PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Evaluating potential EEG-indicators for auditory attention to speech in realistic environmental noise

Ehsan Eqlimi(1), Annelies Bockstael(2), Bert De Coensel(1) and Dick Botteldooren(1)

(1)WAVES research group, Department of Information Technology (INTEC), Ghent University (UGent), Belgium
(2)École d'orthophonie et d'audiologie, Université de Montréal, Québec, H3N 1X7, Canada

Abstract

The human brain is remarkably capable of perceiving relevant sounds in noisy environments but the underlying interplay of neurophysiology and acoustics is still being investigated. Cortical processing of these sounds in the brain depends on attentional demand. One of the most important issues is how to identify whether a person is paying attention to the relevant sounds or not. The aim of this study was to explore the potential of single-trial electroencephalography (EEG) indicators to distinguish the cortical representation of three sequential tasks — attentive listening to lectures in background noise, attentive and inattentive listening to background noise alone. Three types of environmental noise, including multi-talker babble, fluctuating traffic and highway sounds were employed as the background during the first task and the stimulus during the second and third tasks. 23 healthy volunteers were exposed to these three tasks while 64-channels EEG signals were recorded. Alpha-band spectral characteristics (peak frequency and power) were investigated as potential indicators of attention and cortical inhibition. Furthermore, based on the hypothesis of self-similarity as excitation-inhibition balance, long-range temporal correlation of alpha-band activity was quantified based on detrended fluctuation analysis. Finally, the hypothesis of speech envelope entrainment of brain activity motivated to estimate the delta absolute power for investigating the attended sound. Considering the participant as a random factor, a linear mixed-effect regression was employed to model the estimated indicators as a function of listening task, EEG channel cluster, and background noise. Strong significant differences were found that support our hypotheses that auditory attention to speech can be observed via EEG-indicators.

Keywords: auditory attention, EEG, cortical inhibition, alpha activity, speech in noise

1 INTRODUCTION

Humans in everyday life are often flooded with a cacophony of target sounds and irrelevant information originating from ecologically valid contexts. Auditory attention plays a key role in sound perception by directing both cognitive and sensory resources to the target sounds. Environmental noise (e.g., multi-talker babble, fluctuating traffic and highway sounds) impairs speech intelligibility and may distract attention away from the narrative. In this regard, one of the most important issues is how to identify whether a person is paying attention to the presented verbal information or not. Multichannel electroencephalography (EEG) at the scalp as a cheap and portable tool with high temporal resolution provides the brain electrical activity, which could be adopted for identifying neural correlates of the auditory attention.

In cognitive sciences, attention is referred to as the mechanism that controls information processing and allocates processing resources to target stimuli [1]. Since attention is not a single and unidirectional process, it could be modulated by externally (bottom-up) driven factors (e.g., sudden attention to a phone ringing) or internally (top-down) factors (e.g., a student focusing on a teacher's lecture in presence of background noise) [2]. Two cognitive control functions are exerted due to attention: (1) excitation and maintenance of task-relevant processes and (2) inhibition of task-irrelevant processes [3].

The alpha-as-inhibition suggests that alpha-band activity (approx. $8-13$ Hz), beyond neural idling state, could reflect a top-down mechanism that inhibits task-irrelevant information [4, 5]. It has been proposed that increased task-related alpha power could be linked to inhibitory control [6]. Indeed, synchronized alpha reflects neural inhibition while desynchronized alpha indicates release from the inhibitory control. Moreover, it has been argued that task-relevant regions exhibit decreased alpha power while task-irrelevant regions exhibit increased alpha power [7]. There are several studies suggesting that the resting-state alpha power could predict the efficiency of cognitive task performance. For example, increased resting-state alpha power in the upper alpha band could be a clue for the good performance of the actual task [8].

In addition to alpha power, several studies have been argued about the relationship between task demand and alpha peak or center of gravity frequency (also known as individual alpha frequency). For instance, it has been demonstrated that alpha peak frequency is linked to the input level (equivalently the number of spiking neurons) such that if the input level increases respect to baseline level, the alpha peak frequency accelerates until the oscillation becomes unstable, being then replaced by slow alpha peak frequency [9, 10, 11]. On the other hand, there are several studies suggesting that with increasing attentional demand and cognitive load, the alpha frequency decreases [1, 12, 13, 14]. This decreased alpha frequency could likely be linked to the unstable state and overloaded attention capacity. Furthermore, it has been suggested [10] that increased alpha peak frequency could be accompanied by a decrease in alpha peak power, which may imply that task-relevant regions could exhibit the accelerated alpha peak frequency during task performance compared to task-irrelevant regions.

