Induced activity in EEG in response to auditory stimulation

Pegah Tayaranian Hosseini[a,*], Steven Bell[a], Shouyan Wang[b], David Simpson[a]

[a]*Signal Processing and Control Group, Institute of Sound and Vibration Research, University of Southampton, Southampton, UK*
[b]*Suzhou Institute of Biomedical Engineering and Technology, Suzhou, China*

Abstract

The response of the brain to a sensory stimulus may present itself in the electroencephalogram (EEG) as evoked and/or induced activity. While the evoked response is given by peaks and troughs in the signal, time-locked and phase-locked to the stimuli, the induced response is time- but not phase-locked, and can be considered as an increase or a decrease in the power of EEG in a specific frequency band at a specific time range with regard to the stimulus onset. The induced response does not have the same phase following successive stimuli. It is believed that cognition and perception of a stimulus present themselves primarily as the induced response in the EEG. In this paper, the induced response of the brain to auditory speech stimuli is investigated and different approaches to detect induced activity are compared. It is shown that there is an increase in theta and delta power in response to words compared to the baseline, starting around 500 ms after their onset. During this time, there is also an increase in pairwise coherence between the posterior electrodes. In response to tone bursts, a change in pairwise coherence was observed in the beta band starting around 200 ms. To the best of our knowledge, this is the first time such responses have been described using simple protocols without complex stimulus manipulations being involved. Responses in the EEG to speech rather than the more conventional tone-bursts or clicks suggests that it may be feasible to use the EEG as an objective means to demonstrate brain activation to salient real world stimuli. This would be of particular benefit in investigating access to speech in patients who are unable or unwilling to reliably respond to conventional subjective experimental protocols, such as infants.

Keywords: Induced response, Evoked response, EEG, Pairwise coherence, Auditory stimuli

1. Introduction

The evoked response is phase-locked to a stimulus such that peaks and troughs occur at fixed times (latencies) following the stimuli. Coherent (ensemble) averaging can be used to enhance these responses and suppress the usually overwhelming background EEG activity which hides the response in the raw signals [1]. On the other hand, induced responses

---

*Corresponding author at: Signal Processing and Control Group, Institute of Sound and Vibration Research, University of Southampton, University Road, Southampton, SO171BJ, United Kingdom.

Email addresses: `p.t.hosseini@gmail.com` (Pegah Tayaranian Hosseini), `s.l.bell@soton.ac.uk` (Steven Bell), `swang@sibet.ac.cn` (Shouyan Wang), `ds@isvr.soton.ac.uk` (David Simpson)

are defined as time-locked but not phase-locked brain responses to the presented stimuli. Thus, at a specific time following a stimulus, there may be an increase or decrease in the power of the signal. This increase or decrease in amplitude does not consistently involve the same phase angle in all stimulus presentations and that is why it may be cancelled out using time domain ensemble averaging. The induced response is best investigated in the frequency domain using time-frequency analysis of the signal [1].

An evoked response is believed to be the result of coherent firing of neurons in response to the stimulus, whereas an induced response is thought to be related to higher order processes of the brain which are not directly related to the stimulus but are caused indirectly by nonlinear interactions of neurons following the stimulus [1]. The induced response can be linked to cognitive functions of the brain such as perception, attention, and learning [2, 3, 4] which can be considered to be higher-order processes. Different frequency bands can be affected, depending on the task being performed. The gamma band, for example, has been associated with recalling information in learning tasks [2] and was also observed in response to auditory mismatch negativity stimuli [3] while participants counted the deviant tones. Alpha band power has been found to increase in response to flicker stimuli [4] and beta power has been related to visual attention [5].

The auditory induced response of the brain has been studied with specific stimulus types such as perturbed and unperturbed sequences of pure tones [6, 7, 8, 9], repeated identical words [10], targeted words in specifically designed sentences [11], and repeating modified and unmodified sentences (changing the pitch of the last word) [12, 13]. For example, the gamma band is activated in response to pure-tones with temporal perturbations [6], it is also affected when subjects are presented with different types of keywords in sentences [11], and an increase in 40 Hz power has been observed in response to standard but not deviant repeating words [10]. In all these studies, specific presentation protocols were used to generate an attention or cognitive related response and aim to test quite specific hearing tasks and are not usually easy to reproduce. Whether such sophistication is necessary to find induced responses, has remained an open question. To the best of our knowledge, none of these studies has ever targeted the brain's induced response to a group of words with little restriction on their frequency or lexical content, or to repeating tone bursts, as commonly used in evoking the Auditory Late Responses (ALR). Furthermore, there is no study that may be particularly relevant to assessing hearing-impaired patient's access to speech.

The brain's response to repeating pure tones has attracted much attention in audiology and has been shown to be beneficial for objective estimation of the hearing threshold [14, 15, 16]. For the purpose of finding hearing threshold, the brain response is typically analysed in the time domain focusing on the evoked response [14] to different sound levels of pure tones. Considerable research effort has gone into improving the detection rate of this response [17] for different conditions, such as in children with auditory neuropathy spectrum disorder [18]. However, to the best of our knowledge, little research has focused on the presence of induced responses to pure tones [6, 7, 8, 9] and none on induced responses to repeating pure tones without manipulating the presentation protocol. This may be due to the assumption that a simple tone burst stimulus may not provide any higher order perception cues that would lead to the induced response, or because it is presumed that habituation of the brain to a repetitive sound may cause existing induced

responses to vanish rapidly in repetitive stimulus presentations. Because the tone is simple and commonly used in clinical recordings, it may produce some insight into perception abilities of the participant rather than their hearing ability only, which presents itself primarily as an evoked activity. A comparison of induced responses to tone-bursts and words may also provide additional insights into the specific effect of speech.

As mentioned before, there are a limited number of studies on the brain's response to normal speech such as words and sentences [10, 12, 13, 11, 19, 20, 21]. In addition, not all of these studies target the induced response but rather focus on the evoked response to these stimuli. There are some studies on patients including cochlear implant and hearing aid users [22], children with auditory neuropathy spectrum disorder [18], patients who have had a stroke [23], and hidden hearing loss in people who have been subjected to loud noise for too long [24]. The latter group sometimes show normal hearing thresholds when listening to pure tones but still report hearing problems in response to more complicated sounds such as words in quiet or noisy environments, i.e. they can hear the word but they do not understand it. Thus, it is important to investigate if these groups of subjects follow the perception cues in a way similar to that of normal hearing healthy subjects. Testing if their induced activity to speech is the same as normal hearing people could be a starting point for the development of new diagnostic methods.

Recent publications have utilised different methods to investigate the existence and the significance of the induced response in the brain in response to various types of passive or active tasks. For example, time-frequency analyses have been carried out using Fourier Transform [25] or Wavelet Transform [11, 3], normalisations have been performed using the baseline power [26, 27, 28, 3] or the power of the whole time window being investigated [6], and statistical methods such as non-parametric Wilcoxon test [29], Analysis of Variance (ANOVA) tests [30], and bootstrapping [31] have been used to test the significance of the observed response.

Given that it is suggested in the literature that the induced response reflects higher-order processing of auditory stimuli [2, 3, 4, 1], the present work aims to investigate these responses following a relatively simple reproducible protocol involving repeating pure tones and lists of unrelated but commonly used words. The current study also compares different induced response detection methods. In addition to more conventional power analyses, the paper also extends previous work by using coherence analysis for the detection of induced responses to tone and word stimuli and compares the sensitivity of the two approaches. The overarching aim of this work is to improve knowledge and understanding of electrophysiological responses to auditory stimuli as well as investigating different approaches of induced response detection to move closer to objective assessment of patients' access to speech, with a view to improve fitting hearing aids or cochlear implants in subjects who are unable or unwilling to collaborate with subjective testing.

2. Materials and Methods

In this section, the process of recording and preprocessing the EEG data, parameter estimation, and analyses procedures are explained. It should be noted that the electrode placement and data recording protocols of tone-burst and word stimuli are different. The two different protocols are identified by "Pure-Tones (PT)" and "Words (WS)" protocols throughout the paper.

2.1. Participants

Both protocols had the University of Southampton ethics approval and all subjects in both protocols voluntarily consented to take part in the experiments. Their normal hearing and middle-ear health were confirmed by pure tone audiometry ($\leq$20 dB HL in all frequencies) and tympanometry, respectively, before the start of the experiment.

PT Protocol: Ten normal hearing subjects (5 males and 5 females) aged 18 to 30 years old (27.2$\pm$2.5 years) participated in this experiment. No other exclusion criteria were applied in selecting the participants.

WS Protocol: Fifteen right-handed normal hearing native English speakers (8 males and 7 females) aged 18 to 45 years (24.9$\pm$7.4 years) participated in this study. Subjects suffering from tinnitus, ear diseases, epilepsy, or any psychiatric condition and subjects taking any neuroactive drugs were not included in the study.

