IJHS

Original Article

The effects of audio stimuli on auditory-evoked potential in normal hearing Malay adults

Ibrahim Amer Ibrahim[1,2,3,4], Hua-Nong Ting[2], Mahmoud Moghavvemi[1,4,5]

[1]Department of Electrical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia, [2]Department of Biomedical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia, [3]Department of Biomedical Engineering, Al-Khwarizmi College of Engineering, Baghdad University, Baghdad, Iraq. [4]Center of Research in Applied Electronics (CRAE), Faculty of Engineering, University of Malaya, Kuala Lumpur, Malaysia, [5]University of Science and Culture, Tehran, Iran

Address for correspondence:
Prof. Dr. Mahmoud Moghavvemi, Department of Electrical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia. Tel: + (603)79676817.
E-mail: mahmoud@um.edu.my

WEBSITE: ijhs.org.sa
ISSN: 1658-3639
PUBLISHER: Qassim University

ABSTRACT

Objectives: The hearing process in the brain is very complicated and hard to solve. However, an understanding of the hearing process is an essential issue and needed in many rehabilitation or treatment applications. This study investigates and compares the effects of simple and complex sounds on latency and amplitude of various event-related potential (ERP) components to male ethnic Malay adults. Comparisons were made with previous studies.

Materials and Methods: Simple and complex sounds were used (pure tones and the naturally produced Malay consonant–vowels [CVs]) to evoke the cortical auditory-evoked potential (CAEP) signals. Moreover, this study analyzed the influence of related CAEP components that are distinct to the selected population and determined which of the ERP components among (CAEP) components is most affected by the two distinct stimuli. Moreover, the study used classification algorithms to discover the ability of the brain in distinguishing CAEP evoked by stimuli contrasts.

Results: The results showed some resemblance between our results and ERP waveforms outlined in previous studies conducted on native speakers of English. On the other hand, it was also observed that the P1 and N2 had a significant effect in amplitude due to different stimulus.

Conclusion: The results show high classification accuracy for the brain to distinguish auditory stimuli. Moreover, the results indicated some resemblance to previous studies conducted on native English speakers using similar tones and English CV stimuli. However, the amplitudes and latencies of the P1 were found to have a significant difference due to stimuli complexity.

Keywords: Auditory event-related potentials, classifications, consonant–vowels, pure tones

Introduction

Hearing enriches our lives and gives us the ability to communicate with others. Hearing enables us to socialize, work, interact, and even relax. It also helps keep us safe by warning us of potential dangers. Moreover, hearing provides us with enormous sources of information. In fact, hearing is essential for us to be able to live and participate in a more meaningful life. Hearing loss can affect our everyday situation and quality of life. Problems with our hearing may lead to feelings of isolation, and even depression.[1]

Auditory information is analyzed through many brain centers, as it flows to the superior temporal gyrus, or auditory cortex, which is the part of the brain involved in perceiving sound. However, auditory-evoked potential (AEPs), or measured brain response with an auditory stimulus, are usually classified based on the latency of the brain waves into early, middle, and late latency responses. While early and middle latency responses involve no cognitive processes by the participants, late responses involve neural processes such as discrimination of pure tones that vary in frequency or complex signals, such as speech.[2] The terms AEP and ERP are interchangeable in this study, with the notion that AEPs are ERPs, and both are both evoked by auditory stimuli.

Auditory long latency ERPs are a series of ERPs occurring between 50 to 500 ms after the onset of stimuli. Long-latency ERPs are comprised of a series of positive and negative peaks labeled P1-N2-P2, along with the mismatch negativity (MMN) and P300 response.[3] The components of the cortical auditory ERP can be used to analyze neural processes involved in discriminating pure tones that vary in frequency or complex signals, such as speech, which also varies along some acoustic continuum, such as the frequency spectrum.[4]

Previous studies have established the effectiveness of auditory ERP as one of the powerful diagnostic tools in examining the human auditory process in the hearing impaired population.[5] Furthermore, quite a few studies showed ERPs can be produced by many speech sounds, including tones[6] and naturally produced CV syllables.[7] In fact, AEP could be classified according to its features. Therefore, many systems use the AEP signal classifications process in their application. AEP signal classifications are used in brain–computer interface applications (BCI),[8] brain hearing problems,[9] and others.[10]

Evidence for the impact of speech characteristics on ERPs were discussed in,[4] where the effects of stimulus frequency and complexity on the components of ERPs during passive listening experiment were detailed. The study used tone bursts in the speech frequency range (400/440, 1500/1650, and 3000/3300 Hz), words (/bad/vs. /dad/), and CVs (/ba/vs. /da/). The authors concluded that the magnitude of N1 and MMN for tones are closely intertwined, with both reflecting the tonotopicity of the auditory cortex. In a different study[11] evoked the potential components that were tested in response to a 10% or 50% frequency increased from 250 or 4000 Hz tones, presented at 500-ms intervals. The results showed that the P50, N100, and P200, and N100, and P200 components exhibited double peaks at bilateral and right temporal sites, and the authors concluded that brain activity changes with respect to frequency changes direction, base frequency, and magnitude.

Oates[12] expounded on the understanding of the subject, where they revealed the significant response difference occurring between the experimental subjects. The study outlined the attenuation effect on the auditory ERPs amplitudes experienced by individuals suffering from sensorineural hearing loss when receiving CVs stimuli compared to the healthy group. Interestingly, sensorineural hearing loss subjects showed prolonged timing response toward ERPs late components (N2/P3) compared to earlier components (N1). In another important study, researchers used two auditory stimuli (1 and 3 kHz) at two intensities (50 and 70 dB sound pressure level [SPL]) and 3 sample groups (18 adults, normal children, and children with sensory processing disorders). The results showed the amplitude and latency measurements to the N1, P2, N2, and P3 components from the averaged ERP for each auditory stimuli.