Furthermore, it has been proposed that ongoing neural oscillations hold a memory of their own dynamics on fairly long time scales based on long-range temporal correlation (LRTC) idea [15, 16, 17]. This behavior could reflect a critical state due to an optimal balance between excitation (maintaining relevant information) and inhibition (suppressing irrelevant information) states during cognitive tasks. The relationship between LRTC and cognitive performance such as attention, working memory and decision making has been investigated. For instance, [18] showed that decreased alpha-band LRTC compared to the resting-state correlates with better attentional performance.

On the other hand, it has been suggested that low-frequency speech envelope entrainment of brain activity could be robust against different background noise [19]. Moreover, it has been proposed that attention could modulate low frequency (approx. < 8 Hz) cortical representation of speech based on the envelope following [20]. This effect is termed the attentional gain of cortical speech representation. Although the coupling between delta activity (approx. $1-4$ Hz) and speech envelope usually is employed to capture the speech-brain entrainment, in the present study, we have focused on the delta absolute power as a measure of speech envelope following and attentional gain.

The above discussions inspired us to investigate the mentioned hypotheses in a scenario with three listening tasks. For this purpose, twenty-three participants were exposed to three sequential listening conditions while 64-channels EEG signals were recorded. Firstly, all participants were instructed to pay attention to lectures in background noise and then to background noise alone. The background sounds were also presented separately during a resting state where participants were asked not to pay attention to sound. EEG-indicators (alpha peak frequency, alpha peak power, alpha-band long-range temporal correlation, and delta absolute power) were separately modeled as a function of listening task, type of noise and EEG channel cluster to discover the statistically significant differences using linear mixed-effect regression with participant included as a random factor.

2 MATERIALS AND METHODS

2.1 Participants

Twenty-three normal hearing and English speaking volunteers (young adults and approximately same age) with no history of neurological complications participated in this experiment. All participants signed the informed consent and the protocol was approved by the ethical committee.

2.2 Experimental protocol

All participant were exposed to three sequential tasks: (1) lecture attended (LA): focusing attention to the informative lecture in the presence of background noise, mainly an outward-oriented process, (2) background attended (BA): being attentive to the background sound with embedded salient events, which is a sensory-driven process, and (3) background unattended (BUA): not paying attention to any sound. The experimental protocol is schematized in Figure 1.

The order of presentation in lecture attended condition was completely random in both lecture and background noise while assuring the two lectures in silence were not presented in succession. Furthermore, each lecture was presented only once so that each participant was exposed to thirteen different lectures. Each participant completed a written exam on presented lectures and salient events in LA and BA tasks, respectively. They were instructed to keep in mind the information from the lectures during LA task and the number of individual audible salient events (phone ringing, car horn's honk, and emergency siren) during BA task.

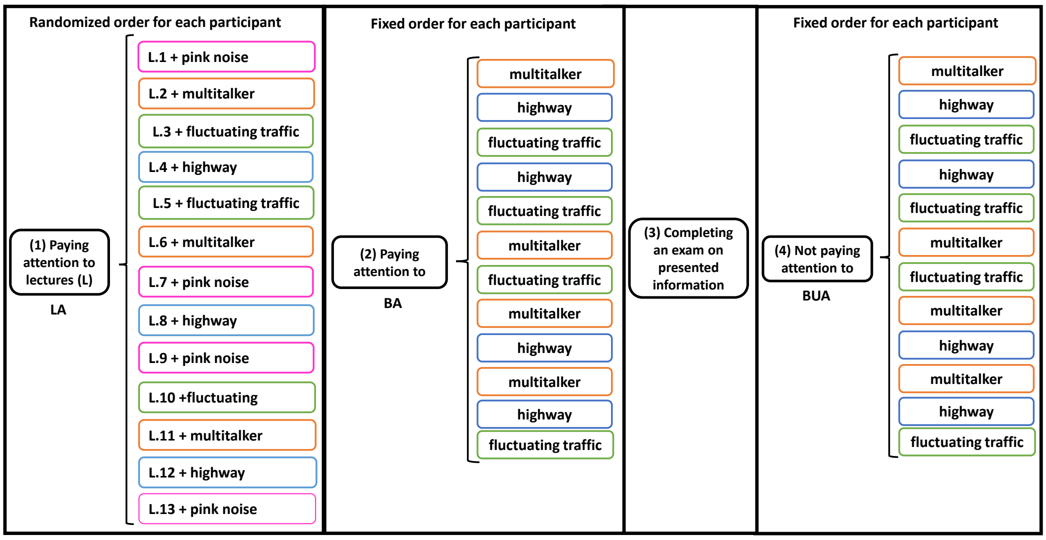

Figure 1. Schematic protocol of attentive & inattentive listening

2.3 Stimuli

During the LA condition, a mixture of target lectures and non-target background noises were presented. Thirteen English lectures were chosen about obscure topics to limit a priori knowledge. The lectures had a duration of about five minutes and were presented at a level of about 68 dB(A). The lectures were read by a native male English speaker with teaching skills and recorded in an anechoic room.