2.2. Stimulus Presentation

PT Protocol: Tone bursts (1 kHz) of 70 ms length with 10 ms linear rise/fall times were presented monaurally (right ear) to normal hearing adults 160 times at 60 dB nHL, with a 1.4 s inter-stimulus interval. This presentation sequence was repeated three times for each subject in the same session. The stimulus was generated at 22050 Hz sampling frequency and was presented using an Eartone 3-A insert earphone (E.A.R Tone, Indianapolis, USA).

WS Protocol: 100 words (average length 540$\pm$80 ms) from the Arthur Boothroyd (AB) word lists [32] -recorded by the Institute of Hearing Research, UK Medical Research Council- were selected randomly and presented to the right ear every about 2 s at 62 dB SPL. The only controlled parameter in these words was that they were three-phoneme words in the shape of consonant-vowel-consonant such as "ship", "bone", "move", and "fan" and they were phonemically balanced, i.e. the phonemic composition was equivalent to that of everyday speech. The words were recorded by a male native English speaker and digitised at 22050 Hz. The stimulus was presented through a Sennheiser HDA200 (Sennheiser Electronic Corporation, Connecticut, USA) headphone using Presentation® software (Neuro Behavioural Systems, California, USA).

2.3. Data Acquisition

PT protocol: the EEG was recorded using a 65-channel EEG cap (Easycap, Falk Minow Services, Germany) pre-fitted with Ag/AgCl electrodes with the 10-20 system set-up. Two of these electrodes were used as ground and nose-reference electrodes. A schematic of scalp electrode locations is presented in Figure 1.a. Three more recording channels were used to record eye-blinks, ECG, and the same sound stimulus as was presented to the ear. The latter was used subsequently to find the stimulus onset and cut the signal into corresponding epochs. EEG was recorded with a sampling rate of 1000 Hz using the NeuroScan4.3 software and Synamps^2 amplifier (Compumedics Neuroscan, Germany) and filtered in the [0.05 100] Hz band.

WS protocol: the EEG was recorded using a 66-channel EEG cap pre-fitted with Ag/AgCl electrodes with an equidistant map. Electrodes 67 and 68 were placed under the eyes to monitor eye movements and eye blinks. A schematic of scalp electrode locations is presented in Figure 1.b. Two more electrodes were used as ground and nose-reference electrodes, respectively. As the sound was presented using Presentation® software, the stimulus trigger was sent automatically and saved on the recorded EEG file. EEG was

recorded with a sampling rate of 1000 Hz using the NeuroScan4.3 software and Synamps[2] amplifier (Compumedics Neuroscan, Germany) and filtered in the [0.05 200] Hz band at the time of the recording.
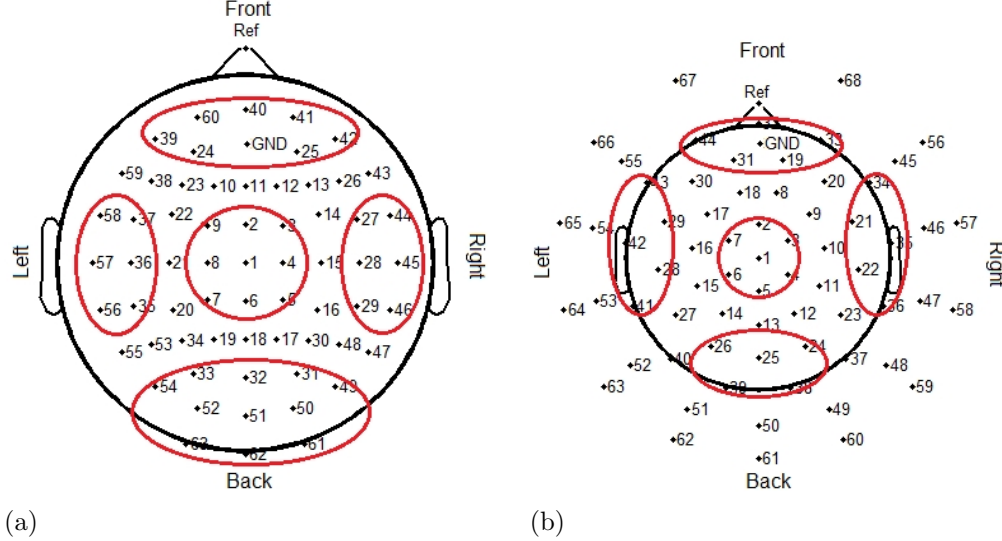

(a)   (b)

Figure 1: Schematic of electrode locations in the electrode cap for recording the brain response in a) PT and b) WS protocols. Red circles group the electrodes in which the power was subsequently averaged prior to analyses. The electrodes that are outside the head circles are the ones that fall below the maximum radius of the head in an axial view.

2.4. EEG Pre-processing

The continuous data were epoched at the onset of each stimulus for both PT and WS protocols. EEG in response to tone bursts was epoched, i.e. segmented, to [-0.5 1] s with respect to the stimulus onset and in response to words this window was [-1 2] s. The time course before the onset of the stimulus was considered as the baseline condition. Independent Component Analysis (ICA) was applied to each epoched dataset and components representing stereotypical artefacts such as ECG and eye-blink were detected visually using the time-domain representation and topographical maps of components, and then were removed [33]. After reconstructing the data from the remaining components, epochs with amplitude values outside the $\pm 100\mu$V range in any channel were removed from further analyses. The data were again visually inspected and additional epochs with non-stereotypical artefacts (caused for example by muscle or head movements) were removed from the data. After this step, two subjects from each protocol were omitted from further computations as the number of their remaining epochs was too small compared to the remainder of the datasets, i.e. less than about 70% of their epochs were left. Note that in some of these procedures, functions from the EEGLAB toolbox [34] were employed.

2.5. Power Analyses

To obtain the power changes after the stimulus onset in the recorded EEG, for each protocol and each subject, the 256-point short time Discrete Fourier Transform (DFT) of each epoch was performed over the whole epoch in 256 ms windows (multiplied by a Hamming window) with a 245 ms overlap. The 256 ms window is centred around

each time point. With a 1000-Hz sampling frequency and 256-point DFT the frequency resolution was 3.9 Hz. To obtain the spectrogram of the evoked response for each dataset, the procedure of [35] was followed. Firstly, the complex DFT values of all epochs were averaged at each time-frequency point, and then the power of this averaged complex value was computed at each point to produce the Evoked Power (EP). To compute the power values of the induced response for each dataset, the power of the complex DFT value in each epoch was computed at each time-frequency point, and then the power was averaged over all epochs at each point, denominated as Total Power (TP). TP is a linear combination of the evoked and the induced power [35]. The Induced Power (IP) was obtained from IP=TP-EP at each time-frequency point.

It is worth noting that we could have removed the evoked activity by subtracting the ensemble average in the time domain from each epoch. However, the adopted approach facilitates visualising the spectrogram of the evoked response and comparing it to that of the induced response. Given that the shape of evoked responses to tone bursts are documented in previous work [14, 15, 16], visualising its spectrogram can act as a quality check for methods used. Also, in the WS protocol, it was not clear from previous work if there was an induced response present, or in which frequency band the signal had to be filtered to observe an evoked response. The evoked power spectrogram helps to assess these issues.

In order to compare the time-varying changes in different frequency bands, that show large changes over time and frequency, normalisation was required. Thus, at each frequency, the evoked and induced power values over time were divided by the average power over the whole epoch at that frequency. This type of normalisation brings all frequency bands into a similar range and enhances the visibility of peaks [6]. It is worth mentioning that normalisation to the baseline power was also tested with simulated data but it was observed to introduce a bias in power values after the onset and increase the false positive rate of significance tests that compared the power values after the onset with those of the baseline. One possible reason for this problem might have been the shape of the power distribution which was not normal but skewed towards zero which, in normalisation, shifted the values of power after the onset to slightly higher values compared to the baseline values. After normalisation, in the PT protocol, the spectra of the three repeated recordings (see section 2.2) were averaged before further analyses to reduce random variations in the recordings.