In the Pettigrew, 2004 research, English consonant–vowels were used on 10 native English speakers to determine the MMN responses to a variety of speech stimuli (/de:/,/ge:/,/deI/"day," and/geI/"gay") in a multiple deviant paradigm. The results showed that MMN responses to a good acoustic speech contrast [d/g] (e.g., "de" vs. "ge", "day" vs. "gay") did not reach to significant effects. However, a significant and larger MMN response was obtained. This MMN was collected at an earlier latency against the real word deviants among non-word standards with a similar initial consonant (i.e., de-day, ge-gay) when compared with the responses across non-word deviants among word standards (day-de, gay-ge). The study results showed that the MMN could be evoked by speech stimuli with large, single acoustic deviances, within a multiple deviant paradigm protocol.[13]

In the area of consonants sounds, the English language experiences fewer sounds with the consonant sounds than the Malay language. This is based on the fact that a higher number of consonants which are attested in the Malay language do not occur in the English language. On the other hands, the Malay language had fewer vowels' sounds because of some vowels attested in English language but absent in Malay.[14] This study answers on if there were a change in the brain AEP recorded from Malay subjects.

This study aims to: (1) Determine the effects of pure tones and naturally produced Malay CVs stimulus on the latency and amplitude of ERP components on ethnic Malay subjects and compare the effects of ERP components with previously reported effects of the widely used English native subjects, (2) determine which ERP component among P1, N1, P2, N2, and P3 is dominant in both the stimuli contrast, and (3) To classify the brain responses to the auditory stimuli using KNN and support vector machine (SVM).

The paper is organized in the following order: Section 1 introduces the study, followed by Section 2, which details the materials and methodology for the experiment, Section 3 explains the results, while Section 4 discusses it, and Section 5 concludes the current study.

Materials and Methods

Participant/subjects

This study was performed on 15 right-handed Malay male adult subjects (mean age = 23.5 year, standard deviation = 3.2 years). They were tested and certified fit by the Otorhinolaryngology (ENT) department in a Medical Centre for normal hearing threshold sensitivity through the routine pure tone audiometry measurement. All of the subjects exhibited normal audiological presentation in both ears (air conduction thresholds 15 dB hearing level (HL) from 125 to 4000 Hz bilaterally, 25 dB HL at 6000 and 8000 Hz, and pure-tone averages (average 20 dB HL from 500 to 4000Hz).

The experimental protocols were approved by the Human Ethics Committee of the University of Malay. The study was carried out on the approval of the Ethics (Medical Ethics Committee, University Malaya Medical Center. Ethics Committee/IRB Reference Number: 1045.22).

Each participant submitted written consent form before the experiments. A simple mini-mental state examination test was also conducted before the experiment to ensure that the subject's mental abilities, memory capabilities, and attention and language proficiency met the required standards.[15]

All participants showed 100% awareness, with healthy normal abilities and responses. However, the test had accounted for more than 15 subjects. Certain subjects' recordings suffered from artifacts, noises, recording calibration, and device setting problems. The study selected the recording signals for subjects that had the most successful, free of artifacts, and noiseless signals. It turned out that only 15 subjects successfully met these criteria.

Stimuli

The study consisted of two disparate types of auditory stimuli; pure tone frequency burst (1 vs. 4 kHz) and speech consonant–vowel (CV) (/ba/vs./da/), presented at approximately 85–90 dB SPL. The tone stimulus was 200 ms in duration, generated by a software program in MATLAB R2013b from (mathwork.com), with (fall time = 10 ms, plateau time = 190 ms) represented at two different frequencies; 1000 and 4000 Hz tone stimuli.[2]

The/ba/and/da/tokens were characterized by their contrasting voiced/voiceless articulatory features of speech, where/ba/ has a higher formant frequency and onset frequency of the formant transitions compared to/da/.[12,16] The speech stimulus was recorded at 44100 Hz sampling rate from the natural speech produced by a female Malay speaker. The CVs were edited into 200 ms in duration by removing the vibration of the vocal cords portion, the final part of the steady-state vowel, and windowing of the offset. The stimuli were presented with a pseudorandomized oddball sequence of 80% standard and 20% deviant presentations, with an interstimulus interval of 800 ± 500 ms, and delivered through Sennheiser HD 428 closed circumaural headphones to both ears. The oddball paradigm is an experimental design procedure, in which a sequences of a repetitive auditory stimuli are frequently interrupted by a deviant stimulus. Deviant stimuli are hidden in a rare occurrence issue among a series of more common standard stimuli. In this study, the pure tone stimulus had standard stimuli of (1 kHz) and deviant stimuli of (4 kHz). Furthermore, the CV stimulus had a standard stimulus of (da) and a deviant stimulus of (ba). The stimuli presented were calibrated at ear level using a KEMAR ear-and-cheek simulator (G.R.A.S. Sound and Vibration, 43AG) and a type 1 integrating sound level meter (Norsonic, nor140).[17]

The tone and CV stimuli contrast were delivered separately and tested in two trials. Each trial consisted of 350 stimuli, that is, 70 deviant stimuli and 280 standard stimuli. Thus, there were 140 deviant stimuli and 560 standard stimuli presented over two trials. The order in which the stimuli were presented ensured that there were 3–5 standard stimuli between each deviant ones. There was no counterbalance for this study, that is, the (1000 Hz/da) stimulus was always the standard, while the (4000 Hz/ba) stimulus was always deviant.

