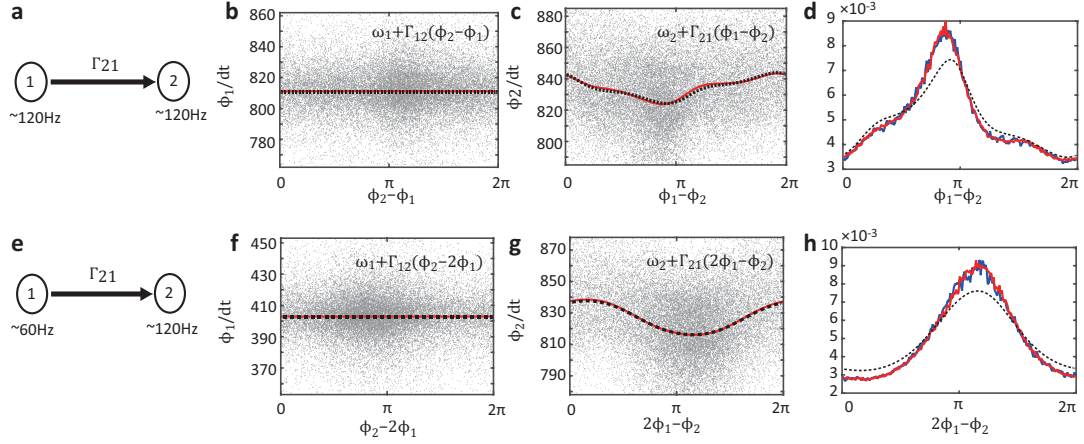


Figure 3-4: Estimated coupling function of electronic circuit. (a) The diagram shows the coupling direction between oscillators of the same frequency. The first oscillator was coupled to the second oscillator. (b) The red line shows the estimated phase coupling function with the natural frequency in the same-frequency coupling case. The dashed black line shows the theoretical coupling function. The coupling function from the second to first oscillator Γ_{12} is identically zero. When there is no interaction, the coupling function is nearly zero. The gray dots show the experimental data points. (c) The coupling functions from the first to second oscillator Γ_{21} . (d) The blue line shows the phase difference histogram of the experimental data in the case of 1:1 phase locking (experimental histogram). The red line shows the simulated histogram calculated in the phase oscillator model estimated from the experimental data (estimated histogram). The dashed black line shows the simulated histogram calculated in the phase oscillator model using the theoretical natural frequencies and coupling functions (theoretical histogram). (e) In the cross-frequency coupling case, the slow oscillator was coupled to the fast oscillator. (f) The coupling function from the fast to slow oscillator is identically zero. (g) The coupling function from the slow to fast oscillator. (h) The experimental, estimated, and theoretical histogram in the 1:2 phase-locking case..

trollable experimental conditions. In other words, the electronic circuit experimental data did not follow the theoretical equations exactly; however, the coupling functions between the experimental van der Pol oscillators were the same as the theoretical coupling functions. In fact, when the theoretical histograms were computed based on the theoretical coupling functions and the estimated natural frequencies rather than the theoretical natural frequencies, the experimental and estimated histograms coincided relatively well with the theoretical histograms. These results indicate that the proposed method works well with real data, even if the data contain observational errors or additional unknown disturbances.

3.3.3 Human EEG experiment

Finally, we applied the proposed method to the EEG data. We estimated the coupling functions between the theta oscillation and the envelope of speech stimulus. The theta oscillation was observed to be synchronized with the speech envelope [5] and is considered to play an important role in speech perception. Generally, the speech envelope consists of syllabic (3–6 Hz) and prosodic (1–3 Hz) rhythms. Thus, we estimated the phase oscillator model under the same-frequency and cross-frequency conditions, i.e., theta oscillation and syllabic rhythm (Figure 3-5), and theta oscillation and prosodic rhythm (Figure 3-6). In these experiments, the phase-difference histograms between the EEG and the speech sound showed 1:1 and 1:2 phase locking (Figures 3-5a and 3-6a). Note that there is obviously no interaction from EEG activity to speech sound. Therefore, we can use this fact to confirm the validity of the estimation results.

In the theta oscillation and syllable data, we assumed that the syllabic rhythm modulated the theta oscillation. The instantaneous phase of the theta oscillation is denoted ϕ_θ , and the phase of the syllable is denoted ϕ_s . The phase difference between the theta oscillation and the syllable is defined as $\phi_s - \phi_\theta$. The phase-difference histogram of each participant showed 1 : 1 phase locking (Figure 3-5a). We estimated