Normalised power was averaged over groups of electrodes marked in Figure 1 and then, in order to reduce the number of multiple statistical tests, it was averaged within specific time-frequency windows in the spectrogram of the evoked or the induced response. These windows were selected as shown in Figure 2. For different types of stimuli, different time widths were selected which is why time units are missing from this figure. According to the average evoked and induced responses observed for each specific stimulus, 100/200 ms and 200/500 ms windows were selected for evoked and induced powers for PT/WS protocols, respectively. Note that the windows do not overlap and that the onset is always the starting point of the first window after the onset and the ending point of the window before the onset. Obviously, the number of windows over time was different depending on the selected time width and the length of the epoch in each stimulus type, however, the frequency bands were kept constant.
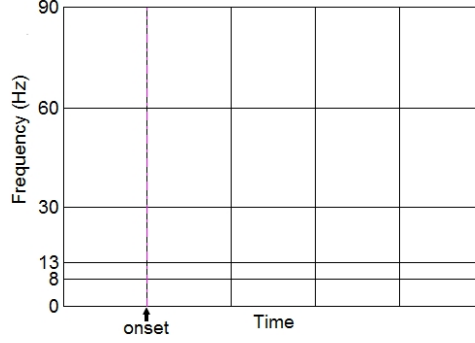
Figure 2: Windows in the spectrogram over which the average power was calculated. The magenta line indicates the stimulus onset. The width of windows over time were selected according to the stimulus type so time units are missing from this figure. 100/200 ms windows were used for the evoked power in response to tones/words and 200/500 ms windows were used for the induced power in response to tones/words.

2.6. Coherence Analyses

It has been shown that the evoked response to a rhythmic stimulation is easier to detect with coherence analyses (coherence with the stimulus) rather than power analyses [36, 37]. This is tested in the current paper for induced activity and in response to both pure tones and word stimuli. Magnitude Squared Coherence (MSC) between two discrete signals $x[k]$ and $y[k]$ is a measure of linear interdependence between two signals and can be calculated from Equation (1) [36]:

$$|\hat{\gamma}_{xy}|^2(f) = \frac{|\sum_{i=1}^{M} X_i^*(f) Y_i(f)|^2}{\sum_{i=1}^{M} |X_i(f)|^2 \cdot \sum_{i=1}^{M} |Y_i(f)|^2}. \tag{1}$$

where $X_i(f)$ and $Y_i(f)$ are, respectively, the DFT of two different EEG channel segments (see Figure 2 and the next paragraph) x_i and y_i where i is the index of an epoch, $*$ denotes the complex conjugate, ˆindicates the estimate, and M is the number of windows the signal has been divided into, which in the case of this study is the same as the number of epochs. The numerator of this fraction computes the cross-spectrum of the two signals and the denominator shows the auto-spectrum of each signal. So, for each segment of the data, the coherence is calculated separately.

To calculate the coherence, the epoched data were first downsampled to 250 Hz to reduce the computation time. Before downsampling, a 5th-order Butterworth anti-alias filter with a cut-off frequency of 100 Hz was applied in the forward and reverse direction (to give zero phase). After downsampling, the ensemble average (evoked response) of EEG in each electrode was subtracted from each epoch (in the time domain) in that electrode so that we were only left with the induced response. Second, the values in each epoch were divided by the standard deviation of that epoch so that all epochs contributed to the coherence equally and to avoid the possibility of one segment with perhaps exceptionally large amplitudes dominating the estimates. Then, depending on the protocol, epochs were segmented into non-overlapping 200/500 ms (for PT/WS protocol) windows using a rectangular window starting 200/500 ms before the stimulus onset (as shown in Figure 2). The two lengths were selected this way so that the results were comparable to those of the induced activity obtained from power analyses.

The pairwise coherence was estimated in each time segment for all channel pairs after calculating the DFT of each segment. The number of points in the DFT was the same as the length of each segment, a frequency resolution of 2 or 5 Hz was thus achieved for words and tones, respectively. The coherence was then averaged in the frequency bands shown in Figure 2 for each channel pair; note that 0 Hz was excluded from averaging in the [0 8] Hz band as coherence at 0 Hz is not meaningful. Tests of significance of any change over time in these values will be described in section 2.8. In the PT protocol, the values of coherence were averaged over the three recordings for each subject.

2.7. Significance Analyses for Power Spectra

To test the significance of power changes after the stimulus onset and thus infer the existence of the evoked or induced activity in the data, Friedman's tests were employed. For this purpose, the average spectrograms over groups of electrodes marked in Figure 2.3 were calculated and the power was again averaged in each time-frequency window (as shown in Figure 2) of the evoked or the induced spectrogram. Friedman's tests were then applied to groups of powers over time (e.g. for tone burst the five groups were [-200 0]ms, [0 200]ms, [200 400]ms, [400 600]ms, and [600 800]ms) in each frequency band. Friedman's test was thus applied five times, once for each frequency band. The number of observations in each group was the same as the number of subjects for that stimulus type. With post-hoc (Tukey's test) analysis, the groups that had powers significantly lower/higher than the baseline power were identified to have significant event-related decrease/increase in the power. Test significance was set at the 5% level.

2.8. Significant Tests for Coherence

Significant coherence changes from the baseline to different segments of data after the stimulus onset would show the existence of event-related activity in those windows after the onset. As the evoked activity had been removed from the data, if a significant change was observed in any segment after the onset, it can be related to the induced activity. Since neighbouring electrodes in EEG data showed high coherence values (close to one, mainly due to the electrical conduction effects rather than neuronal activity)) even without stimulus-related activity in the data, we focused on the change in the coherence over time rather than the actual coherence values between two electrodes.

As for power analyses (see section 2.7), Friedman's tests with 5% significance level were applied to coherence values of each channel pair and each frequency band separately. If a significant change over time was detected in a frequency band, post-hoc (Tukey's test) analyses were applied to find the time windows that had significantly different coherence values compared to the baseline. The post-hoc analysis was used to identify the patterns in connectivity. The significance level here was used primarily as a relative measure of significance, given that with multiple dependent tests, overall significance could not easily be determined. In order to reduce the number of potentially spurious connections identified, a 1% significance level was used at this stage. The significance value of 0.01 was chosen as a compromise: too high a value would give many false positives and confusing results, too low a value would show few connections and thus not be helpful in generating hypotheses about connectivity. Future work with more data will be required to test a smaller number of more targeted hypotheses, guided by this exploratory study.

3. Results

3.1. Power Analyses

PT protocol: As explained in section 2.5, the evoked and induced response powers were calculated, normalised, and averaged for groups of electrodes marked by red circles in Figure 1.a. The spectrograms were then averaged in time-frequency windows as shown in Figure 2 with 100 ms- and 200 ms-long windows for the evoked and the induced power, respectively. Significance tests were then applied according to section 2.7. Figures 3 and 4 show the average time-frequency spectra over groups of electrodes and subjects and the windows marked black show a significant power increase compared to baseline, in the same frequency band.

Figure 3 shows that there are evoked responses in all areas of the brain from approximately 0 to 0.3 s, with a lower frequency response in posterior areas compared to other lobes. Furthermore, there was no significant change ($p<0.05$) of the induced power compared to the baseline in any of the brain areas (Figure 4). Note that other combinations of electrodes were also tested for tone bursts but no significant induced response could be detected in any combination.

WS protocol: As in section 2.5, the evoked and induced power spectra were calculated in response to words, averaged over groups of electrodes, and then averaged in time-frequency windows as in Figures 2 with 200 ms- and 500 ms-long windows for the evoked and the induced power, respectively. Significance analyses were applied on these averaged powers as in section 2.7. Figures 5 and 6 show the average spectra over groups of electrodes and subjects and the windows marked black are those showing a significant power increase compared to the baseline in the same frequency band.

These figures show that evoked and induced responses both occur in lower frequency bands, i.e. delta, theta, and alpha. The evoked response lasts longer (600 ms) in the WS protocol compared to those obtained from the PT protocol (200-400 ms) which may be explained by longer stimuli (540 ± 80 ms) compared to tone bursts. In Figure 6, a significant increase in the induced power in delta and theta bands can be observed starting around 500 ms and lasting up to 1 s after the stimulus onset, thus starting around the average end of words (540 ± 80 ms) and ending a few hundred milliseconds later. The induced response is observed in all areas except the frontal electrodes. Also, there seems to be a decrease in beta power for about 200-300 ms starting around 300-400 ms after the stimulus onset, but this did not reach significance. Still, when a time frequency band of [400 600] ms and [13 30] Hz was selected and compared to the baseline with a Wilcoxon signed-rank test, this decrease from the baseline was also found to be significant ($p<0.05$) in all electrode groups except the frontal ones.

Although detailed results are not presented in the current report, when similar analyses were performed on EEG in response to left ear or binaural word presentations, similar results were obtained, i.e. evoked and induced power increase was observed in frequency bands lower than alpha and induced responses arose later than the evoked responses. Also, a decrease in beta power around the average end time of the words was observed in the induced power.

3.2. Coherence Analyses

PT protocol: As explained in section 2.6, after removing the ensemble average from each epoch of EEG in response to tone bursts, the pairwise coherence was estimated in 200 ms windows without overlap, starting 200 ms before the stimulus onset and then coherence values were averaged in different frequency bands (as in Figure 2).