ERP *recording*

Subjects were seated in a comfortable armchair inside a sound-proof chamber. They were instructed to minimize, and if possible eliminate, any eye blinking and muscle movements. The recording was done in various sessions at ~35 min each. To ensure the continuation of passive listening conditions, written short stories were presented throughout the experiment. The recording was done at 500 Hz sampling rate using the wireless Enobio EEG/ERP acquisition system.[18]

Data were recorded from eight Ag/AgCl electrodes mounted on Neoprene EEG cap and located over the following scalp sites: Four electrodes were located on the midline of the head, Fz, Cz, Pz, FPZ, and other four electrodes were located on the left-hand side of the scalp, F7, C3, P7, and T5 (according to the modified International 10–20 System).[19] Electroencephalography activity from each electrode was recorded with the common mode sense active electrode, while the driven right leg passive electrode was referred to the linked mastoid. The recording device Enobio EEG/ERP provided an online filter. The online filter consists of a bandpass filter, with passband (2–40 Hz) second order Butterworth FIR.

ERP *waveform and component analysis*

After ERP data collection, the responses evoked by the standard and deviant stimuli for both stimulus types (pure tone and CV) first went through pre-processing to correlate the baseline drift and filtered offline at 2–30 Hz using second-order Butterworth FIR bandpass filter. These evoked responses were averaged for each trial separately, and then these average evoked responses were averaged another time, with other trials in a session. Some recording sessions contained more than two trials, and some session resulted in bad trials (corrupted by artifacts and noises). The averaged trials were taken from successful runs free from artifacts, noises, and clearly evoked auditory ERP signals. This averaging process was done for each used electrode, separately. However, the standard average responses excluded responses to the stimulus occurring immediately after the deviant stimulus and vice versa for the deviant average response. However, there are many cleaning methods that could be used to denoise the raw signals.[20] The empirical mode decomposition (EMD) technique was used to denoise the recorded EEG signals. All standard and deviant waveforms evoked responses were initially denoised by the EMD technique and inspected visually.[21]

The criteria used to determine ERP response presence or absence were[1] using visual inspection where the ERP is present if individual ERP peaks were larger than the level of the pre-stimulus baseline[2] and using ERP analysis included baseline-to-peak amplitude and latency comparison with a typical standard ERP waveform as described by Näätänen R, McPherson *et al*., and Davies *et al*.,[22-24] where N1 and N2 were defined as the most negative peaks occurring 80–150 ms and 180–250 ms after the onset of stimulus, respectively. P1, P2, and P3 were also defined as the most positive peaks between 55–80 ms, 145–180 ms, and 220–380 ms, respectively.

MMN was defined between 100 and 250 ms. In some trials, P1 and P2 were below the baseline, that is, a negative value, in which case the latency of the peak was measured and the amplitude recorded as missing. All measurements reported here are from responses recorded at the Cz electrode since it was at this electrode where the ERP was largest, while other electrodes were used for the purpose of comparison.[25,26] Furthermore, MMN waveform was extracted using the difference waveform, where the deviant and standard stimuli were used (deviant minus standard).

Segmentation and feature extraction

The averaged cortical auditory-evoked potentials (CAEP) signals were segmented individually into time segments according to the CAEP latencies components. In which, for the P1 (latency window 20–100 ms), N1 (latency window 60–160 ms), P2 (latency window 140–240 ms), N2 (latency window 160–300 ms), and P3 (latency window 240–420 ms) [4,22,27] The latencies were obtained visually and using automated latency detection algorithms. This was done for each stimulus responses separately.

Non-linear feature extraction methods were used in this study such as root mean square Hilbert (RMSHilbert), Kolmogorov–Sinai entropy (KolmogEnt.), Sample Entropy (SampleEnt.), and Approximate Entropy (ApproxEnt.). This was due to the fact that brain neurons are controlled by nonlinear phenomena, such as the threshold and saturation processes. Therefore, its behavior can be regarded as non-linear, and non-linear dynamic analysis can be regarded as an integral approach in detecting mental tasks because it provides more information compared to that reported by traditional linear methods.[28] Approximate Entropy (ApproxEnt.) evaluates the instability of variation in the signal, while Sample Entropy (SampleEnt.) measures the regularity of physiological signals. Furthermore, Kolmogorov–Sinai entropy (KolmogEnt.) evaluates the uncertainty of any signal with respect to time.

Classification of auditory ERP

Classification helps us to understand how a certain auditory brain response is related to other brain responses. Therefore, the classification of auditory response is crucial for patients suffering from hearing loss, as well as people with normal hearing.

The performance of the classifier is tested on different sets of instances. A cross-validation method was used to determine the trained and tested sets. The cross-validation process could be done through multiple approaches (i.e., K-fold cross-validation, Holdout validation, etc.). This study used K-fold cross-validation, with k = 4. This will make 75% of the data in the classification matrix used to develop an automated system to get the features that used to train the classifier, while 25% were used to test the classifier performances. The feature matrix generated from the successful feature extraction method was 240 × 5 elements for both types of stimuli. The matrix contains 4 features (RMSHilbert, KolmogEnt., SampleEnt., and ApproxEnt.), 4 stimuli type (1 kHz, 4 kHz, da and ba) for 15 subjects CAEP's response, and for 5 time intervals, which were (P1, N1, P2, N2, and P3) intervals.

Therefore, for a classification matrix consists of 240 × 5 elements and using 4-fold cross-validation. The training matrix was 180 × 5 and the test matrix was 60 × 5, whereas the test matrix used to evaluate the classification accuracy; classification accuracy given a good indicator of the human brain responses to different auditory stimulus. However, to determine the brain's ability to distinguish between the auditory evoked responses resulting from two auditory stimulus contrasts; pure tone frequency burst (1 kHz vs. 4 kHz) and speech consonant–vowel (CV) (/ba/versus/da/), and if there is a significant difference in the ERP to both stimulus.[29] The classifiers used here to classify the brain responses (CAEP signals) into different auditory stimuli. In this study, a pure tone and CV stimuli were used as the auditory stimuli to evoke the brain ERP. A SVM with kernel function to learn and classify and K-nearest neighbor (KNN) classification methods were used to distinguish between the auditory-evoked responses to both stimulus contrasts. The SVM used a kernel trick to transform the data points into a higher dimensional space, and then separate them by a hyperplane at a maximal margin. The KNN determined a testing sample's class by the majority class of the k nearest training samples.[30,31] A feature extraction methods were applied for the EEG signals in the time domain, which were sample entropy, approximate entropy, Kolmogorov entropy, and RMS Hilbert transformation, these features were non-linear features.[32]

In other hands, the classifier's performance were determined using performance parameter (accuracy),[33] defined as:

$$\text{Accuracy} = \frac{\begin{matrix}\text{Number of correctly}\\ \text{classified observation}\end{matrix}}{\text{Number of total observation}} \quad (1)$$

Statistical *analysis*

Statistical t-test was used to determine the significant difference in CAEP components due to both auditory stimuli, and to distinguish of the effective CAEP component related to each stimulus exclusively.