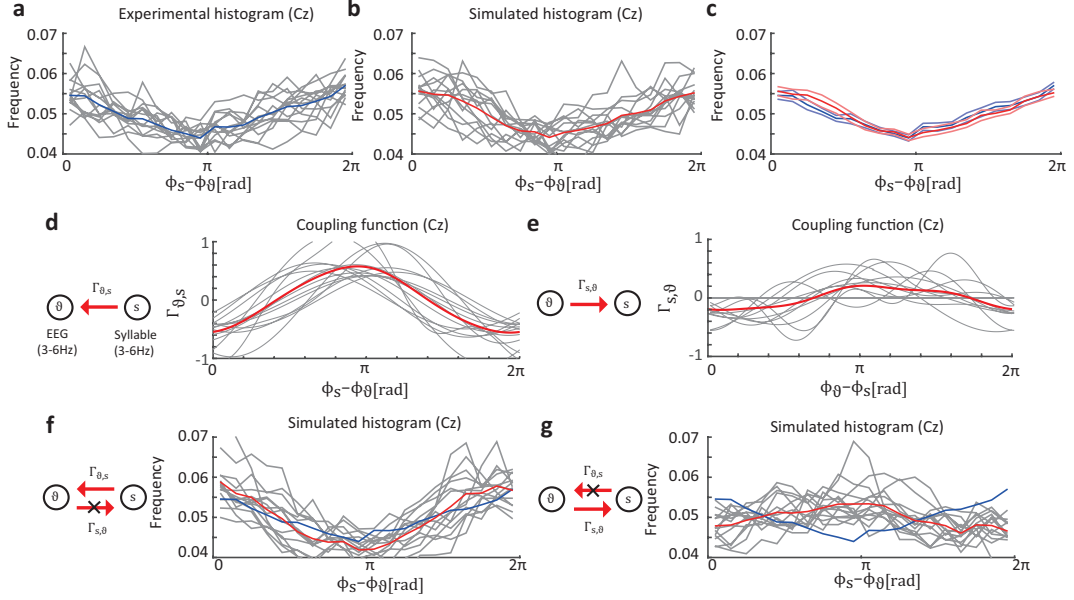


Figure 3-5: Estimated distribution of phase difference between EEG data and syllable envelope. (a) Experimental histogram of phase difference between the theta oscillation on the Cz electrode and the syllabic rhythm (the histograms show phase locking). The gray lines represent histograms of individual participants and the blue line represents the histogram averaged over all participants. (b) Histograms obtained from the simulated data in the estimated phase oscillator model. The averaged histogram is similar to the experimental histogram. The gray lines represent the phase difference histograms of individual participants. The red line represents the average of the simulated histograms. (c) Blue lines represent the averaged experimental histogram and the standard error of mean (SEM). Red lines represent the averaged simulated histograms and the SEM. (d) Estimated coupling functions $\Gamma_{\theta,s}$ from syllabic rhythm to theta oscillation. The gray and red lines represent the results of individual participants and the average results of all participants, respectively. (e) Estimated coupling functions $\Gamma_{s,\theta}$ are considerably smaller than the opposite directional coupling functions. (f) Simulated histograms where the coupling functions $\Gamma_{s,\theta}$ are removed. The effect on the original phase-locking state was negligible. (g) Histograms where $\Gamma_{\theta,s}$ were removed are nearly flat.

the phase oscillator models of 1 : 1 phase locking as follows:

$$\frac{d\phi_\theta}{dt} = \omega_\theta + \Gamma_{\theta,s}(\phi_s - \phi_\theta) + \eta_\theta(t), \quad (3.22)$$

$$\frac{d\phi_s}{dt} = \omega_s + \Gamma_{s,\theta}(\phi_\theta - \phi_s) + \eta_s(t). \quad (3.23)$$

To confirm that the estimated dynamical system can explain the experimental data, we simulated phase synchronization based on the estimated phase oscillator model. Figure 3-5b shows the phase difference histograms obtained from the simulated data. Note that, for the simulated data, phase differences were calculated by numerical simulation performed using the estimated phase oscillator model (Eqs. (3.22)–(3.23)). The histograms of the simulation data were similar to those of the experimental data (Figure 3-5c), which indicates that the estimated phase oscillator model can explain the experimental data. Furthermore, the estimated coupling functions $\Gamma_{\theta,s}$ were consistent among all participants (Figure 3-5d). In contrast, the estimated coupling functions $\Gamma_{s,\theta}$ had smaller amplitudes than $\Gamma_{\theta,s}$ (Figure 3-5e) and were not consistent among all participants. These results are reasonable in terms of the relationship between EEG and speech sound because direct interaction from theta oscillation to speech sound never exists. To examine the effects of each coupling function on the dynamics, we computed the simulated histogram under the condition that either $\Gamma_{s,\theta}$ or $\Gamma_{\theta,s}$ was set to identically zero. In the case of $\Gamma_{s,\theta} = 0$ (Figure 3-5f), the resultant histogram shows that, compared to the original dynamics in Figure 3-5b, the synchronized state is almost maintained. This implies that the coupling function $\Gamma_{s,\theta}$ does not contribute to the realization of 1 : 1 phase locking. In the case of $\Gamma_{\theta,s} = 0$, the synchronized state disappeared, as shown in the flatter histograms (Figure 3-5g). Consequently, the results indicate that the coupling function $\Gamma_{\theta,s}$ primarily contributed to 1 : 1 phase locking.

In the theta oscillation and prosody data, we assumed that the prosodic rhythms modulated the theta oscillation. Here, let ϕ_p denote the prosody phase. Considering the 1 : 2 phase-locking state, the phase difference between theta oscillation and prosody is reasonably defined as $2\phi_p - \phi_\theta$. The phase difference histograms of each

participant showed 1 : 2 phase locking (Figure 3-6a). Next, we considered the phase oscillator model for 1 : 2 phase locking as follows:

$$\frac{d\phi_\theta}{dt} = \omega_\theta + \Gamma_{\theta,p}(2\phi_p - \phi_\theta) + \eta_\theta(t), \quad (3.24)$$

$$\frac{d\phi_p}{dt} = \omega_p + \Gamma_{p,\theta}(\phi_\theta - 2\phi_p) + \eta_p(t), \quad (3.25)$$

We confirmed that the estimation result can explain the experimental data by calculating the simulated phase difference histograms under the 1 : 2 phase-locking condition (Figure 3-6b). The simulated histograms were similar to the experimental histograms (Figure 3-6c). The results indicate that the estimation phase interaction functions can explain the experimental data, as well as the 1 : 1 phase-locking condition. The estimated coupling functions of all participants were consistent (Figures 3-6d and 3-6e). Furthermore, the estimated coupling functions $\Gamma_{p,\theta}$ showed small amplitude or were identically zero (Figure 3-6e). These results clearly show that there were no coupling function from EEG to speech sound, which is reasonable given the relationship between EEG and speech sounds in the experiments. In the case of $\Gamma_{p,\theta} = 0$, the simulated phase difference histograms also showed phase locking (Figure 3-6f), as well as Figure 3-6b. In contrast, phase locking disappeared for $\Gamma_{\theta,p} = 0$ (Figure 3-6g). These results indicate that the coupling function $\Gamma_{\theta,p}$ contributed to 1 : 2 phase locking.