Significant analyses were then performed according to section 2.8 on these averaged coherence values. Using Friedman's test as explained in 2.8, in each frequency band, channel pairs with significant ($p < 0.05$) coherence change over time were identified. Post-hoc analyses were then applied on the identified pairs to find if their coherence value in any window after the onset was significantly ($p < 0.01$) different from the baseline. In Figure 7, arrows show connections between channel pairs that were indicated in post-hoc analyses as significant (using EEGLAB [34] headshape and the arrow.m file by Erik Johnson [38]). Red/blue arrows indicate increase/decrease in the coherence compared to the baseline. Rows and columns are the same time-frequency windows as in Figure 2. Each subfigure is an axial view of the head. Although it was expected that there could be a few false-positives in each time-frequency window of Figure 7 (given the selected significance levels and multiple non-independent tests), the number of arrows in [13 30] Hz and [0.2 0.6] s windows stand out. Areas of high connectivity also contrast quite clearly with time-frequency regions where no significant connections were found. There thus appears to be clear increase in the coherence in posterior areas in the beta band when listening to pure tones; initially contralateral and after a while, ipsilateral to the side of auditory stimulation (presentation to the right ear). This can be attributed to induced rather than evoked responses, as the latter have been removed from the signal prior to coherence estimation.

WS protocol: Coherence was estimated in 500 ms windows (starting 500 ms before and ending 1500 ms after the onset) for each electrode pair of EEG data in response to words and averaged in different frequency bands (see Figure 2) in each time window.

Significant analyses were then performed in the same way as for the PT protocol, with results shown in Figure 8. As with PT, although it is expected that there are some false-positives in each time-frequency window, the large number of significant arrows especially in the time-frequency window of [0 8] Hz and [0.5 1] s are notable and again this contrasts quite clearly with other time-frequency regions, in which no significant connections were found. There is a notable increase in the coherence in posterior areas (concentrated towards the right side, ipsilateral to the stimulus) when listening to words and this agrees also with the results from induced response power analyses.

It is worth mentioning that in order to further investigate the word stimuli, three datasets were generated by randomly selecting 50 epochs out of the full set of 100 epochs (in response to words) of each subject and the same coherence procedure was repeated on the new datasets. Every time, similar results were obtained, i.e. a significant increase in the coherence in [0 8] Hz band in [0.5 1] s window compared to the baseline coherence. Also, similar results were achieved from EEG data when the stimuli were presented to the left ear and from the data when the stimuli were presented binaurally (details not shown).

4. Discussion

The main aim of this paper was to study the existence of induced activity in EEG data in response to pure tones and words. Moreover, it was investigated if coherence may be more sensitive than power analyses in detecting this activity in the data.

PT protocol: The EEG obtained in response to tone bursts shows a significant increase in the evoked response as evidenced by time-frequency analysis up to 300 ms after the stimulus onset in all electrodes (see Figure 3). The evoked power included higher frequency components in central and frontal areas of the brain and the posterior electrodes showed the lowest frequency components. This result confirms the previous findings that tone bursts usually activate fronto-central areas [39, 40] of the brain more than other areas. Moreover, posterior electrodes appear to have smaller early components (earlier than 100 ms) comparing to other areas. While a change in evoked responses is clearly evident in the signals, no significant induced response is identified in time-frequency analysis following PT stimulation (see Figure 4). This does not exclude the possibility that such changes are present, but if so, they are too weak to be detected by the current approach in this relatively small sample. However, it is clear that they are much less evident than the evoked responses. It has been shown [36] that power analyses are less sensitive in response detection than methods that include phase information such as coherent averaging or coherence, so in the next step the existence of induced activity was investigated using coherence analyses.

Unlike power analyses that can not detect any induced activity in this set of data (see Figure 4), pairwise coherence analyses demonstrate that there is a clear change of activity in the beta band in the [0.2 0.6] s time range (see Figure 7). The role of beta activity is not fully understood in auditory tasks, however, beta de-synchronisation has been shown to be present in some auditory oddball and language perception studies [41, 42, 43]. It is interesting to note that changes in the coherence of beta activity are detected here even with this very simple auditory stimulus protocol, that may not elicit much higher-level processing (and thus induced responses) in the brain.

The difference between the results coming from the two methods (power and coherence) may imply either that the coherence is a more sensitive method than power analyses (as [36] also claims) for detecting induced activity in the signal or that there is little increase or decrease in the power in beta frequency band but that there is phase synchronisation in the [0.2 0.6] s time range which can be picked up by coherence but not power analyses. It has been repeatedly suggested [44] that evoked responses may be the result of synchronisation of the phase in on-going EEG activity rather than additive neuronal activity. This may contribute to the observation that phase may provide a more sensitive means of detecting evoked responses [7, 8] than the power spectrum. Since the evoked response was subtracted prior to the current coherence analysis, the result suggests that there is also increased synchronisation between channels that goes beyond a simple time-locking to the stimulus. One explanation might be that there is jitter in the phase following stimuli [45, 1, 46], but such jitter is coherent between different channels.

WS protocol: The EEG in response to words, according to power analyses, shows a long-lasting (600 ms post-onset) increase in the evoked response in lower frequency bands (lower than alpha) in all electrodes (see Figure 5). Components with only the lower frequencies and with a later onset were noted in posterior electrodes compared to other

brain areas. These results match those of the tone bursts. By looking more closely at the spectra for the evoked response, it can be seen that its power decreases in delta and theta bands after about 300 ms and again increases around 500 ms. As the lengths of the words are around 540 ms ($\pm$80 ms), the increase around 500 ms occurs around the average ending of the words. The peak at this time point may be due to the response of the brain to the offset of words [47].

Moreover, the induced activity in lower frequency bands starts much later than the evoked response, around the average end time of the words (540$\pm$80 ms) and continues for a few hundred milliseconds afterwards (see Figure 6). As the induced response is considered to be related to cognitive tasks of the brain [2, 3, 4], it is expected to start with a delay compared to the evoked response. Only when the word is about to finish, can the brain start the perception process and it takes some time for it to perceive the whole word. It may well be that during this time, the induced response is observed in the EEG. Other studies have shown an increase in theta power [48, 49] after the target word onset in speech perception studies though with different auditory stimulus presentation approaches. Weiss and Rappelsberger [43] and Schack and Weiss [19] also reported an increase in theta coherence in response to recalled versus non-recalled German nouns although they do not seem to have subtracted the evoked response from data before coherence analyses. Also, a pairwise Wilcoxon signed-rank test on WS protocol showed that there was a significant decrease in the beta power for about 200-300 ms starting around 400 ms after the onset in the [13 30] Hz frequency band. As explained before, this frequency band has been associated with language perception in previous studies [41, 42, 50].

Furthermore, coherence estimation of the EEG in response to words showed that there is a very notable increase in the coherence in the [0 8] Hz frequency band mainly in the [0.5 1] s time range (see Figure 8). These changes were seen in electrode pairs falling behind the coronal plane tending towards the right hemisphere, i.e. ipsilateral to the presentation ear. The main speech and language related areas close to these electrodes are the Wernicke's area and bilateral superior temporal gyri that encompass auditory cortices. These results are consistent with those obtained from power analyses, but rather more striking.

In additional data collection (not presented in this paper), both power and coherence analyses were repeated for EEG data recorded from left ear and binaural word presentation. Similar analyses were also applied on data generated by randomly selecting epochs from EEG data in response to words presented to the right ear. In all these cases (left ear, binaural, and random epochs), the results were similar to those obtained from the original EEG recorded from right ear presentation and were consistent with the results shown in this paper, i.e. an increase in power and coherence in the [0 8] Hz frequency band and a decrease in beta power in temporal, central, and posterior electrodes with an ipsilateral tendency in monaural stimulation.

5. Conclusion

The current paper investigates the existence of induced activity in EEG in response to tone bursts and words using power analyses and coherence estimation. To the best of our knowledge, this is the first time that the induced response has been identified with a simple protocol consisting of tones and words, rather than more sophisticated protocols

specifically involving higher order processings, such as word recognition or memory retrieval. Even in these simple protocols, a clear evidence of induced responses was noted. Coherence proved to be able to detect responses in multi-channel EEG following tone-bursts that were not accounted for by evoked activity and time-locking to the stimulus. Power-spectral analysis, on the other hand, did not show significant changes in tone bursts. With word-stimuli, induced responses were evident clearly using both power spectra and coherence changes. The results for word stimuli suggest the potential for their use in objectively assessing hearing and speech perception beyond the conventional approaches based on the audiogram and evoked responses. While it clearly appears that coherence is more sensitive than power-spectral analysis in detecting induced brain responses to stimulation, this result should be interpreted with caution. Changes in between-channel coherence may occur as a result of re-organisation of spontaneous EEG activity, synchronised across channels and with the stimulus, rather than additional neuronal activity, usually assumed when discussing induced activity.