Results

In this study, auditory ERPs were recorded from 15 Malay male adults with normal hearing in response to tone bursts of 1 and 4 kHz and CV speech stimuli of/ba/and/da/at an intensity of (85–90) dB SPL. Both stimuli contrast were presented in an oddball paradigm in passive listening condition. Amplitude and latency measurements were obtained for auditory ERP components from each stimulus. The results indicated that

ERP correlates of Malay language spoken by native speakers have partly different brains' hearing process as in native English speakers.

The experiments were conducted and the results obtained. Only the Cz electrode was selected for further process and analysis, as it was the most significant toward CAEP waveform in response to auditory stimuli. Furthermore, this electrode demonstrates the highest signal-to-noise ratio as opposed to other electrodes.[5] Samples of the average ERP response before EMD and after EMD denoising are shown in Figure 1.

CAEP components

The all averaged CAEP responses for the tone stimulus (4 kHz) are shown in Figure 2. However, a sample of the average ERP waveforms for both contrast stimuli is shown in Figure 3.

The mean and standard deviation of the peak amplitudes and latencies of the P1, N1, P2, N2, and P3 components for each group are shown in Table 1. However, the percentages of appearance for the ERP components in this work were as follows: P3 was present in 100%, N2 was present in 100%, P2 was present in 83%, N1 was present in 70.8%, and P1 in 73% of all average CAEP responses.

In general, the results showed that all the average amplitudes of the CAEP components (P1, N1, P2, N2, and P3) were larger in standard stimuli compared to deviant stimuli for both stimulus

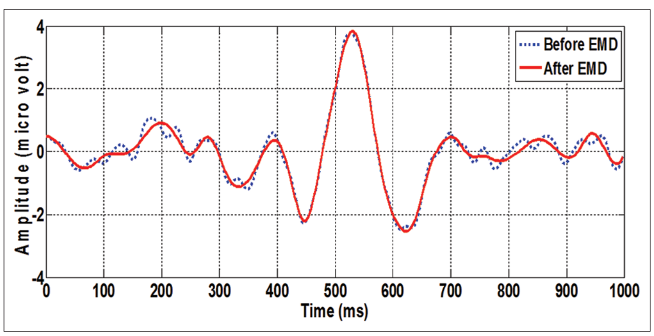

Figure 1: A sample of the average event-related potential response before empirical mode decomposition (EMD) and after EMD denoising

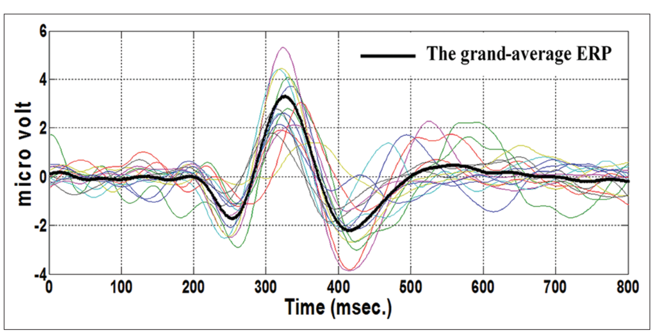

Figure 2: The all averaged cortical auditory-evoked potentials responses for the tone stimulus (standard stimuli, Cz)

types (pure tone and CV). Furthermore, the CAEP components P1, P2, N2, and P3 with response to CV stimulus showed larger latencies in standard stimuli compared to the deviant stimuli. The only exception was N1 wave, which exhibited the opposite pattern. While in the response to frequency stimuli, the CAEP components (P1, N1, P, and P3) showed larger latencies in responses to standard stimuli compared to responses to deviant stimuli. The only exception was N2 wave, which exhibited the opposite pattern. Furthermore, a t-test used for distinguishing the significantly effected CAEP component, which evoked by the auditory stimulus used on Malay subjects. Table 2 shows the t-test results for the amplitudes and latencies of the CAEP components in both stimulus type(s). From Table 2, we can infer that the P1 had a significant effect in amplitude due to CV stimuli, and there was no significant effected CAEP component elsewhere.

Classification

The training and testing were conducted in four times and the classification accuracy is averaged over the 4 trials. Table 3 listed the classification accuracies obtained from the two classifiers that used in this study. The non-linear feature extraction methods that used in this study resulted in a good classification accuracy.

However, the classification accuracy results show that the SVM classification analysis differentiated the groups of stimulus, pure tone frequency stimulus, and CV stimulus with high accuracy, revealing important interrelationship between the components. Furthermore, the kernel type for the support vector classifier was very efficient and separated the two classification groups by a good margin. Thus, it resulted in a good classification accuracy. The performance of the classifier *was assessed* by measuring its accuracy (accuracy = 84.58%).

Discussion

This study is considered as one of the first studies in comparing an auditory-evoked brain potential for different races (English and Malay). Furthermore, it resulted in a classification process for the brain-evoked ERP stimulant through different multiple stimuli. However, the study was conducted as a pilot study to provide a clinical assessment of hearing impairment for aural rehabilitation. The result would have a profound effect on assessing hearing abilities since it is a study toward the aural rehabilitation of the hearing-impaired native Malay population.