In both cases, our method could estimate whether there was a relationship between the EEG activity and speech sound even though there was some variance due to estimation inaccuracies. In the case of the theta oscillation and prosody data, the estimated coupling functions $\Gamma_{p,\theta}$ showed small amplitudes or were identically zero, clearly demonstrating the asymmetry of the relationship between the EEG activity and speech sound. In contrast, for the theta oscillation and syllable data, the estimated coupling functions $\Gamma_{s,\theta}$ showed somewhat larger amplitudes than $\Gamma_{p,\theta}$, giving no clear indication of an asymmetric relationship. Therefore, in order to determine whether there was a asymmetry relationship, we estimated the coupling functions using surrogate data (Figure 3-7). The surrogate data consisted of randomly time-

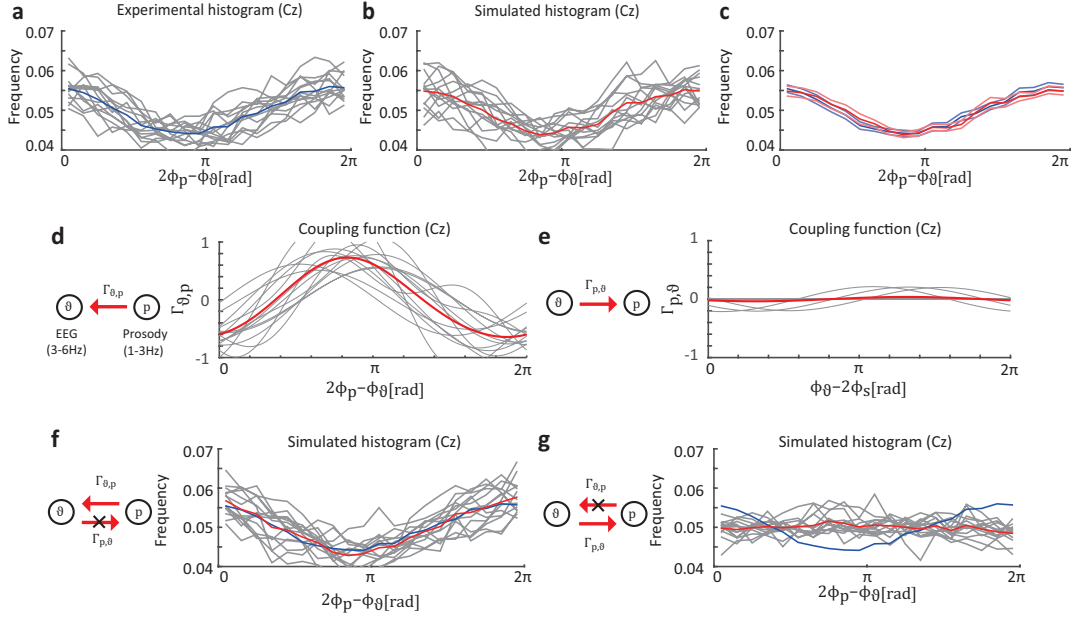


Figure 3-6: Estimated distribution of phase difference between EEG data and prosody envelope. (a) Experimental phase difference histograms for 1:2 phase locking. (b) Simulated histograms based on the estimated phase oscillator model. (c) Blue lines represent the average and SEM of experimental phase difference histograms. Red lines represent the average and SEM of simulated histograms. (d) Estimated coupling functions $\Gamma_{\theta,p}$. (e) Estimated coupling functions $\Gamma_{p,\theta}$. (f) Simulated histograms where coupling functions $\Gamma_{p,\theta}$ are removed. (g) Simulated histograms where $\Gamma_{\theta,p}$ is removed are uniform.

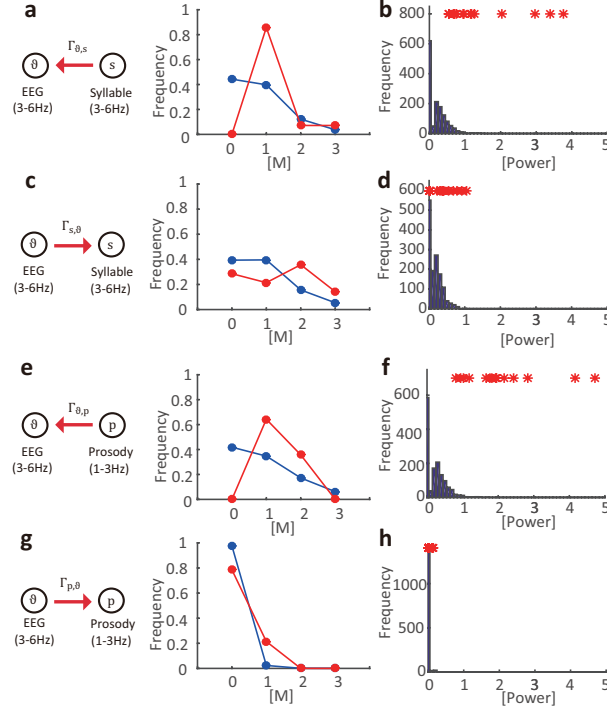


Figure 3-7: Histograms of property of estimated coupling functions in surrogate data. (a) Histograms of the M values which were selected based on logarithmic evidence for the coupling functions $\Gamma_{\theta,s}$. The blue line represents the model selection histogram for the 1,400 surrogate datasets, while the red line represents the model selection histogram for the 14 participants' original data. (b) Histogram of all coupling function powers for the surrogate data (including $M=0,1,2,3$). The red stars represent the coupling function powers for the original data. (c) Model selection histograms for the coupling functions $\Gamma_{s,\theta}$. (d) Histogram of powers of the coupling functions $\Gamma_{s,\theta}$. (e) Model selection histograms for the coupling functions $\Gamma_{\theta,p}$. (f) Histogram of powers of the coupling functions $\Gamma_{\theta,p}$. (g) Model selection histograms for the coupling functions $\Gamma_{p,\theta}$. (h) Histogram of powers of the coupling functions $\Gamma_{p,\theta}$.

shifted speech-sound phase data and original EEG data that had not been time-shifted. Note that there was no temporal relationship between the time-series, speech-sound phase data and the original EEG signals. This random shifting process was repeated 100 times for each of the 14 participants. Then, we estimated the coupling functions using these 1,400 surrogate datasets and computed histograms of the model selection results for the appropriate Fourier modes based on the logarithmic evidence and the coupling function powers $\int_0^{2\pi} |\Gamma(\psi)|^2 d\psi$. For the coupling functions $\Gamma_{\theta,s}$ and $\Gamma_{s,\theta}$, the model selection histograms showed that the $M = 0$ model was selected more often than the other models (Figures 3-7a and 3-7e), while the original data results showed that none of the participants selected the $M=0$ model. Furthermore, the integrated values of coupling function power showed that the original data results did not follow the same histograms as the surrogate data results (Figures 3-7b and 3-7f). For the coupling functions $\Gamma_{s,\theta}$ and $\Gamma_{p,\theta}$, the model selection results for the original data were not largely different from those that for the surrogate data (Figures 3-7c and 3-7g), and the integrated coupling function values were relatively small (Figures 3-7d and 3-7h). Consequently, these results suggest that the coupling functions, $\Gamma_{s,\theta}$ and $\Gamma_{p,\theta}$, showed no relationship between the EEGs and speech sounds.