Background noise fragments included three types of environmental noise: (1) HW: highway sound with matching salient sounds of emergency vehicles (2) FT: fluctuating city street sound dominated by individual passage of cars, some birds singing and pass-by of honking car as salient sounds (3) MT: multi-talker babble with matching salient sounds of ring tone of a telephone. The background noise and lecture signals have been mixed with equivalent levels of 63 dB(A) and 68 dB(A), respectively. This assures that the noise does not mask the lecture energetically. The FT noise during LA had as expected more instances of bad SNR (below 0 dB) compared to HW noise, but also more instances of optimal SNR (more than 5 dB). The MT noise is generally more hampering than the HW noise, but compared to the FT noise had fewer instances of bad SNR. As shown in Figure 1, during LA condition, in addition to the mixture of lecture and background noise, four lectures of duration five minutes with low pink (PK) noise (i.e. power spectral density is proportional to $\frac{1}{f}$) at a level of 35 dB(A) were presented. During BA and BUA conditions, each background noise fragment had a duration of about three minutes and was presented at a level of 63 dB(A).

2.4 EEG recordings and preprocessing

Continuous EEG signals were recorded with a BioSemi system from 64 scalp locations (10-20 standard layout). Mono playback was done during the EEG recordings with one loudspeaker at one meter from the participant. The data were recorded with Fpz reference and at a sampling frequency of 2048 Hz. Off-line preprocessing was carried out in EEGLAB [21] package. The data re-referenced to nose electrode and down-sampled to 512 Hz. A finite impulse response (FIR) and hamming windowed-sinc bandpass filter of order 3380 between $0.5-134$ Hz was applied to remove extremely slow drifts and sharp oscillations.

EEG signals were cleaned up in two steps. At first, non-repeating big artifacts were removed based on visual inspection. In the second step, infomax independent components analysis (ICA) with EEGLAB [21] default settings was applied to identify and remove eye blink and movement artifacts.

Since playing audio file typically has a latency of a few milliseconds, the sound was recorded together with the EEG on a free channel and this channel was used to were needed to synchronize with the presented audio signal. For this purpose, at first, the presented audio files were re-sampled to 512 Hz and then the cross-correlation between re-sampled audio signals and recorded sound signals together with the EEG was calculated. The lag corresponding to maximum cross-correlation is the delay in audio files with respect to EEG measurement. To compensate for this delay, all 64-channels EEG signals were shifted with estimated delays.

2.5 EEG signal processing

The EEG-indicators for auditory attention were calculated for each participant per EEG channel and listening task. Firstly, alpha peak frequency and alpha peak power (APF and APP, respectively) were estimated based on multiple signal classification algorithm in order to investigate alpha-as-inhibition. APF is defined as the frequency of alpha-band oscillations which demonstrates maximal power. Second, in order to capture the dynamic properties of EEG signals and estimate the power-law scaling exponent, alpha-band long-range temporal correlation (LRTC) was quantified based on detrended fluctuation analysis. Finally, delta absolute power (DAP, approx. $1-4$ Hz) as a measure of speech envelope following of brain activity was estimated using Welch's power spectral density (PSD) method (a Hamming window of length 512 with 256 samples overlap). The next subsections briefly explain our approaches to estimate APF, APP, and LRTC.

2.5.1 Alpha peak frequency and power based on multiple signal classification

APF and APP could be quantified based on PSD of EEG signals. There are several techniques to estimate PSD such as fast Fourier transform (FFT)-based and subspace-based methods. Subspace-based methods (e.g., root-multiple signal classification (root-MUSIC) [22]) provide higher frequency resolution compared to FFT-based methods (e.g., periodogram method).