However, the major purpose of this study was to investigate the effects of tones and Malay CV on CAEPs response on Malaysians of ethnic Malay descent. ERPs recordings *were obtained* from all participants for both *p*ure *t*ones and CVs stimuli. Overall, the results showed that the components N1, P1, P2, N2, and P3 of CAEP were clearly visible in response to the contrasting stimuli. The results showed some resemblance between the present study and previous studies using tone *an*d

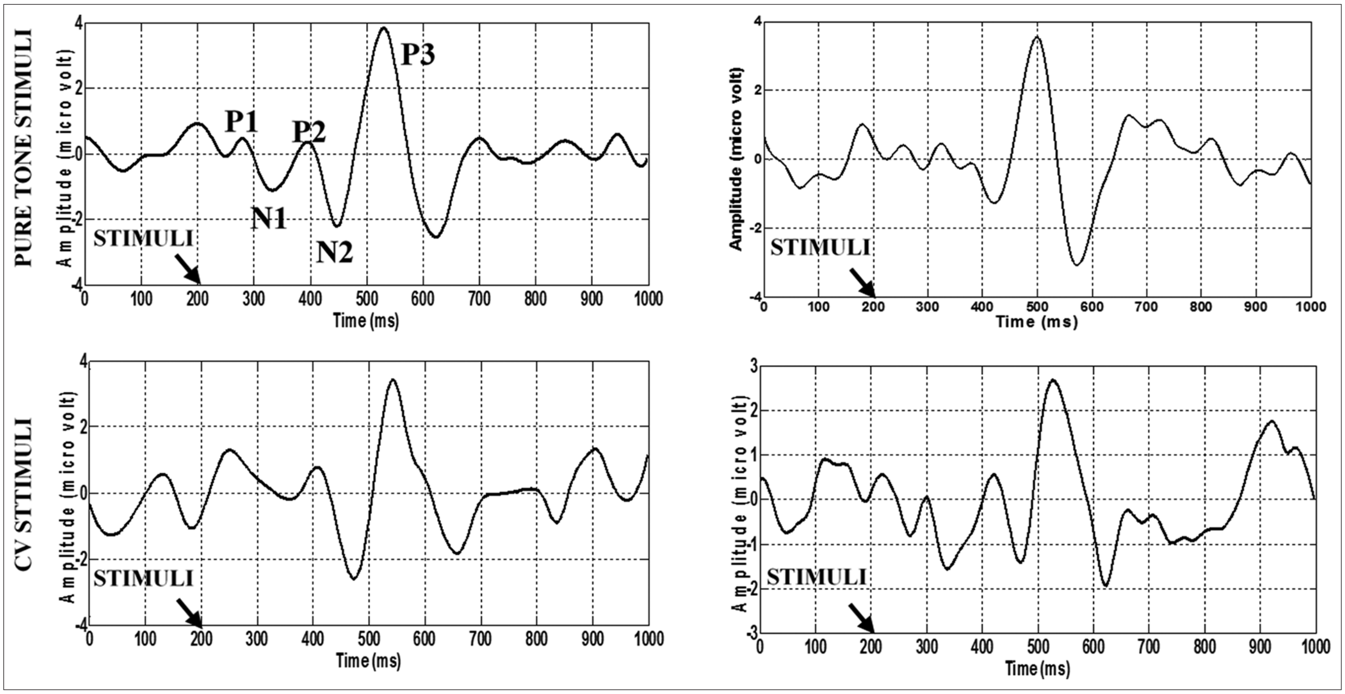

Figure 3: A sample of the average event-related potential waveforms

CV stimuli on native English speakers.[34] as discussed above.

In addition, this paragraph used for comparing between our results and results from previous studies. Our results listed in Table 1 indicate that the latencies of CAEP components (P1, N1, P2, and N2) were larger in comparison with previous studies, *w*hile P3 was smaller in comparison with same previous studies.[4,22-24] This results comparison was true in both *p*ure *t*one stimuli and CV stimuli for the standard or the deviant stimulus. On the other hand, the amplitudes of the CAEP components (N1, P2, and N2) evoked by *p*ure *t*one Stimuli were smaller in comparison with previous studies.[4,22,23] The only exception was a P3 wave, which exhibited the opposite pattern. In fact, this results comparison was true in both *p*ure *t*one stimuli and CV stimuli for the standard or the deviant stimulus. Furthermore, P1 shows a similarity in amplitude as compared with previous studies.

Table 2 shows the *t*-test results for the amplitudes and latencies of the CAEP components in both stimulus type. In which, only P1 had a significant effect in amplitude due to pure tone and CV stimuli and N2 had a significant effect in amplitude due to CV stimuli, which *is* highlighted in "bold." Other CAEP component did not have any significant effect on both amplitude and latency and for both stimulus type (Pure Tone and CV).

The CAEP component (P1) is the early identified component due to different analysis approaches.[35] However, many studies reported that P1 is generated in the primary auditory cortex in the temporal lobe.[36] In addition, *t*his component shows a significant effect since their amplitudes grow by giving more attention to auditory stimuli or by increasing the stimulus interval. The P1 response is an obligatory response, meaning that this component is essentially related to the stimulus characteristics and not to the predictions of the auditory brain system by the listener, *w*hich has *happen* in this study according to the CV stimuli. Although this study used a passive listening protocol.[37]

N1/P2 complex came from combining of N1 and P2; this combining leads the P2 component to share many characteristics from the N1 components.[38] The N1/P2 complex identified in varieties of cognitive tasks including selective attention and stimulus change. However, the CAEP responses were recorded from the passive protocol, where action responses are not required. Furthermore, the N1/P2 complex has a low contrast variability and high repetition. Thus, this study did not show any significant effect on the N1/P2 complex in amplitudes and latencies for both auditory stimuli.[39]