The observed latencies of induced responses in this paper are in agreement with previous work suggesting that induced responses probably reflect higher-order processing of the brain in normal hearing people. It is thus not surprising that they are more evident in response to words than tone bursts. The findings of this study thus may pave the way to develop objective measurement of speech process in the brain. The next step is to find a way to link the observed activity to the amount of speech being understood, i.e. access to speech. This type of research may be of great interest in investigations involving vulnerable subjects who are unable or unwilling to respond reliably to the speech being presented to them. In particular, they may be helpful in developing methods for the validation/optimisation of hearing aid fitting in subjects who report that they can hear the presented speech but they may not understand it and also in finding out why this happens. In follow-on work we are also planning to assess changes in speech perception in response to controlled word stimuli (as opposed to this research with little control over words) and also extend the research to EEG data recorded following the fitting of cochlear implants.

Acknowledgement

[1] O. David, J. M. Kilner, K. J. Friston, Mechanisms of evoked and induced responses in MEG/EEG, Neuroimage 31 (2006) 1580–1591.

[2] T. Gruber, A. Keil, M. M. Müller, Modulation of induced gamma band responses and phase synchrony in a paired associate learning task in the human EEG, Neuroscience Letters 316 (2001) 29–32.

[3] I. Gurtubay, M. Alegre, A. Labarga, A. Malanda, J. Artieda, Gamma band responses to target and non-target auditory stimuli in humans, Neuroscience Letters 367 (2004) 6–9.

[4] U. Graichen, H. Witte, J. Haueisen, Analysis of induced components in electroencephalograms using a multiple correlation method, BioMedical Engineering OnLine 8 (2009) 1–8.

[5] M. Gola, M. Magnuski, I. Szumska, A. Wróbel, EEG beta band activity is related to attention and attentional deficits in the visual performance of elderly subjects, International Journal of Psychophysiology 89 (2013) 334–341.

[6] T. P. Zanto, E. W. Large, A. Fuchs, J. S. Kelso, Gamma-band responses to perturbed auditory sequences: Evidence for synchronization of perceptual processes, Music Perception 22 (2005) 535–552.

[7] L. Fuentemilla, J. Marco-Pallarés, C. Grau, Modulation of spectral power and of phase resetting of EEG contributes differentially to the generation of auditory event-related potentials, NeuroImage 30 (2006) 909–916.

[8] L. Fuentemilla, J. Marco-Pallarés, T. Münte, C. Grau, Theta EEG oscillatory activity and auditory change detection, Brain Research 1220 (2008) 93–101.

[9] C. Haenschel, T. Baldeweg, R. J. Croft, M. Whittington, J. Gruzelier, Gamma and beta frequency oscillations in response to novel auditory stimuli: A comparison of human electroencephalogram (EEG) data with in vitro models, Proceedings of the National Academy of Sciences 97 (2000) 7645–7650.

[10] C. M. Krause, P. Korpilahti, B. Pörn, J. Jäntti, H. A. Lang, Automatic auditory word perception as measured by 40 hz EEG responses, Electroencephalography and Clinical Neurophysiology 107 (1998) 84–87.

[11] J. Obleser, S. A. Kotz, Multiple brain signatures of integration in the comprehension of degraded speech, NeuroImage 55 (2011) 713–723.

[12] M. F. Howard, D. Poeppel, Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension, Journal of Neurophysiology 104 (2010) 2500–2511.

[13] M. F. Howard, D. Poeppel, The neuromagnetic response to spoken sentences: Comodulation of theta band amplitude and phase, NeuroImage 60 (2012) 2118–2127.

[14] G. Lightfoot, V. Kennedy, Cortical electric response audiometry hearing threshold estimation: accuracy, speed, and the effects of stimulus presentation features, Ear Hear 27 (2006) 443–456.

[15] L. Carter, M. Golding, H. Dillon, J. Seymour, The detection of infant cortical auditory evoked potentials (CAEPs) using statistical and visual detection techniques, Journal of American Academy of Audiology 21 (2010) 347–356.

[16] M. Golding, H. Dillon, J. Seymour, L. Carter, The detection of adult cortical auditory evoked potentials (CAEPs) using an automated statistic and visual detection, International Journal of Audiology 48 (2009) 833–842.

[17] K. Ikeda, A. Hayashi, O. Matsuda, T. Sekiguchi, An ignoring task improves validity of cortical evoked response audiometry, Neuroreport 21 (2010) 709–715.

[18] S. He, H. F. Teagle, P. Roush, J. H. Grose, C. A. Buchman, Objective hearing threshold estimation in children with auditory neuropathy spectrum disorder, Laryngoscope 123 (2013) 2859–2861.

[19] B. Schack, S. Weiss, Quantification of phase synchronization phenomena and their importance for verbal memory processes, Biological Cybernetics 92 (2005) 275–287.

[20] S. Weiss, H. M. Muller, The non-stop road from concrete to abstract: high concreteness causes the activation of long-range networks, Frontiers in Human Neuroscience 7 (2013) 526.

[21] J. E. Peelle, J. Gross, M. H. Davis, Phase-locked responses to speech in human auditory cortex are enhanced during comprehension, Cerebral Cortex 23 (2013) 1378–1387.

[22] T. Y. Ching, H. Dillon, A brief overview of factors affecting speech intelligibility of people with hearing loss: implications for amplification, American Journal of Audiology 22 (2013) 306–309.

[23] L. Maiorova, O. Martynova, O. Fedina, A. Petrushevskii, An fMRI study of impairments to speech perception in patients with vascular sensory aphasia, Neuroscience and Behavioral Physiology 44 (2014) 740–747.

[24] R. Tyler, N. Tye-Murray, The relationship between speech perception and psychoacoustical measurements in noise-induced hearing loss subjects, in: R. Salvi, D. Henderson, R. Hamernik, V. Colletti (Eds.), Basic and Applied Aspects of Noise-Induced Hearing Loss, volume 111 of *NATO ASI Series*, Springer US, 1986, pp. 323–333.

[25] F.-B. Vialatte, M. Maurice, J. Dauwels, A. Cichocki, Steady state visual evoked potentials in the delta range (0.5-5 Hz), in: M. Köppen, N. Kasabov, G. Coghill (Eds.), Advances in Neuro-Information Processing, volume 5506 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2009, pp. 400–407.

[26] R. Hannemann, J. Obleser, C. Eulitz, Top-down knowledge supports the retrieval of lexical information from degraded speech, Brain Research 1153 (2007) 134–143.

[27] A. Sanderson, Cortical representation of coded speech, Master's dissertation, University of Southampton, 2010.

[28] E. V. Friedrich, R. Scherer, C. Neuper, Stability of event-related (de-) synchronization during braincomputer interface-relevant mental tasks, Clinical Neurophysiology 124 (2013) 61–69.

[29] M. Karrasch, C. M. Krause, M. Laine, A. Lang, M. Lehto, Event-related desynchronization and synchronization during an auditory lexical matching task, Electroencephalography and Clinical Neurophysiology 107 (1998) 112 – 121.

[30] C. M. Krause, T. Åström, M. Karrasch, M. Laine, L. Sillanmäki, Cortical activation related to auditory semantic matching of concrete versus abstract words, Clinical Neurophysiology 110 (1999) 1371–1377.

[31] B. Telenczuk, S. N. Baker, A. V. M. Herz, G. Curio, High-frequency EEG covaries with spike burst patterns detected in cortical neurons, Journal of Neurophysiology 105 (2011) 2951–2959.

[32] A. Boothroyd, Developments in speech audiometry, British Journal of Audiology 2 (1968) 3–10.

[33] S. Makeig, T.-P. Jung, A. J. Bell, D. Ghahremani, T. J. Sejnowski, Blind separation of auditory event-related brain responses into independentcomponents, Proceedings of the National Academy of Sciences 94 (1997) 10979–10984.

[34] A. Delorme, S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, Journal of Neuroscience Methods 134 (2004) 9–21.

[35] P. Trautner, T. Rosburg, T. Dietl, J. Fell, O. Korzyukov, M. Kurthen, C. Schaller, C. Elger, N. Boutros, Sensory gating of auditory evoked and induced gamma band activity in intracranial recordings, NeuroImage 32 (2006) 790–798.

[36] A. Miranda de Sá, A. Infantosi, D. Simpson, A statistical technique for measuring synchronism between cortical regions in the EEG during rhythmic stimulation, Biomedical Engineering, IEEE Transactions on 48 (2001) 1211–1215.

[37] A. Miranda de Sá, A. Infantosi, D. Simpson, Coherence between one random and one periodic signal for measuring the strength of responses in the electro-encephalogram during sensory stimulation, Medical and Biological Engineering and Computing 40 (2002) 99–104.

[38] Erik Johnson, Draw a line with an arrowhead., 2009. [Online; accessed 4-March-2014].

[39] T. Picton, S. Hillyard, H. Krausz, R. Galambos, Human auditory evoked potentials. i: Evaluation of components, Electrophysiology and Clinical Neurophysiology 36 (1974) 179–190.