Root-MUSIC as a subspace-based method estimates the frequency content of a signal using an eigenspace method. This method assumes that a signal consists of P complex exponentials in the presence of Gaussian white noise. Given a real-valued observation vector, $\mathbf{x} \in \mathbb{R}^{1\times N}$ (N is the number of signal samples), a rectangular Toeplitz matrix, $R_x \in \mathbb{R}^{(N-2P+1)\times 2P}$ could be estimated such that generates an autocorrelation estimate for $\mathbf{x}$ based on a $(2P-1)$th-order prediction error model. If eigenvalues of $R_x^T R_x \in \mathbb{R}^{2P\times 2P}$ are sorted in decreasing order, the eigenvectors corresponding to the P largest eigenvalues (i.e. directions of largest variability) span the signal subspace. The remaining P eigenvectors span the orthogonal space, where there is only noise. A polynomial, $\mathbf{d}$, could be formed consisting of a sum of polynomials given by the product of the noise subspace eigenvectors and the reversed and conjugated version. Finally, the frequency estimates, $\mathbf{f}$, could be determined as the angular positions of P roots of $\mathbf{d}$ which are located nearest the unit circle [23]. In this paper, the preprocessed EEG signals were first bandpass filtered at $8-13$ Hz. Second, above mentioned root-MUSIC algorithm was performed on each filtered EEG channel with $P=2$ as the dimension of the signal subspace. Third, the maximum power (APP in μV^2) and corresponding frequency (APF in Hz) were found.

2.5.2 Long-range temporal correlation based on detrended fluctuation analysis

The processes that do not have a characteristic scale (i.e., scale-free processes) cannot be described completely in terms of the average concept (e.g., peak frequency). There is sufficient evidence that EEG time series exhibit the scale-free dynamics [24]. One of the successful methods to analyze these scale-free signals is long-range temporal correlation (LRTC). LRTC has been developed to quantify how much future dynamics of a signal are influenced by past temporal events [16].

In fractal geometry, LRTC could be interpreted by a self-similarity behavior, which suggests the signal dynamics are similar in different time scales. One of the most common techniques to quantify LRTC is detrended fluctuation analysis (DFA) [25]. The presence of a trend in the signal can cause an overestimation for LRTC. DFA tries to eliminate the trend effect on LRTC. Indeed, DFA is employed to quantify how slowly the autocorrelations of signals decay in power-law behavior (i.e., a, scaling exponent). If $0.5 < a < 1$, the signal likely exhibits strong LRTC and obeys the power-law behavior.

We employed the DFA algorithm which has been suggested in [26] to quantify LRTC for each EEG channel. First, we defined a set of window size, $\mathbf{s}$, which are equally spaced on a logarithmic scale between $2.5-180$ seconds (180 seconds is the duration of signals during BA and BUA). The amplitude envelope of the bandpass filtered signal ($8-13$ Hz), $\mathbf{e}$, based on Hilbert transform was estimated. Then, the cumulative sum of the amplitude envelope, $\mathbf{c}$, was split into a set of n separated time series ($\mathbf{W}^t$) of length $t \in \mathbf{s}$, which have 50% overlap. For each window, $\mathbf{w}_i^t \in \mathbf{W}^t$, the linear trend was removed using a least squares method and obtained the detrended version, $\hat{\mathbf{w}}_i^t$. After calculating the standard deviation of $\hat{\mathbf{w}}_i^t$ i.e., $\sigma_{\hat{\mathbf{w}}_i^t}$, fluctuation function as the mean standard deviation of all windows was computed, $\bar{\mathbf{f}}(t) = \frac{1}{n}\sum_{i=1}^{n} \sigma_{\hat{\mathbf{w}}_i^t}$. Finally, we plotted the fluctuation function, $\bar{\mathbf{f}}(t)$, along t on logarithmic axes. We computed the slope of the trend line (scaling exponent) in the range of $5-18$ seconds (such that filter effect is negligible [26]) using the linear regression as a measure for LRTC.

3 Results

APF, log(APP), log(DAP) and alpha LRTC were modeled based on a linear mixed regression as a function of listening task (LA, BUA, BA), channel cluster (occipital, parietal, central, frontal, left and right temporal), and type of noise (highway (HW), fluctuating traffic (FT), multi-talker babble (MT) and pink (PK)). The participant was included as a random factor. Tables 1 and 2 report the significance relationships for inter-noise comparison of task effect and inter-task comparison of background noise effect on EEG-indicators, respectively. Effect of channel cluster on APF is described in subsection 3.3. In each 4×4 and 3×3 sub-tables, upper triangular elements denote p-values, lower triangular elements denote which noise or task results in higher values in terms of given indicator and main diagonal elements denote which noise or task results in the maximum/minimum values in terms of the given indicator. For example, in Table 1, MT log(DAP) is significantly higher than that of HW during LA because $(1,2)$ and $(2,1)$ elements of matrix corresponding to LA and log(DAP) are < 0.001 and MT, respectively. Note (i, j) indicates the indices of the row and column of the matrix, respectively.