3.4 Discussion

We have proposed an estimation method to identify the phase dynamics of cross-frequency synchronization using rhythmic time-series data. By identifying the dynamics, we can reveal the direction of coupling and the role of each coupling function in synchronization. To confirm the reliability of the proposed method, we estimated a dynamical system from time-series data obtained by numerical simulation and experimentation using an electronic circuit, and we showed that these results were estimated successfully. In addition, we applied the proposed method to scalp EEG data and evaluated its validity based on the estimation results.

3.4.1 Validity of estimation method for cross-frequency synchronization

We can obtain the theoretical coupling function from the numerical simulation and data from the electronic circuit experiment. To confirm the reliability of the proposed method, we compared the estimated results to theoretical results. In the simulation, time-series data were generated by numerical simulation used in the given phase oscillator model. In this case, we knew the true instantaneous phase of the time-series data and the correct coupling functions. The proposed method worked well with the simulation data (Figure 3-7), and the complexity parameter for the coupling function M_{ij} was selected correctly.

We also estimated a dynamical system using electric potential data, i.e., real time-series data. In this situation, the corresponding theoretical coupling function was computed using the adjoint method from which we constructed a theoretical model of real electronic circuits. The results demonstrated that the proposed method can correctly estimate the coupling functions (Figure 3-4). Furthermore, to confirm that the estimated phase oscillator model can reproduce real time-series data, we compared the phase difference histogram of real time-series data with those of the simulated data. The results demonstrate that the estimated coupling functions and the noise strengths can explain the real data, including any additional disturbance imposed on the system.

In the EEG data, the correct coupling function to be compared to the estimated function is unknown. Therefore, we must consider an alternative procedure to examine the validity of the estimation results. To this end, we considered the following three steps. In the first step, we focused on the coupling functions from the EEG activity to the speech stimulus (Figures 3-5e and 3-6e). Under this EEG experimental condition, EEG activity did not influence speech sound because the timing of the external speech sound was given by a recorded sound. Our estimated dynamics showed that the coupling function from EEG to speech sound had smaller amplitude than the coupling function in the opposite direction (Figures 3-5d and 3-6d) and did

not influence the phase difference histograms (Figures 3-5f and 3-6f). Therefore, the estimated network structure is consistent with the real EEG and speech system under this experimental condition. In the second step, we confirmed whether the simulated phase difference histograms were similar to the experimental histograms. Our estimation results and the experimental data both showed phase locking (Figures 3-5a–b and 3-6a–b). In addition to results of the first step, these results suggest that the phase oscillator model can explain the EEG and speech sound data. In the final step, we confirmed that the estimated coupling functions were consistent across all participants to check whether the above results occurred by chance. Our results indicated that the coupling functions and phase difference histograms were similar across all participants on the Cz electrode (Figures 3-5 and 3-6). Furthermore, the estimation results for neighbor electrodes (e.g., FCz, Pz, CP, and CP2) showed results similar to those obtained on the Cz electrode. Based on the results obtained by performing these three steps, the mechanism between EEG and speech sound can be explained by the dynamical phase oscillator system.

3.4.2 Remarks on estimation method

To apply the proposed method to EEG data, it is necessary to consider whether the systems to be estimated can be considered a weakly coupled oscillator system. It is well known that EEG phases are often locked by external input among trials [21, 2]. In addition, it is possible that phase locking is generated by phase-resetting through strong external inputs. If phase-resetting depends on the timing of external inputs rather than the phase difference, the interaction between the EEG phase and the stimulus cannot be explained by the phase oscillator model. Therefore, it is necessary to avoid phase resetting caused by sudden and strong external inputs. Event related phase locking is often induced at stimulus onset. Therefore, to prevent such phase-resetting, we presented noise (increased linearly over 0.5 s) prior to presenting speech sound. Furthermore, to avoid a situation where a strong external input induces phase-resetting at speech onset, we employed a bandpass filter to decrease the strong periodic speech sound signals.

Note that the proposed method cannot estimate the coupling function if the EEG phase is completely synchronized. Under such synchronization conditions, each phase difference between the two oscillators takes only a specific value. Therefore, except for this specific value, there is no information about the coupling function on the other phase difference value. To obtain the full range of the coupling function, the phase differences in the data must be distributed in the range 0 to 2π , as shown in Figure 3-3.

This study focused primarily on $1 : p$ phase synchronization; however, other types of cross-frequency synchronization exist [57, 58], e.g., phase-amplitude, amplitude-amplitude, and phase-frequency synchronization. These synchronizations are also important from a cognitive function perspective; however, the proposed method cannot be applied to such dynamical systems. In future, we plan to construct a method that is applicable to the experimental data of these synchronizations.

Methods to quantify the causality between different frequency rhythms [91, 92, 93, 94] have been proposed. Such methods, including transfer entropy and Granger causality, may reveal more general causality than the proposed method, which can only estimate the coupling function related to the $p_i : p_j$ phase synchronization. However, the proposed method can reveal causality and quantify the phase interaction function; thus, it can examine the role of connection in phase synchronization from a dynamical system perspective.

The proposed method can estimate the coupling functions of simulation data and experimental electronic circuit data accurately. In the EEG experiment, we estimated the dynamical system of EEG and speech sound. Note that we examined the dynamics of phase synchronization between a single EEG activity and speech sound rather than between two EEG activities. We applied the proposed method to synchronization between EEG and speech because the direction of causality between EEG and speech sound is clear, whereas that of EEG phases is unknown. Therefore, EEG and speech data were used to verify the estimation results. It is expected that the proposed method can serve as a useful tool to reveal the role of connectivity and causality in neural oscillations.