[40] R. Knight, S. Hillyard, D. Woods, H. Neville, The effects of frontal and temporal-parietal lesions on the auditory evoked potential in man, Electrophysiology and Clinical Neurophysiology 36 (1980) 112–124.

[41] J. S. Kim, C. K. Chung, Language lateralization using MEG beta frequency desynchronization during auditory oddball stimulation with one-syllable words, NeuroImage 42 (2008) 1499–1507.

[42] M. Hirata, S. Koreeda, K. Sakihara, A. Kato, T. Yoshimine, S. Yorifuji, Effects of the emotional connotations in words on the frontal areasa spatially filtered MEG study, NeuroImage 35 (2007) 420–429.

[43] S. Weiss, P. Rappelsberger, Long-range EEG synchronization during word encoding correlates with successful memory performance, Brain Research Cognitive Brain Research 9 (2000) 299–312.

[44] S. Makeig, M. Westerfield, T. Jung, S. Enghoff, J. Townsend, E. Courchesne, T. J. Sejnowski, Dynamic brain sources of visual evoked responses, Science 295 (2002) 690–694.

[45] C. Tallon-Baudry, O. Bertrand, Oscillatory gamma activity in humans and its role in object representation, Trends in Cognitive Sciences 3 (1999) 151–162.

[46] T. Gruber, C. M. Giabbiconi, N. J. Trujillo-Barreto, M. M. Müller, Repetition suppression of induced gamma band responses is eliminated by task switching, The European Journal of Neuroscience 24 (2006) 2654–2660.