3.1 Inter-noise comparison of task effect on EEG-indicators

The results of statistical analysis for inter-noise comparison of task effect on EEG-indicators were reported in Table 1. During LA in silence (PK), log(DAP) is significantly higher than that of HW and MT. This finding might be due to the strong speech envelope following where there is no background noise. However, there is no significant difference in log(DAP) between PK and FT during LA. On the other hand, log(DAP) during LA is significantly the lowest in HW i.e., the lowest envelope following. During BA task, log(DAP) in FT is slight significantly higher than that of MT. There are no significant differences in terms of log(DAP) for other comparisons during BA. However, during BUA, log(DAP) in MT and FT are significantly the lowest and the highest, respectively. During BUA, log(DAP) difference between FT and HW is slightly significant. Since task demand during BUA is inattentive listening, decreased envelope following are expected especially in MT.

APF during LA task in HW is significantly the highest compared to other types of noise. Since increased attentional demand in an unstable state could tend to reduce APF (as we discussed in Section 1), the drop in attentional performance during LA in HW is the lowest compared to other types of noise. It is worth to note that the drop in attentional performance could be interpreted by the overloaded attentional capacity. During BA task in MT, APF is significantly the lowest due to increased task demand or attentional drop. This finding is consistent with the fact that the multi-talker babble is more hampering and could be more difficult to pay attention to it. Contrary, APF during BUA in MT is significantly the highest, which implies attracting attention when participants try not to attend the multi-talker babble sounds. In fact, during inattentive listening to MT, the attentional demand increases in a stable state and tends to accelerated APF.

Log(APP) during LA in HW is significantly the highest, which reflects the easiest inhibition compared to other noises. Moreover, log(APP) during BUA in MT and HW is significantly the lowest and the highest, implying the most difficult and the easiest inhibition compared to other noises, respectively. These findings support the hypothesis [1] that small upper alpha event-related desynchronization (ERD) values (equivalently increased APP), which may be accompanied by large lower alpha ERD, could reflect easier task performance. During the BA task, there are no significant differences in terms of log(APP) likely due to no need for inhibition.

LRTC of alpha activity during LA task in MT is significantly the highest. Since LRTC could capture self-similarity behavior as excitation-inhibition balance, increased LRTC leads to increased excitation-inhibition balance. As a result, both excitation and inhibition are required to pay attention to lecture in MT. In agreement with previous researches [18], during the attentional task, unlike the resting state, LRTC and attentional performance could have an inverse relationship. During LA in silence (PK), it is expected that excitation is performed more than the inhibition and tend to decreased LRTC. LRTC of alpha activity during BA task in FT is significantly the lowest due to more activity in excitation mode (less excitation-inhibition balance). In the case of BUA (could be equivalent with the resting state) in MT, LRTC is significantly the highest because participants try to inhibit the sound but their attentions are drawn to verbal information (more balance between excitation and inhibition and LRTC).

3.2 Inter-task comparison of background noise effect on EEG-indicators

The results of statistical analysis for inter-task comparison of background noise effect on EEG-indicators were reported in Table 2. Log(DAP) is significantly the highest during MT and FT in LA compared to other tasks due to more speech envelope following. Moreover, log(DAP) during HW in LA is slight significantly higher than that of HW in BUA.

As we discussed in Section 1, decreased APF could be an indicator of the attentional demand in an unstable state due to overloaded attentional capacity. According to Table 2, APF is significantly the lowest during MT in BA compared to other tasks, which implies MT in BA is likely the most difficult condition to perform the task. On the contrary, MT in BUA exhibits the maximum APF compared to other tasks. However, APF during FT noise exhibits an opposite result i.e., it is the maximum for BA and the minimum for BUA likely due to the more attentional drop during BUA compared to LA and BA (i.e., a stable state).

On the other hand, log(APP) is significantly the lowest during MT in BA, implying the most difficult task compared to other tasks. In addition, log(APP) is significantly the highest during HW and FT in BUA, which implies the easier inhibition compared to LA. This finding is in line with [8] where has been suggested that the decreased alpha power during task performance (likely equivalent to LA and BA) could be predicted by high levels of resting-state alpha activity (likely equivalent to BUA).