The N2 component *i*s described by higher contrast variability and has many psychological explanations through orienting response and stimulus characteristics. In facts, N2 has responses that vary based on the features of the stimuli, such as modality and trial structure. Thus, in this study, the N2 component show a significant effec*t* due to CV stimuli in the amplitude field.[35]

The P3 waves were extracted in decision-making process. Furthermore, P3 is an endogenous potential since its appearance not related to the physical property of stimuli, but it related to human's reaction from the stimuli.[24] However, *i*n

Table 1: Mean amplitudes in (µV) and latencies in (ms) with standard and deviations for the recorded ERP components with different stimulus (pure tone and CV stimulus) with normal hearing Malay subjects

Stimulus	Amplitude (uV)w									
	P1		N1		P2		N2		P3	
	Stand. [1]	Devi. [2]	Stand.	Devi.	Stand.	Devi.	Stand.	Devi.	Stand.	Devi.
Tone Stim. [3]										
Mean	0.6176	0.2332	0.5103	0.4111	0.4604	0.2349	1.9666	1.5677	3.6142	3.0102
SD[4]	0.2227	0.1829	0.2109	0.2323	0.2282	0.2116	0.7860	0.6522	1.2951	1.1395
CV Stim										
Mean	0.6711	0.1076	0.5384	0.4505	0.2754	0.2069	1.5775	1.1071	2.4544	2.3443
SD	0.3147	0.1018	0.2559	0.1464	0.1864	0.1710	0.6668	0.5538	1.0433	0.7166
Stimulus	**Latency (ms)**									
	P1		N1		P2		N2		P3	
	Stand. [1]	Devi. [2]	Stand.	Devi.	Stand.	Devi.	Stand.	Devi.	Stand.	Devi.
Tone Stim.										
Mean	85.733	85.411	147.73	144.06	195.21	189.46	252.13	254.11	327.73	325.46
SD	14.488	20.141	27.106	29.453	21.431	22.032	10.756	13.373	15.153	18.306
CV Stim.										
Mean	86.533	83.266	140.93	153.41	196.21	203.66	276.26	272.81	356.93	352.53
SD	17.840	10.538	25.443	40.216	59.255	40.872	21.591	32.538	27.680	30.877

[1]Stand.=Standard. [2]Devi.=Deviant. [3]SD: Standard deviation. CVs: Consonant–vowels, ERP: Event-related potential

Table 2: The ***t***-test results for the amplitudes and latencies of the CAEP components in both stimulus type

Stimulus	Amplitude					Latency				
	P1	N1	P2	N2	P3	P1	N1	P2	N2	P3
Pure Tone										
t	1.8929	−0.7779	0.8888	−1.4959	0.9337	0.0442	0.3499	0.7161	−0.4325	−0.2906
p	0.0364	0.2219	0.1912	0.0731	0.1802	0.4825	0.3645	0.2399	0.3343	0.3867
CV										
t	3.8146	−0.8153	0.1518	−1.9999	0.1757	0.6012	−1.0186	−0.4044	0.3388	0.5332
p	0.0003	0.2116	0.4402	0.0278	0.4309	0.2767	0.15926	0.3446	0.3688	0.2991

CVs: Consonant–vowels, CAEP: Cortical auditory-evoked potentials

this study, the CAEP responses were recorded from the passive protocol, where action responses are not required. Thus, this action protocol often not show*s* a significant effect for the P3 component.

Furthermore, the auditory MMN is an automatic neuronal response to changes in events that occur close together in time. However, after extracting the different waveform between the deviant and standard stimuli, the results were expected to show the MMN, but it did not occur. However, the results showed unclear MMN waveforms, which were then rolled out of the rest of analysis. This is due to the stimuli sequence and repeated used in this study, which only had 70 deviant stimuli, whereas it was recommended by some studies that the stimulus sequence continues until at least 100 deviant stimuli is recorded. It *was* also recommended that different waveforms should be made between the same type of stimuli (da as deviant subtracted from da as standard) (MMN review*s*).[27,40]

Finally, another goal of this study was to discover if the brain processes the *p*ure *t*one and CV stimuli separately. The classification accuracy resulted in this study were very high [Table 3]. Thus, this gave us some facts, which are, the high classification accuracy proved that the auditory evoked responses were significantly related to the type of stimulus and significantly different due to the type of stimuli. Furthermore, the high classification accuracy showed that the brain processes both frequency and CV stimuli separately and that difference attentional processes were engaged in responding to CVs compared to tones. This finding will support many systems that used the AEP signal classifications process in their application. For example, AEP signal classifications are used in BCI,[8] brain hearing problems*,*[9] and others.[10]

Effects *of frequency tone on ERP components*

The tone frequency stimuli with a higher frequency (4 kHz) elicited a smaller amplitude response and smaller latencies relative to that of the lower frequency tone (1 kHz), which

Table 3: The performance parameter of classifier

	Classifier			
	KNN		SVM	
	Pure Tone	CV	Pure Tone	CV
Pure Tone	95	25	98	22
CV	18	102	15	105
Total Accuracy	197/240=0.8208		203/240=0.8458	

elicited a larger amplitude responses with bigger latencies. This finding agrees with previous studies. There are already some evidence from previous studies on MEG,[41] PET,[42] and FMRI,[43] which reasoned that the area with neurons, located deep in the brain, responds better to high-frequency sounds, whereas the area with neurons, located superficially, responds better to low-frequency sounds. The only exception to this was the N2 wave, which showed different patterns with lower latency with the lower frequency stimulus. This study used a passive listening protocol*,* and the N2 component could identify in varieties of cognitive tasks including selective attention and stimulus change.[39]

The results also showed unclear P1-N1-P2 complex in some subjects, where sometimes N1 and P1 were missing (the overall mean showed N1 was present in 70.8%, P1 in 73%, and P2 in 85% of all subjects*)*. This might be due to the fact that the P1-N1-P2 complex is sensitive to frequency stimuli[44] and that the P2 amplitude dominates the response compared to P1 and N1 in adults.[45]

P3 was elicited in all subjects, and it had a 100% appearance in both standard and deviant stimuli. Since the results were collected from the Cz electrode, P3 is regarded as P3b, which is usually recorded from the electrodes located over the centroparietal regions of the scalp (such as Cz and Pz sites).[34,46] However, we preferred to use the term P3 in this article. An explanation of why P3 was the most detected component in this study is the fact that it is best elicited when the stimuli are delivered in oddball paradigm manner (for reviews see[46] and[47]).