Chapter 4

Conclusion

4.1 Mechanisms of Synchronization between Speech Sound and EEG Oscillation

The results of estimation in chapter 3 suggest that phase velocity modulation without phase resetting induces synchronization between the speech rhythms and theta oscillations. This synchronization plays a role in speech encoding and in the segmentation of speech components [9]. A previous study has reported that phase resetting at the speech onset or at temporal speech edges induces synchronization [12], whereas the results of our study revealed that a coupling function in the condition restraining the phase resetting induces synchronization (Figures 3-5d and 3-6d). In this condition, speech and EEG oscillation are not completely synchronized; nevertheless, a phase difference value equal to zero is most frequently observed (Figures 3-5a and 3-6a). The coupling function may induce this synchronization by modulating the phase velocity of the EEG, which depends on the phase difference. In this result, the natural frequency of the speech envelope is lower compared with that of the theta oscillation (theta oscillation: 4.29 ± 0.36 Hz, syllabic rhythm: 3.80 Hz, prosody: 1.77 Hz). Then, when the phase difference is around zero, the phase velocity of the theta oscillation is decreased by the coupling function and approaches the rhythms of speech sound. Our results suggest that this mechanism induces the phase synchronization. Note that,

our hypothesis does not deny that the phase synchronization is induced by phase resetting, because our experimental condition restraining the phase resetting. Our results indicate that phase velocity modulation depends on the phase difference between the EEG and the rhythms of speech sounds when speech stimulus is relatively weak. In general speech sounds, both modulation of the phase velocity and phase resetting may induce phase synchronization.

However, if the natural frequency of the speech envelope is higher compare with that of the theta oscillations, our estimated coupling function cannot generate zero-phase synchronization. Further study is warranted to reveal the mechanisms of neural synchronization in more detail.

4.2 Functional Role of Phase effect and Synchronization

We have demonstrated that there are two types of theta oscillations: one modulating the behavioral performance (in chapter 2) and one synchronized with speech rhythms (in chapter 3). Note that, the electrodes with which these theta oscillations were recorded were different (for the theta oscillations in chapter 2: P3 and CP1 electrodes; in chapter 3: FCz, Cz, Pz, CP, and CP2 electrodes). Furthermore, we cannot estimate consistent coupling functions across all participants for the theta oscillation recorded with the P3 and CP1 electrodes and the delta oscillations. Thus, we consider that these neural oscillations occur in different areas and may play different functional roles.

In chapter 2, we suggest that the behavioral performance is modulated by theta and delta phases and that these EEG activities are associated with the motor cortical activity. Speech recognition has been considered to be enhanced by the identification of a speaker's motor signals [24, 25, 40]. Viewing the speaker's face or listening to him/her speech allows one to identify the motor signals [29]. The facial movements occur before speech production. Thus, viewing the facial movements allows one to

predict the timing of speech production [5]. Furthermore, the motor cortical activity influences speech comprehension [23]. In particular, delta and theta oscillations in the motor cortex modulate the phases of neural oscillations in the auditory cortex to improve the synchronization between neural oscillation in the auditory cortex and speech rhythms [15]. A previous study has suggested that synchronization contributes to speech encoding in each linguistic component [9]. Therefore, it is possible that viewing facial movements changes the theta and delta phase predictively to an optimal state for improving the synchronization.

In chapter 3, we have shown that theta oscillations are synchronized with speech rhythms using estimated coupling functions. Many previous studies have reported that delta, theta, and gamma oscillations in the auditory cortex are synchronized with speech components such as prosody, syllables, and phonemes [2, 12, 21]. Synchronization plays the roles of encoding speech and separating the different speech components [9]. Thus, a synchronized neural oscillation may be considered to extract the speech components, and synchronization contributes to the performance for speech recognition. If theta and delta phase-modulating behavioral performance enhances synchronization, this should help in encoding speech, resulting in speech comprehension being enhanced. However, our results do not provide sufficient corroboration of this hypothesis. In particular, we could not directly show that the theta and delta phases enhance the synchronization between neural oscillations and speech rhythm. Further work is needed to find the interactions which enhances the synchronization.

THIS PAGE INTENTIONALLY LEFT BLANK

Bibliography

- [1] G. H. Bishop. Cyclic changes in excitability of the optic pathway of the rabbit. *American Journal of Physiology–Legacy Content*, Vol. 103, No. 1, pp. 213–224, 1932.
- [2] P. Lakatos, A. S. Shah, K. H. Knuth, I. Ulbert, G. Karmos, and C. E. Schroeder. An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J Neurophysiol*, Vol. 94, No. 3, pp. 1904–1911, 2005.
- [3] Mircea Steriade, David A McCormick, and Terrence J Sejnowski. Thalamocortical oscillations in the sleeping and aroused brain. *Science*, Vol. 262, pp. 679–679, 1993.
- [4] József Fiser, Chiayu Chiu, and Michael Weliky. Small modulation of ongoing cortical dynamics by sensory input during natural vision. *Nature*, Vol. 431, No. 7008, pp. 573–578, 2004.
- [5] C. E. Schroeder, P. Lakatos, Y. Kajikawa, S. Partan, and A. Puce. Neuronal oscillations and visual amplification of speech. *Trends Cogn Sci*, Vol. 12, No. 3, pp. 106–113, 2008.
- [6] B. S. Ng, T. Schroeder, and C. Kayser. A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. *J Neurosci*, Vol. 32, No. 35, pp. 12268–12276, 2012.
- [7] N. A. Busch, J. Dubois, and R. VanRullen. The phase of ongoing eeg oscillations predicts visual perception. *J Neurosci*, Vol. 29, No. 24, pp. 7869–7876, 2009.
- [8] N. A. Busch and R. VanRullen. Spontaneous eeg oscillations reveal periodic sampling of visual attention. *Proc Natl Acad Sci U S A*, Vol. 107, No. 37, pp. 16048–16053, 2010.
- [9] A. L. Giraud and D. Poeppel. Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci*, Vol. 15, No. 4, pp. 511–517, 2012.
- [10] J. Ward. *The student’s guide to cognitive neuroscience*. Psychology Press, 2015.
- [11] P. Fries. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn Sci*, Vol. 9, No. 10, pp. 474–480, 2005.