Alpha LRTC may increase when both excitation and inhibition are evoked. Based on this hypothesis, alpha LRTC is significantly the lowest during MT in BA (less need for inhibition). However, alpha LRTC during MT in BUA is significantly higher than MT in LA. This finding is in agreement with [18], where decreased LRTC has been observed during attentional demand (LA) compared to rest (BUA). The same finding was obtained in FT but alpha LRTC during FT in BA is slight significantly lower than that of LA. For HW case, alpha LRTC values in LA and BUA are significantly the lowest and the highest compared to other tasks, respectively.

3.3 Effect of channel cluster on APF

In connection with the inter-task comparison of channel cluster effect on APF, during LA and BUA tasks due to effort for inhibition, APF in right and left temporal clusters are significantly higher than occipital cluster ($p < 10^{-5}$ and $p < 0.001$, respectively). This finding is in agreement with the hypothesis that APF (unlike APP) increases in the task-related region (i.e., auditory cortex) compared to the occipital region. In the opposite direction, during the BA task likely due to less need for inhibition, occipital APF is significantly higher than that of temporal clusters ($p < 0.001$). On the other hand, regarding the inter-noise comparison of channel cluster effect on APF, during LA task with pink noise (LA in silence), there is no significant difference between occipital and right/left temporal clusters (likely due to less need of inhibition). Contrariwise, during multi-talker, highway and fluctuating traffic due to effort for inhibition, APF in right/left temporal clusters is significantly higher than the occipital region ($p < 0.001$). During BA with three types of noise, occipital APF is significantly higher than that of temporal clusters ($p < 10^{-4}$) likely due to less need for inhibition. During BUA task with multi-talker noise, APF in the right temporal cluster is significantly higher than the occipital APF ($p < 0.001$). However, there is no significant difference between temporal and occipital clusters during BUA in highway and fluctuating traffic.

Table 1. Significance relationships for inter-noise comparison of task effect on EEG-indicators based on a linear mixed-effect model

		LA				BA			BUA		
		MT	HW	FT	PK	MT	HW	FT	MT	HW	FT
Log(DAP)	MT	-	< 0.001	-	0.0381	-	-	0.0872	Min	0.0088	< 0.001
	HW	MT	Min	< 0.001	< 0.001	-	-	-	HW	-	0.0777
	FT	-	FT	-	-	FT	-	-	FT	FT	Max
	PK	PK	PK	-	-	-	-	-	-	-	-
APF	MT	-	$< 10^{-4}$	-	-	Min	$< 10^{-5}$	$< 10^{-5}$	Max	$< 10^{-4}$	$< 10^{-4}$
	HW	HW	Max	$< 10^{-4}$	$< 10^{-4}$	HW	-	-	MT	-	-
	FT	-	HW	-	-	FT	-	-	MT	-	-
	PK	-	HW	-	-	-	-	-	-	-	-
Log(APP)	MT	HW	$< 10^{-4}$	-	-	-	-	-	Min	< 0.001	< 0.001
	HW	-	Max	$< 10^{-4}$	$< 10^{-4}$	-	-	-	HW	Max	0.0084
	FT	-	HW	-	-	-	-	-	FT	HW	-
	PK	-	HW	-	-	-	-	-	-	-	-
Alpha LRTC	MT	Max	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$	-	-	$< 10^{-4}$	Max	$< 10^{-4}$	$< 10^{-4}$
	HW	MT	-	-	$< 10^{-6}$	-	-	$< 10^{-4}$	MT	-	-
	FT	MT	-	-	1.3×10^{-6}	MT	HW	Min	MT	-	-
	PK	MT	HW	FT	Min	-	-	-	-	-	-

Table 2. Significance relationships for inter-task comparison of background noise effect on EEG-indicators based on a linear mixed-effect model

		MT			HW			FT		
		LA	BA	BUA	LA	BA	BUA	LA	BA	BUA
Log(DAP)	LA	Max	$< 10^{-6}$	$< 10^{-6}$	-	-	0.0295	Max	$< 10^{-5}$	$< 10^{-5}$
	BA	LA	-	3.27×10^{-6}	-	-	-	LA	-	-
	BUA	LA	BA	Min	LA	-	-	LA	-	-
APF	LA	-	$< 10^{-10}$	$< 10^{-10}$	-	-	$< 10^{-4}$	-	0.00165	< 0.001
	BA	LA	Min	$< 10^{-10}$	-	-	$< 10^{-4}$	BA	Max	< 0.001
	BUA	BUA	BUA	Max	LA	BA	Min	LA	BA	Min
Log(APP)	LA	-	$< 10^{-4}$	-	-	-	$< 10^{-4}$	-	0.0192	< 0.001
	BA	LA	Min	$< 10^{-4}$	-	-	$< 10^{-4}$	LA	Min	< 0.001
	BUA	-	BUA	-	BUA	BUA	Max	BUA	BUA	Max
Alpha LRTC	LA	-	$< 10^{-5}$	$< 10^{-5}$	Min	$< 10^{-4}$	$< 10^{-4}$	-	0.0834	< 0.001
	BA	LA	Min	$< 10^{-5}$	BA	-	$< 10^{-4}$	LA	Min	< 0.001
	BUA	BUA	BUA	Max	BUA	BUA	Max	BUA	BUA	Max