Effects of CVS on ERP components

As documented in T*abl*e 1, the mean amplitude for all ERP components w*as* larger for responses to standard stimuli compared to the responses to deviant stimuli. One likely explanation for higher amplitude in responses to standard stimuli (/da/) is the spectral differences that exist in these CVs, where the formant frequencies and the onset frequencies of the formant transitions of*/*da*/*are lower than that of*/*ba/. Recently, the authors[48] provided evidence using the FMRI technique on the area of the human cortex, located more superficially to the scalp, which responds better to low-frequency information, whereas the area located deeper in the brain *responds* better to high-frequency information. Therefore, the ERPs were recorded using surface electrodes and CVs with lower frequency energy that would mostly activate the more superficial regions of the scalp, thus producing larger amplitudes than CVs, dominated by higher frequency energies.

The mean latencies for all CAEP components were larger in the responses to standard stimulus compared to the responses to deviant stimulus for all used auditory stimuli. The present results agree with those reported in previous studies. First, we showed the latencies of CAEP responses to (85–90) dB SPL/ d*a/*stimuli being larger compared to responses to*/*b*a*/. Second, the amplitude of CAEP components was larger in*/*da*/*versus/ ba*/*responses. This was due to the fact that brain neurons are responsive to sounds in certain *specialized* brain areas. This response depends on sound frequencies and amplitude contractions values.[49] However, the only exception to this was the N1 wave, which showed different patterns with large latency in deviant CV stimuli as compared with standard CV stimuli. This is due to, N2 are an endogenous potential, and its appearance is not related to the physical property of stimuli, but it is related to human*'*s reaction from the stimuli.[50]

Classification accuracy

The accuracy of the results of the SVM classifier (which was 84.58%, Table 3) indicated that the human brain had a strong ability to distinguish auditory stimuli. However, the high accuracy of the SVM classifier gives us the opportunity to use different auditory stimuli in future to control certain applications using our minds.[51] Furthermore, this classification protocol will support many systems that use the AEP signal classifications process in their application. For example, AEP signal classifications are used in BCI,[8] brain hearing problems*,*[9] and others.[10] In fact, the limitation of sample size utilized in this study resulted in highly precise values for some of the utilized classifiers.

Conclusion

This study investigated the ERP correlates of tones and Malay CV on Malaysian males of ethnic Malay descent. The auditory ERPs were recorded from 15 right-handed adult Malay male subjects (mean age = 23.5 ± 3.2 year) with normal hearing while listening to tone bursts (1 and 4 kHz) and *CV* speech stimuli of*/*ba*/*and*/*da*/*of intensity 85 *dB* SPL. The results showed that components N1, P1, P2, N2, and P3 of ERP were clearly evoked in response to the contrasting stimuli. The results indicated some resemblance to previous studies conducted on native English speakers using similar tones and English CV stimuli. However, the amplitudes and latencies of the P1 were found to have a significant difference due to stimuli complexity. Due to the fact that the auditory stimuli of different languages have different influences compared to native English speakers, this study found an important fact about the influences of auditory stimulus on the CAEP components elected from non-English native speakers. It should also be pointed out that English CVs are commonly used for aural assessment tests, which were not proper for all cases such as non-English native speakers. Hence, the result would have a positive effect in the area of aural rehabilitation conducted in local hospitals. Furthermore, this study is one of the first attempts of classifying an auditory brain response signals. This study faced several limitations. For instance, the sample size was small, and an additional analysis with a large database should be performed in the future.