[47] E. Skoe, N. Kraus, Auditory brain stem response to complex sounds: a tutorial, Ear and Hearing 31 (2010) 302–324.

[48] M. C. Bastiaansen, J. J. van Berkum, P. Hagoort, Event-related theta power increases in the human EEG during online sentence processing, Neuroscience Letters 323 (2002) 13–16.

[49] M. Bastiaansen, P. Hagoort, Oscillatory neuronal dynamics during language comprehension, in: C. Neuper, W. Klimesch (Eds.), Event-Related Dynamics of Brain Oscillations, volume 159 of *Progress in Brain Research*, Elsevier, 2006, pp. 179–196.

[50] S. Weiss, H. M. Mueller, "Too many betas do not spoil the broth": The role of beta brain oscillations in language processing, Frontiers in Psychology 3 (2012) 201.
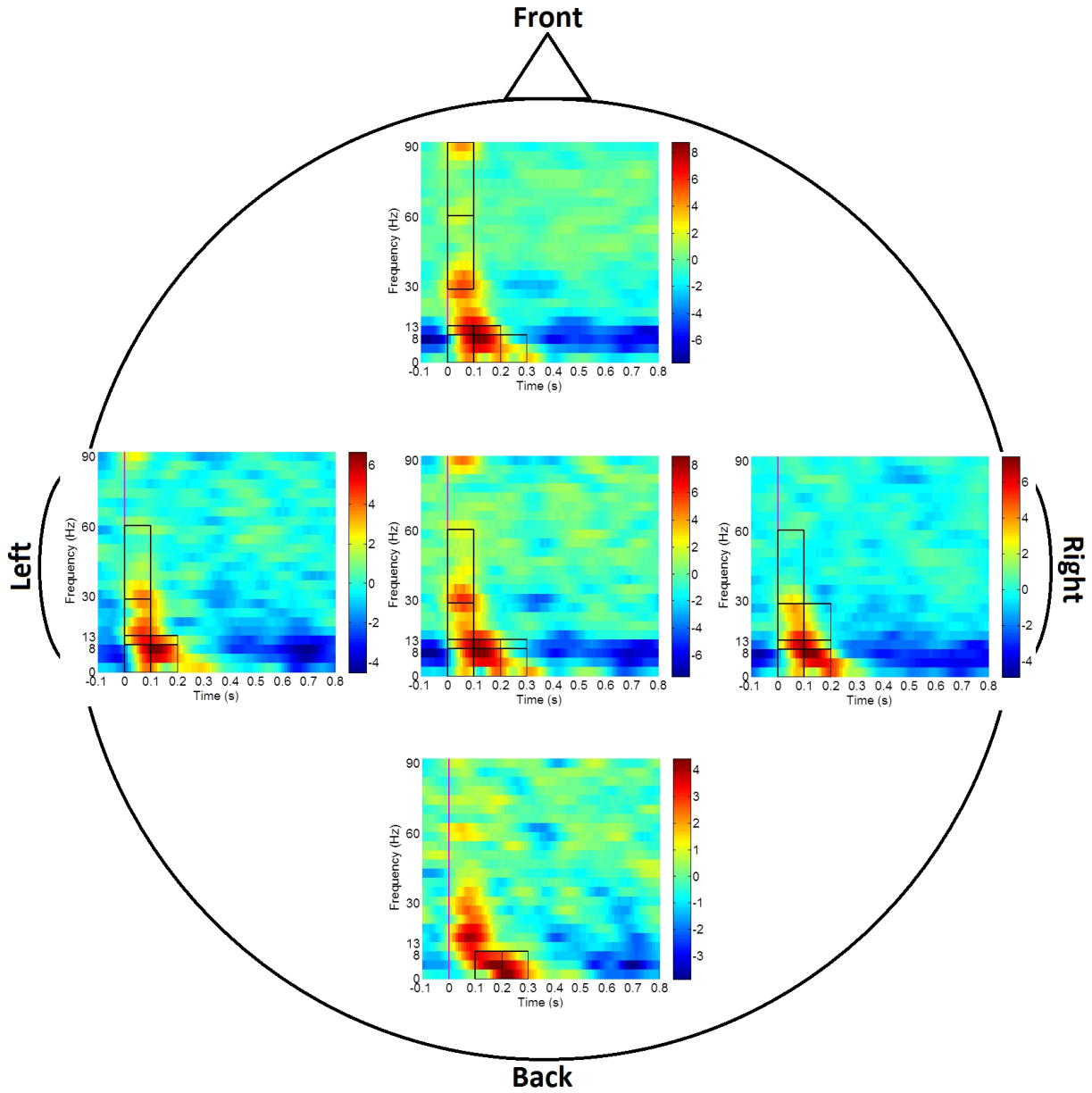
Figure 3: Evoked power spectra in the PT protocol averaged over subjects and groups of electrodes marked by red circles in Figure 1.a. Black windows mark significant (p<0.05) power increase compared to the baseline power in the same frequency band. The vertical magenta line indicates the stimulus onset.
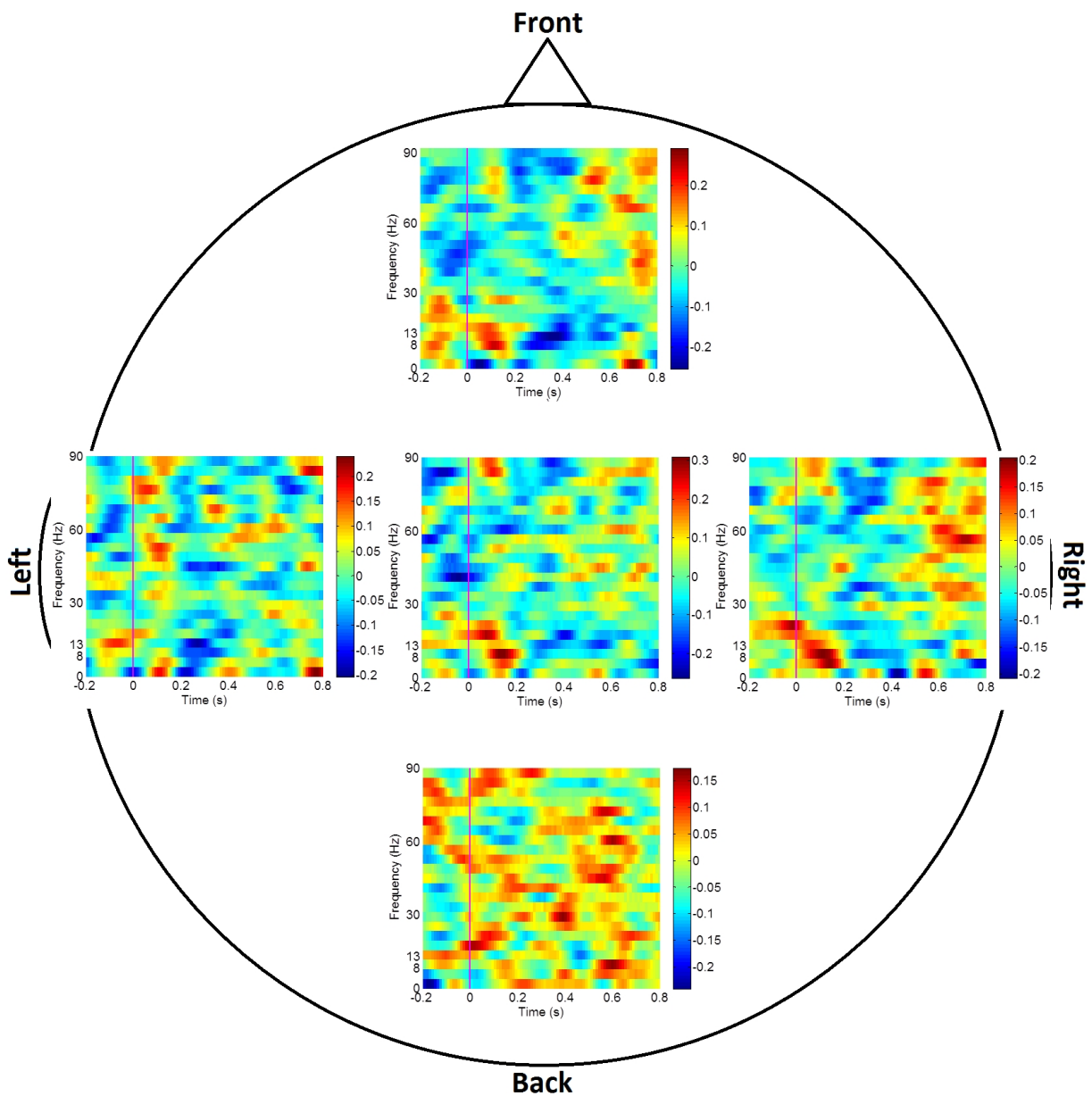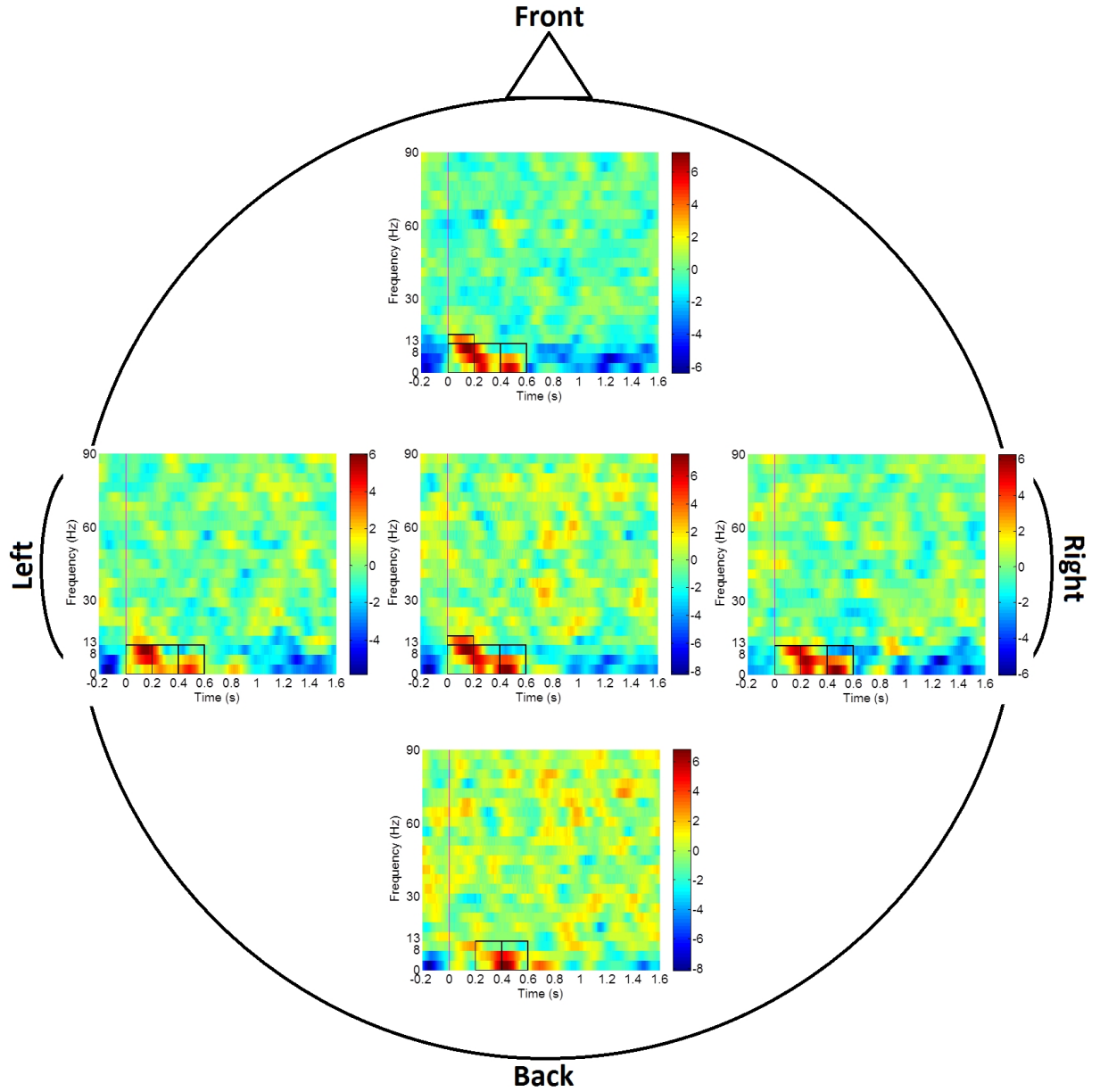
Figure 4: Induced power spectra in the PT protocol averaged over subjects and groups of electrodes marked by red circles in Figure 1.a. No significant changes were observed. The vertical magenta line indicates the stimulus onset.