4 DISCUSSION AND CONCLUSIONS

In this study, we aimed to identify the potential EEG-indicators for auditory attention. For this purpose, 64-channels EEG signals of twenty-three participants were recorded under three tasks—attentive listening to lectures in background noise, attentive and inattentive listening to background noise alone. The background noises included multi-talker babble, highway and fluctuating traffic sounds. Three hypotheses were investigated—alpha-as-inhibition, self-similarity as excitation-inhibition balance and low-frequency entrainment of brain activity by speech envelope. Frequency and power of alpha peak, alpha-band long-range temporal correlation, and delta absolute power were employed as the EEG indicators for auditory attention based on these hypotheses.

The most distinct results were: (1) alpha peak frequency in left and right temporal clusters increases compared to occipital cluster during conditions that need more inhibition i.e., lecture attended in background noise and unattended multi-talker tasks; (2) decreased overall alpha peak frequency could capture increased attentional drop so that background unattended task in highway and fluctuating traffic exhibits the lowest alpha peak frequency compared to attentive listening tasks; (3) increased alpha peak power could reflect easier inhibition so that alpha peak power increases during lecture attended and background unattended in highway compared to other noises and decreases during background unattended in multi-talker compared to other noises. These make sense because highway noise is the most monotonous background and multi-talker noise is the most difficult background to inhibit. Furthermore, alpha peak power increases during background unattended in highway and fluctuating traffic compared to other tasks (easier inhibition); (4) decreased long-range temporal correlation of the alpha activity could reflect the attentional performance so that the long-range temporal correlation is the minimum during lecture attended in silence and the maximum during lecture attended in multi-talker; (5) delta absolute power is able to capture speech envelope following during listening so that delta absolute power values during lecture attended in multi-talker and fluctuating traffic noises are significantly the highest compared to other tasks with the same noises.

We revealed that overall alpha peak frequency plays two distinct roles. First, increased attentional demand in a stable state could accelerate the alpha peak frequency (e.g., attentive listening to the lecture in highway noise compared to inattentive listening to highway noise). Second, the peak alpha frequency could be decreased due to increased attentional demand in an unstable state when the attentional capacity is overloaded (e.g., attentive listening to multi-talker compared to other noises). The intriguing remaining questions are whether these two roles of alpha peak frequency could be linked to lower and upper alpha frequency and also whether they could be observed in two distinct brain regions.

Interesting avenues of research which we are currently investigating include analyzing the inter-regional connectivity and also coupling between EEG and speech envelope as the further indicators of auditory attention. Furthermore, analysing the cortical representation of salient events (phone ringing, car horn's honk, and emergency siren) and the onset/offset of speech during stimuli to discover the attention-inhibition mechanism are the interesting subjects for future studies.

ACKNOWLEDGEMENTS

This study is part of DUCK project and has received research funding from Belgium Special Research Fund (Bijzonder Onderzoeksfonds, BOF). We gratefully acknowledge the assistance and support of International Laboratory for Brain, Music and Sound Research (BRAMS), Montreal, Canada for collecting the EEG data.

References

[1] Wolfgang Klimesch, Hannes Schimke, and Gert Pfurtscheller. Alpha frequency, cognitive load and memory performance. *Brain topography*, 5(3):241–251, 1993.

[2] Emine Merve Kaya and Mounya Elhilali. Modelling auditory attention. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1714):20160101, 2017.

[3] Michael S Clayton, Nick Yeung, and Roi Cohen Kadosh. The roles of cortical oscillations in sustained attention. *Trends in cognitive sciences*, 19(4):188–195, 2015.

[4] Wolfgang Klimesch, Paul Sauseng, and Simon Hanslmayr. EEG alpha oscillations: the inhibition–timing hypothesis. *Brain research reviews*, 53(1):63–88, 2007.

[5] Andero Uusberg, Helen Uibo, Kairi Kreegipuu, and Jüri Allik. EEG alpha and cortical inhibition in affective attention. *International Journal of Psychophysiology*, 89(1):26–36, 2013.