- [12] J. Gross, N. Hoogenboom, G. Thut, P. Schyns, S. Panzeri, P. Belin, and S. Garrod. Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol*, Vol. 11, No. 12, p. e1001752, 2013.
- [13] Scott Makeig, Marissa Westerfield, T-P Jung, Sonia Enghoff, Jeanne Townsend, Eric Courchesne, and Terrence J Sejnowski. Dynamic brain sources of visual evoked responses. *Science*, Vol. 295, No. 5555, pp. 690–694, 2002.
- [14] M. Lobier, F. Siebenhuhner, S. Palva, and J. M. Palva. Phase transfer entropy: a novel phase-based measure for directed connectivity in networks coupled by oscillatory interactions. *Neuroimage*, Vol. 85 Pt 2, pp. 853–872, 2014.
- [15] H. Park, R. A. Ince, P. G. Schyns, G. Thut, and J. Gross. Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Curr Biol*, Vol. 25, No. 12, pp. 1649–1653, 2015.
- [16] T. Onojima, K. Kitajo, and H. Mizuhara. Ongoing slow oscillatory phase modulates speech intelligibility in cooperation with motor cortical activity. *PLoS ONE*, Vol. 12, No. 8, 2017.
- [17] T. Onojima, T. Goto, H. Mizuhara, and T. Aoyagi. A dynamical systems approach for estimating phase interactions between rhythms of different frequencies from experimental data. *PLOS Computational Biology*, Vol. 14, No. 1, 2018.
- [18] P. Lakatos, G. Karmos, A. D. Mehta, I. Ulbert, and C. E. Schroeder. Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, Vol. 320, No. 5872, pp. 110–113, 2008.
- [19] P. Lakatos, C. M. Chen, M. N. O’Connell, A. Mills, and C. E. Schroeder. Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron*, Vol. 53, No. 2, pp. 279–92, 2007.
- [20] G. Stefanics, B. Hangya, I. Hernadi, I. Winkler, P. Lakatos, and I. Ulbert. Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *J Neurosci*, Vol. 30, No. 41, pp. 13578–13585, 2010.
- [21] R. Galambos, S. Makeig, and P. J. Talmachoff. A 40-hz auditory potential recorded from the human scalp. *Proc Natl Acad Sci U S A*, Vol. 78, No. 4, pp. 2643–2647, 1981.
- [22] W. H. Sumby and I. Pollack. Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, Vol. 26, No. 2, pp. 212–215, 1954.
- [23] A. D’Ausilio, F. Pulvermuller, P. Salmas, I. Bufalari, C. Begliomini, and L. Fadiga. The motor somatotopy of speech perception. *Curr Biol*, Vol. 19, No. 5, pp. 381–385, 2009.
- [24] A. M. Liberman and I. G. Mattingly. The motor theory of speech perception revised. *Cognition*, Vol. 21, No. 1, pp. 1–36, 1985.

- [25] A. M. Liberman and D. H. Whalen. On the relation of speech to language. *Trends Cogn Sci*, Vol. 4, No. 5, pp. 187–196, 2000.
- [26] Y. Du, B. R. Buchsbaum, C. L. Grady, and C. Alain. Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proc Natl Acad Sci U S A*, Vol. 111, No. 19, pp. 7126–7131, 2014.
- [27] I. G. Meister, S. M. Wilson, C. Deblieck, A. D. Wu, and M. Iacoboni. The essential role of premotor cortex in speech perception. *Curr Biol*, Vol. 17, No. 19, pp. 1692–6, 2007.
- [28] J. Alho, M. Sato, M. Sams, J. L. Schwartz, H. Tiitinen, and I. P. Jaaskelainen. Enhanced early-latency electromagnetic activity in the left premotor cortex is associated with successful phonetic categorization. *Neuroimage*, Vol. 60, No. 4, pp. 1937–1946, 2012.
- [29] K. E. Watkins, A. P. Strafella, and T. Paus. Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, Vol. 41, No. 8, pp. 989–994, 2003.
- [30] G. Gratton, M. G. Coles, and E. Donchin. A new method for off-line removal of ocular artifact. *Electroencephalogr Clin Neurophysiol*, Vol. 55, No. 4, pp. 468–484, 1983.
- [31] P. J. Allen, G. Polizzi, K. Krakow, D. R. Fish, and L. Lemieux. Identification of eeg events in the mr scanner: the problem of pulse artifact and a method for its subtraction. *Neuroimage*, Vol. 8, No. 3, pp. 229–239, 1998.
- [32] P. J. Allen, O. Josephs, and R. Turner. A method for removing imaging artifact from continuous eeg recorded during functional mri. *Neuroimage*, Vol. 12, No. 2, pp. 230–239, 2000.
- [33] C. Tallon-Baudry, O. Bertrand, C. Delpuech, and J. Pernier. Oscillatory gamma-band (30–70 hz) activity induced by a visual search task in humans. *J Neurosci*, Vol. 17, No. 2, pp. 722–734, 1997.
- [34] E. Maris and R. Oostenveld. Nonparametric statistical testing of eeg- and meg-data. *J Neurosci Methods*, Vol. 164, No. 1, pp. 177–190, 2007.
- [35] M. Esterman, B. J. Tamber-Rosenau, Y. C. Chiu, and S. Yantis. Avoiding non-independence in fmri data analysis: leave one subject out. *Neuroimage*, Vol. 50, No. 2, pp. 572–576, 2010.
- [36] D. W. Marquardt and R. D. Snee. Ridge regression in practice. *American Statistician*, Vol. 29, No. 1, pp. 3–20, 1975.
- [37] H. Mizuhara, N. Sato, and Y. Yamaguchi. Cortical networks dynamically emerge with the interplay of slow and fast oscillations for memory of a natural scene. *Neuroimage*, Vol. 111, pp. 76–84, 2015.