Figure 5: Evoked power spectra in the WS protocol (right ear presentation) averaged over subjects and groups of electrodes marked by red circles in Figure 1.b. Black windows mark significant (p<0.05) power increase compared to the baseline power in the same frequency band. The vertical magenta line indicates the stimulus onset.
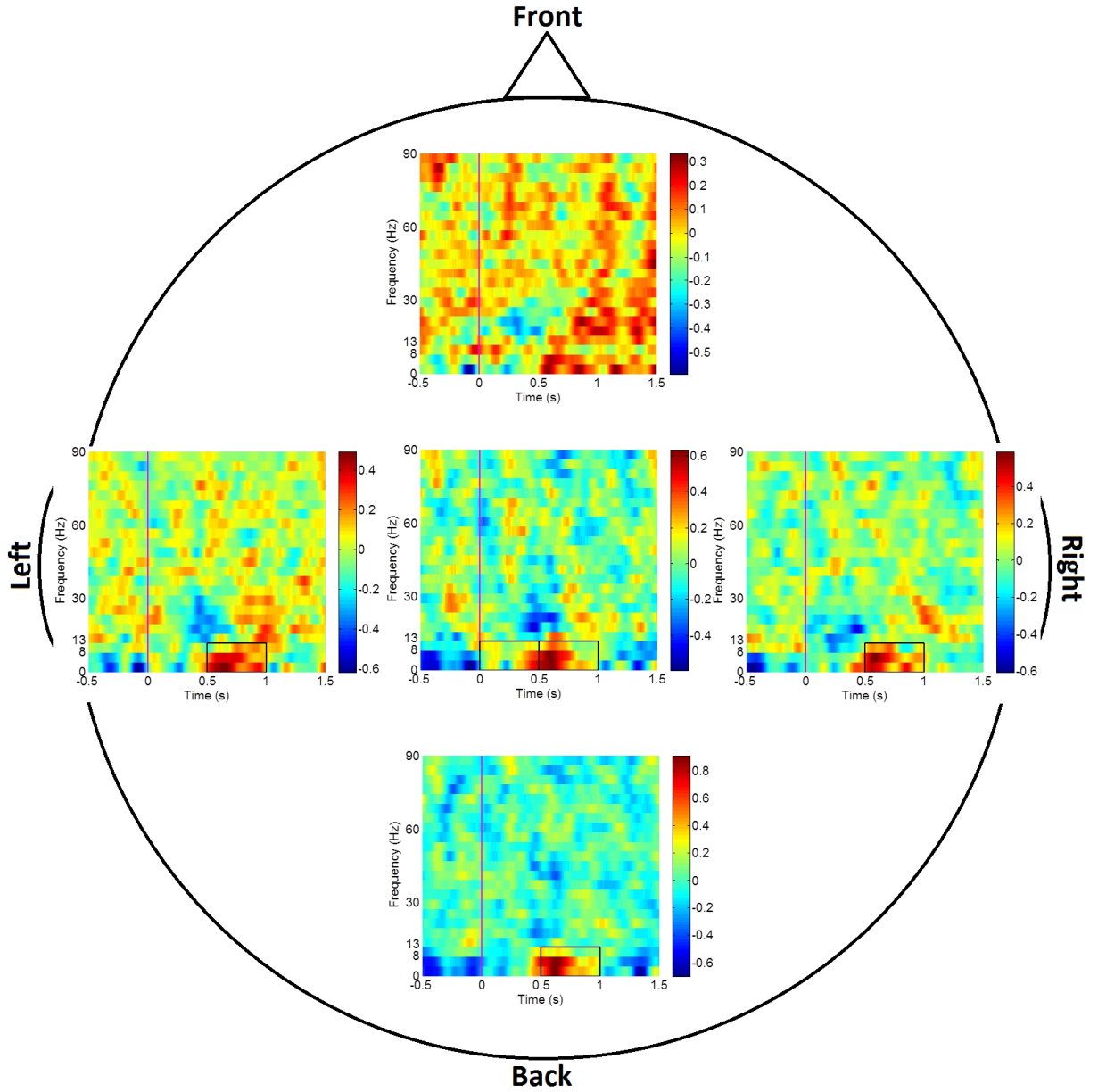
Figure 6: Induced power spectra in the WS protocol (right ear presentation) averaged over subjects and groups of electrodes marked by red circles in Figure 1.b. Black windows mark significant (p<0.05) power increase compared to the baseline power in the same frequency band. The vertical magenta line indicates the stimulus onset.
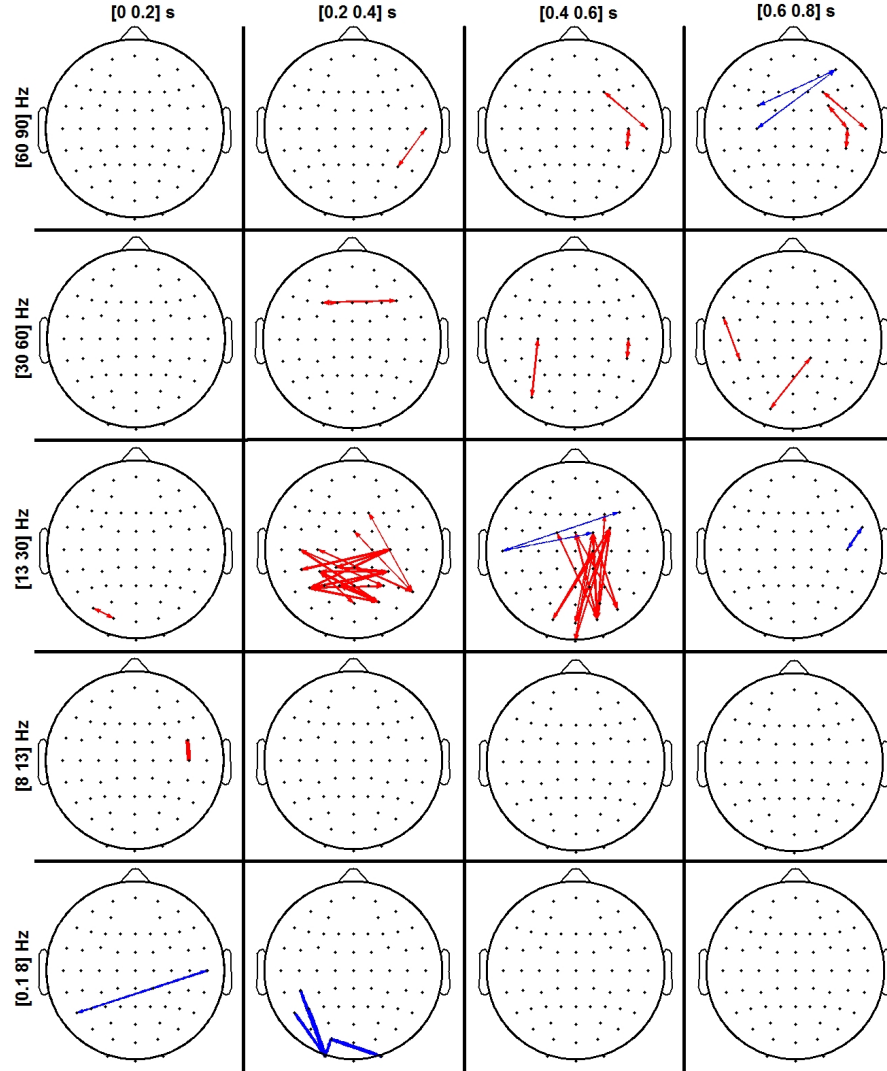
Figure 7: PT protocol: Significant (p<0.01) coherence changes from the baseline in different time-frequency windows are marked by arrows. Red/blue arrows indicate increase/decrease in the coherence compared to the baseline. Thicker arrows indicate larger difference from the baseline. The stimulus onset occurs at zero seconds.
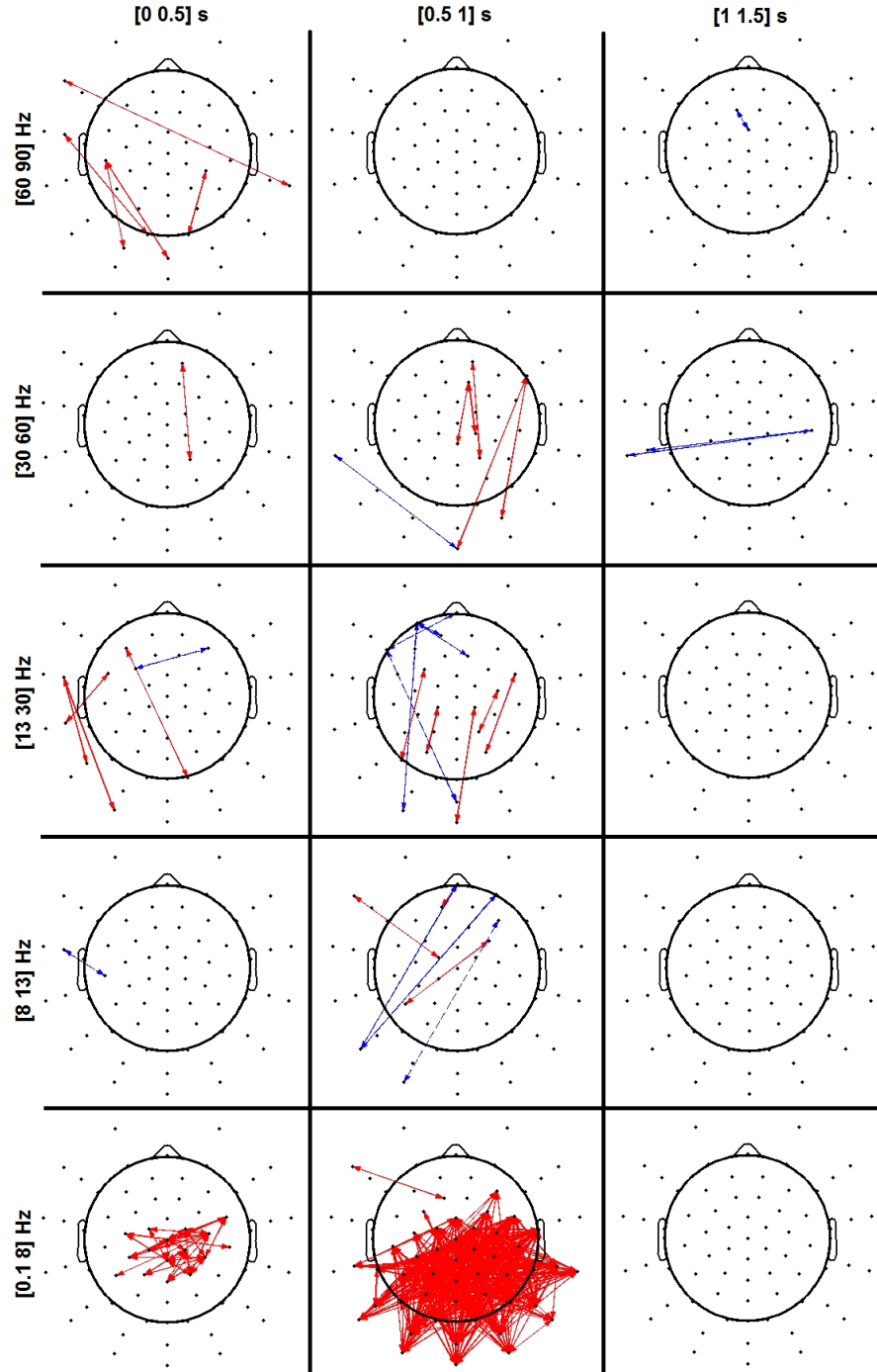
Figure 8: WS protocol (right ear presentation): Significant (p<0.01) coherence changes from the baseline in different time-frequency windows are marked by arrows. Red/blue arrows indicate increase/decrease in the coherence compared to the baseline. Thicker arrows indicate larger difference from the baseline. The stimulus onset occurs at zero seconds.