[6] Ole Jensen, Jack Gelfand, John Kounios, and John E Lisman. Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cerebral cortex*, 12(8):877–882, 2002.

[7] Gert Pfurtscheller and FH Lopes Da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical neurophysiology*, 110(11):1842–1857, 1999.

[8] Wolfgang Klimesch. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, 29(2-3):169–195, 1999.

[9] Andreas Mierau, Wolfgang Klimesch, and Jérémie Lefebvre. State-dependent alpha peak frequency shifts: Experimental evidence, potential mechanisms and functional implications. *Neuroscience*, 360:146–154, 2017.

[10] Axel Hutt, Andreas Mierau, and Jérémie Lefebvre. Dynamic control of synchronous activity in networks of spiking neurons. *PloS one*, 11(9):e0161488, 2016.

[11] Jérémie Lefebvre, Axel Hutt, Jean-François Knebel, Kevin Whittingstall, and Micah M Murray. Stimulus statistics shape oscillations in nonlinear recurrent neural networks. *Journal of Neuroscience*, 35(7):2895–2903, 2015.

[12] Mariko Osaka. Peak alpha frequency of EEG during a mental task: Task difficulty and hemispheric differences. *Psychophysiology*, 21(1):101–105, 1984.

[13] Efthymios Angelakis, Joel F Lubar, Stamatina Stathopoulou, and John Kounios. Peak alpha frequency: an electroencephalographic measure of cognitive preparedness. *Clinical Neurophysiology*, 115(4):887–897, 2004.

[14] Rasa Gulbinaite, Tara van Viegen, Martijn Wieling, Michael X Cohen, and Rufin VanRullen. Individual alpha peak frequency predicts 10 hz flicker effects on selective attention. *Journal of Neuroscience*, 37(42):10173–10184, 2017.

[15] J Matias Palva, Alexander Zhigalov, Jonni Hirvonen, Onerva Korhonen, Klaus Linkenkaer-Hansen, and Satu Palva. Neuronal long-range temporal correlations and avalanche dynamics are correlated with behavioral scaling laws. *Proceedings of the National Academy of Sciences*, 110(9):3585–3590, 2013.

[16] Klaus Linkenkaer-Hansen, Dirk JA Smit, Andre Barkil, Toos EM van Beijsterveldt, Arjen B Brussaard, Dorret I Boomsma, Arjen van Ooyen, and Eco JC de Geus. Genetic contributions to long-range temporal correlations in ongoing oscillations. *Journal of Neuroscience*, 27(50):13882–13889, 2007.

[17] Vadim V Nikulin and T Brismar. Long-range temporal correlations in electroencephalographic oscillations: relation to topography, frequency band, age and gender. *Neuroscience*, 130(2):549–558, 2005.

[18] Mona Irrmischer, Simon-Shlomo Poil, Huibert D Mansvelder, Francesca Sangiuliano Intra, and Klaus Linkenkaer-Hansen. Strong long-range temporal correlations of beta/gamma oscillations are associated with poor sustained visual attention performance. *European Journal of Neuroscience*, 48(8):2674–2683, 2018.

[19] Nai Ding and Jonathan Z Simon. Adaptive temporal encoding leads to a background-insensitive cortical representation of speech. *Journal of Neuroscience*, 33(13):5728–5735, 2013.

[20] Jess R Kerlin, Antoine J Shahin, and Lee M Miller. Attentional gain control of ongoing cortical speech representations in a "cocktail party". *Journal of Neuroscience*, 30(2):620–628, 2010.

[21] Arnaud Delorme and Scott Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21, 2004.

[22] Arthur Barabell. Improving the resolution performance of eigenstructure-based direction-finding algorithms. In *ICASSP'83. IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 8, pages 336–339. Citeseer, 1983.

[23] Petre Stoica, Randolph L Moses, et al. Spectral analysis of signals. 2005.

[24] Biyu J He. Scale-free brain activity: past, present, and future. *Trends in cognitive sciences*, 18(9):480–487, 2014.

[25] C-K Peng, Sergey V Buldyrev, Shlomo Havlin, Michael Simons, H Eugene Stanley, and Ary L Goldberger. Mosaic organization of dna nucleotides. *Physical review e*, 49(2):1685, 1994.

[26] Richard Hardstone, Simon-Shlomo Poil, Giuseppina Schiavone, Rick Jansen, Vadim V Nikulin, Huibert D Mansvelder, and Klaus Linkenkaer-Hansen. Detrended fluctuation analysis: a scale-free view on neuronal oscillations. *Frontiers in physiology*, 3:450, 2012.