- [38] H. Mizuhara, L. Q. Wang, K. Kobayashi, and Y. Yamaguchi. A long-range cortical network emerging with theta oscillation in a mental task. *Neuroreport*, Vol. 15, No. 8, pp. 1233–1238, 2004.
- [39] M. Bourguignon, X. De Tiege, M. O. de Beeck, N. Ligot, P. Paquier, P. Van Bogaert, S. Goldman, R. Hari, and V. Jousmaki. The pace of prosodic phrasing couples the listener’s cortex to the reader’s voice. *Hum Brain Mapp*, Vol. 34, No. 2, pp. 314–326, 2013.
- [40] B. Galantucci, C. A. Fowler, and M. T. Turvey. The motor theory of speech perception reviewed. *Psychon Bull Rev*, Vol. 13, No. 3, pp. 361–377, 2006.
- [41] B. Osnes, K. Hugdahl, and K. Specht. Effective connectivity analysis demonstrates involvement of premotor cortex during speech perception. *Neuroimage*, Vol. 54, No. 3, pp. 2437–2445, 2011.
- [42] R. I. Schubotz and D. Y. von Cramon. Functional-anatomical concepts of human premotor cortex: evidence from fmri and pet studies. *Neuroimage*, Vol. 20 Suppl 1, pp. S120–S131, 2003.
- [43] H. Mizuhara. Cortical dynamics of human scalp eeg origins in a visually guided motor execution. *Neuroimage*, Vol. 62, No. 3, pp. 1884–1895, 2012.
- [44] K. Jerbi, J. P. Lachaux, K. N’Diaye, D. Pantazis, R. M. Leahy, L. Garnero, and S. Baillet. Coherent neural representation of hand speed in humans revealed by meg imaging. *Proc Natl Acad Sci U S A*, Vol. 104, No. 18, pp. 7676–7681, 2007.
- [45] A. L. Giraud, A. Kleinschmidt, D. Poeppel, T. E. Lund, R. S. Frackowiak, and H. Laufs. Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron*, Vol. 56, No. 6, pp. 1127–1134, 2007.
- [46] J. Besle, C. A. Schevon, A. D. Mehta, P. Lakatos, R. R. Goodman, G. M. McKhann, R. G. Emerson, and C. E. Schroeder. Tuning of the human neocortex to the temporal dynamics of attended events. *J Neurosci*, Vol. 31, No. 9, pp. 3176–3185, 2011.
- [47] C. E. Schroeder and P. Lakatos. Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci*, Vol. 32, No. 1, pp. 9–18, 2009.
- [48] A. Riehle, S. Grun, M. Diesmann, and A. Aertsen. Spike synchronization and rate modulation differentially involved in motor cortical function. *Science*, Vol. 278, No. 5345, pp. 1950–1953, 1997.
- [49] P. G. Schyns, G. Thut, and J. Gross. Cracking the code of oscillatory activity. *PLOS Biology*, Vol. 9, No. 5, p. e1001064, 2011.
- [50] E. Rodriguez, N. George, J. P. Lachaux, J. Martinerie, B. Renault, and F. J. Varela. Perception’s shadow: long-distance synchronization of human brain activity. *Nature*, Vol. 397, No. 6718, pp. 430–433, 1999.

- [51] P. R. Roelfsema, A. K. Engel, P. König, and W. Singer. Visuomotor integration is associated with zero time-lag synchronization among cortical areas. *Nature*, Vol. 385, No. 6612, pp. 157–161, 1997.
- [52] R. Srinivasan, D. P. Russell, G. M. Edelman, and G. Tononi. Increased synchronization of neuromagnetic responses during conscious perception. *J Neurosci*, Vol. 19, No. 13, pp. 5435–5448, 1999.
- [53] A. von Stein. Synchronization between temporal and parietal cortex during multimodal object processing in man. *Cerebral Cortex*, Vol. 9, No. 2, pp. 137–150, 1999.
- [54] F. Varela, J. P. Lachaux, E. Rodriguez, and J. Martinerie. The brainweb: phase synchronization and large-scale integration. *Nat Rev Neurosci*, Vol. 2, No. 4, pp. 229–239, 2001.
- [55] T. Womelsdorf, J. M. Schoffelen, R. Oostenveld, W. Singer, R. Desimone, A. K. Engel, and P. Fries. Modulation of neuronal interactions through neuronal synchronization. *Science*, Vol. 316, No. 5831, pp. 1609–1612, 2007.
- [56] R. T. Canolty, E. Edwards, S. S. Dalal, M. Soltani, S. S. Nagarajan, H. E. Kirsch, M. S. Berger, N. M. Barbaro, and R. T. Knight. High gamma power is phase-locked to theta oscillations in human neocortex. *Science*, Vol. 313, No. 5793, pp. 1626–1628, 2006.
- [57] R. T. Canolty and R. T. Knight. The functional role of cross-frequency coupling. *Trends Cogn Sci*, Vol. 14, No. 11, pp. 506–515, 2010.
- [58] O. Jensen and L. L. Colgin. Cross-frequency coupling between neuronal oscillations. *Trends Cogn Sci*, Vol. 11, No. 7, pp. 267–269, 2007.
- [59] J. E. Lisman and O. Jensen. The theta-gamma neural code. *Neuron*, Vol. 77, No. 6, pp. 1002–1016, 2013.
- [60] V. V. Nikulin and T. Brismar. Phase synchronization between alpha and beta oscillations in the human electroencephalogram. *Neuroscience*, Vol. 137, No. 2, pp. 647–657, 2006.
- [61] V. V. Nikulin, G. Nolte, and G. Curio. A novel method for reliable and fast extraction of neuronal eeg/meg oscillations on the basis of spatio-spectral decomposition. *Neuroimage*, Vol. 55, No. 4, pp. 1528–1535, 2011.
- [62] J. M. Palva, S. Monto, S. Kulashekhar, and S. Palva. Neuronal synchrony reveals working memory networks and predicts individual memory capacity. *Proc Natl Acad Sci U S A*, Vol. 107, No. 16, pp. 7580–7585, 2010.
- [63] J. M. Palva, S. Palva, and K. Kaila. Phase synchrony among neuronal oscillations in the human cortex. *J Neurosci*, Vol. 25, No. 15, pp. 3962–3972, 2005.

- [64] P. Sauseng, W. Klimesch, W. R. Gruber, and N. Birbaumer. Cross-frequency phase synchronization: a brain mechanism of memory matching and attention. *Neuroimage*, Vol. 40, No. 1, pp. 308–317, 2008.
- [65] B. Schack, W. Klimesch, and P. Sauseng. Phase synchronization between theta and upper alpha oscillations in a working memory task. *Int J Psychophysiol*, Vol. 57, No. 2, pp. 105–114, 2005.
- [66] F. Siebenhühner, S. H. Wang, J. M. Palva, and S. Palva. Cross-frequency synchronization connects networks of fast and slow oscillations during visual working memory maintenance. *Elife*, Vol. 5, , 2016.
- [67] J. P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela. Measuring phase synchrony in brain signals. *Hum Brain Mapp*, Vol. 8, No. 4, pp. 194–208, 1999.
- [68] P. A. Tass. Stimulus-locked transient phase dynamics, synchronization and desynchronization of two oscillators. *Europhysics Letters*, Vol. 59, No. 2, pp. 199–205, 2002.
- [69] P. A. Tass. Transmission of stimulus-locked responses in two oscillators with bistable coupling. *Biol Cybern*, Vol. 91, No. 4, pp. 203–211, 2004.
- [70] P. Tass, M. G. Rosenblum, J. Weule, J. Kurths, A. Pikovsky, J. Volkmann, A. Schnitzler, and H. J. Freund. Detection of $n : m$ phase locking from noisy data: Application to magnetoencephalography. *Phys Rev Lett*, Vol. 81, No. 15, pp. 3291–3294, 1998.
- [71] O. David and K. J. Friston. A neural mass model for meg/eeg: coupling and neuronal dynamics. *Neuroimage*, Vol. 20, No. 3, pp. 1743–1755, 2003.
- [72] B. H. Jansen and V. G. Rit. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol Cybern*, Vol. 73, No. 4, pp. 357–366, 1995.
- [73] Y. Kuramoto. *Chemical oscillations, waves, and turbulence*, Vol. 19. Springer Science & Business Media, 2012.
- [74] B. Kralemann, L. Cimponeriu, M. Rosenblum, A. Pikovsky, and R. Mrowka. Uncovering interaction of coupled oscillators from data. *Phys Rev E*, Vol. 76, No. 5 Pt 2, p. 055201, 2007.
- [75] B. Kralemann, L. Cimponeriu, M. Rosenblum, A. Pikovsky, and R. Mrowka. Phase dynamics of coupled oscillators reconstructed from data. *Phys Rev E*, Vol. 77, No. 6 Pt 2, p. 066205, 2008.
- [76] B. Kralemann, A. Pikovsky, and M. Rosenblum. Reconstructing phase dynamics of oscillator networks. *Chaos*, Vol. 21, No. 2, p. 025104, 2011.

- [77] K. Ota and T. Aoyagi. Direct extraction of phase dynamics from fluctuating rhythmic data based on a bayesian approach. *arXiv preprint, arXiv:1405.4126*, 2014.
- [78] W. D. Penny, V. Litvak, L. Fuentemilla, E. Duzel, and K. Friston. Dynamic causal models for phase coupling. *J Neurosci Methods*, Vol. 183, No. 1, pp. 19–30, 2009.
- [79] M. G. Rosenblum and A. S. Pikovsky. Detecting direction of coupling in interacting oscillators. *Phys Rev E Stat Nonlin Soft Matter Phys*, Vol. 64, No. 4 Pt 2, p. 045202, 2001.
- [80] I. T. Tokuda, S. Jain, I. Z. Kiss, and J. L. Hudson. Inferring phase equations from multivariate time series. *Phys Rev Lett*, Vol. 99, No. 6, p. 064101, 2007.
- [81] G. Buzsaki and X. J. Wang. Mechanisms of gamma oscillations. *Annu Rev Neurosci*, Vol. 35, pp. 203–225, 2012.
- [82] F. C. Hoppensteadt and E. M. Izhikevich. Oscillatory neurocomputers with dynamic connectivity. *Phys Rev Lett*, Vol. 82, No. 14, pp. 2983–2986, 1999.
- [83] F. C. Hoppensteadt and E. M. Izhikevich. *Weakly connected neural networks*, Vol. 126. Springer Science & Business Media, 2012.
- [84] Y. Terada and T. Aoyagi. Dynamics of two populations of phase oscillators with different frequency distributions. *Phys Rev E*, Vol. 94, No. 1-1, p. 012213, 2016.
- [85] Rosenblum M. Pikovsky, A. and J. Kurths. *Synchronization: a universal concept in nonlinear sciences*, Vol. 12. Cambridge university press, 2003.
- [86] C. Bishop. *Pattern Recognition and Machine Learning*. Springer, New York, 2007.
- [87] K. P. Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.
- [88] N. J. Corron. A simple circuit implementation of van der pol oscillator. <http://ccweb.org/documents/physics/chaos/vdp2006.html>, 01-Feb-2010.
- [89] B. Ermentrout. Type i membranes, phase resetting curves, and synchrony. *Neural Comput*, Vol. 8, No. 5, pp. 979–1001, 1996.
- [90] H. Risken. *Fokker-planck equation*. Springer, 1984.
- [91] A. Duggento, T. Stankovski, P. V. E. McClintock, and A. Stefanovska. Dynamical bayesian inference of time-evolving interactions: From a pair of coupled oscillators to networks of oscillators. *Phys Rev E*, Vol. 86, No. 6, 2012.
- [92] T. Stankovski, A. Duggento, P. V. McClintock, and A. Stefanovska. Inference of time-evolving coupled dynamical systems in the presence of noise. *Phys Rev Lett*, Vol. 109, No. 2, p. 024101, 2012.

- [93] T. Stankovski, S. Petkoski, J. Raeder, A. F. Smith, P. V. McClintock, and A. Stefanovska. Alterations in the coupling functions between cortical and cardio-respiratory oscillations due to anaesthesia with propofol and sevoflurane. *Philos Trans A Math Phys Eng Sci*, Vol. 374, No. 2067, 2016.
- [94] T. Stankovski, V. Ticcinelli, P. V. McClintock, and A. Stefanovska. Coupling functions in networks of oscillators. *New Journal of Physics*, Vol. 17, No. 3, p. 035002, 2015.