References

1. FIRST OP. Why is hearing important?: Oticon Canada; 2016. Available from: http://www.oticon.ca/hearing/facts/hearing/why-is-hearing-important.aspx. [Last accessed on 2016 Sep 06].
2. Alain C, Roye A, Arnott SR. Middle-and long-latency auditory evoked potentials: What are they telling us on central auditory disorders. In: Handbook of Clinical Neurophysiology: Disorders of Peripheral and Central Auditory Processing; 2013. p. 177-99.
3. Swink SD. Auditory Event-Related Potentials Recorded During Passive Listening and Speech Production: East Carolina University; 2010.
4. Wunderlich JL, Cone-Wesson BK. Effects of stimulus frequency and complexity on the mismatch negativity and other components of the cortical auditory-evoked potential. J Acoust Soc Am 2001;109:1526- 37.
5. Schröder A, van Diepen R, Mazaheri A, Petropoulos-Petalas D, de Amesti VS, Vulink N, *et al*. Diminished n1 auditory evoked potentials to oddball stimuli in misophonia patients. Front Behav Neurosci 2014;8:123.
6. Wang M, Liu Z, Zhu Q, Liu J, Wang L, Liu H. Influence of a preceding auditory stimulus on evoked potential of the succeeding stimulus. Chin Sci Bull 2004;49:1032-7.
7. Agung K, Purdy SC, McMahon CM, Newall P. The use of cortical auditory evoked potentials to evaluate neural encoding of speech sounds in adults. J Am Acad Audiol 2006;17:559-72.
8. Gao S, Wang Y, Gao X, Hong B. Visual and auditory brain-computer interfaces. IEEE Trans Biomed Eng 2014;61:1436-47.
9. Sriraam N. EEG based automated detection of auditory loss: A pilot study. Expert Syst Appl 2012;39:723-31.
10. Ravan M, Reilly JP, Trainor LJ, Khodayari-Rostamabad A. A machine learning approach for distinguishing age of infants using auditory evoked potentials. Clin Neurophysiol 2011;122:2139-50.
11. Pratt H, Starr A, Michalewski HJ, Dimitrijevic A, Bleich N, Mittelman N. Auditory-evoked potentials to frequency increase and decrease of high-and low-frequency tones. Clin Neurophysiol 2009;120:360-73.
12. Oates PA, Kurtzberg D, Stapells DR. Effects of sensorineural hearing loss on cortical event-related potential and behavioral measures of speech-sound processing. Ear Hear 2002;23:399-415.
13. Pettigrew CM, Murdoch BE, Ponton CW, Finnigan S, Alku P, Kei J, *et al*. Automatic auditory processing of English words as indexed by the mismatch negativity, using a multiple deviant paradigm. Ear Hear 2004;25:284-301.
14. Oshodi B. A cross-language study of the speech sounds in Yorùbá and Malay: Implications for second language acquisition. Issu Lang Stud 2013;2(1):1-12.
15. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189-98.
16. Zoubek L, Charbonnier S, Lesecq S, Buguet A, Chapotot F. Feature selection for sleep/wake stages classification using data driven methods. Biomed Signal Process Control 2007;2:171-9.
17. Sulaiman AH, Husain R, Seluakumaran K. Hearing risk among young personal listening device users: Effects at high-frequency and extended high-frequency audiogram thresholds. J Int Adv Otol 2015;11:104-9.
18. Barcelona N. Neuroelectric's Wiki; 2015. Available from: http://www.wiki.neuroelectrics.com/index.php/Neuroelectric%27s_Wiki. [Last accessed on 2015 Aug 12].
19. Lee CY, Jaw FS, Pan SL, Lin MY, Young YH. Auditory cortical evoked potentials in tinnitus patients with normal audiological presentation. J Formos Med Assoc 2007;106:979-85.
20. Ibrahim IA, Santhosh J, Moghavvemi M. A new approach for an effective eye movement artifact elimination from EEG signal. IEEE International Conference on Signal and Image Processing Applications (ICSIPA) 2015. Malaysia, KL: IEEE; 2015.
21. Blanco-Velasco M, Weng B, Barner KE. ECG signal denoising and baseline wander correction based on the empirical mode decomposition. Comput Biol Med 2008;38:1-13.
22. Näätänen R. Attention and Brain Function. Hillsdale, NJ: Psycholgy Press; 1992.
23. McPherson DL, Ballachanda B, Kaf W. Middle and Long Latency Auditory Evoked Potentials, Audiology Diagnosis. New York: Ed Thieme; 2000. p. 471-501.
24. Davies PL, Chang WP, Gavin WJ. Middle and late latency ERP components discriminate between adults, typical children, and children with sensory processing disorders. Front Integr Neurosci 2010;4:16.
25. Bardy F, Van Dun B, Dillon H, McMahon CM. Deconvolution of overlapping cortical auditory evoked potentials recorded using short stimulus onset-asynchrony ranges. Clin Neurophysiol 2014;125:814- 26.
26. Hall JW. New Handbook of Auditory Evoked Responses. San Diego, CA: Plural; 2007.
27. Näätänen R, Paavilainen P, Rinne T, Alho K. The mismatch negativity (MMN) in basic research of central auditory processing: A review. Clin Neurophysiol 2007;118:2544-90.
28. Al-Qazzaz NK, Ali SH, Ahmad SA, Chellappan K, Islam MS, Escudero J. Role of EEG as biomarker in the early detection and classification of dementia. Sci World J 2014;2014:1-16.
29. Lotte F, Congedo M, Lécuyer A, Lamarche F. A review of classification algorithms for EEG-based brain-computer interfaces. J Neural Eng 2007;4:24.
30. Chen J, Huang H, Tian S, Qu Y. Feature selection for text classification with Naïve Bayes. Expert Syst Appl 2009;36:5432-5.
31. Polat K, Güneş S. Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform. Appl Math Comput 2007;187:1017-26.
32. Acharya UR, Fujita H, Sudarshan VK, Bhat S, Koh JE. Application of entropies for automated diagnosis of epilepsy using EEG signals: A review. Knowl Based Syst 2015;88:85-96.
33. Amin HU, Malik AS, Ahmad RF, Badruddin N, Kamel N, Hussain M, *et al*. Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques. Aust Phys Eng Sci Med 2015;38:139-49.
34. Giraudet L, St-Louis ME, Scannella S, Causse M. P300 event-related potential as an indicator of inattentional deafness? PLoS One 2015;10:e0118556.
35. Key APF, Dove GO, Maguire MJ. Linking brainwaves to the brain: An ERP primer. Dev Neuropsychol 2005;27:183-215.
36. Mangun GR. Neural mechanisms of visual selective attention. Psychophysiology 1995;32:4-18.
37. Hari R, Kaila K, Katila T, Tuomisto T, Varpula T. Interstimulus interval dependence of the auditory vertex response and its magnetic counterpart:

Implications for their neural generation. Electroencephalogr Clin Neurophysiol 1982;54:561-9.

38. Crowley KE, Colrain IM. A review of the evidence for P2 being an independent component process: Age, sleep and modality. Clin Neurophysiol 2004;115:732-44.
39. Patterson JV, Jin Y, Gierczak M, Hetrick WP, Potkin S, Bunney WE, *et al*. Effects of temporal variability on p50 and the gating ratio in schizophrenia: A frequency domain adaptive filter single-trial analysis. Arch Gen Psychiatry 2000;57:57-64.
40. Duncan CC, Barry RJ, Connolly JF, Fischer C, Michie PT, Näätänen R, *et al*. Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. Clin Neurophysiol 2009;120:1883-908.
41. Pantev C, Bertrand O, Eulitz C, Verkindt C, Hampson S, Schuierer G, *et al*. Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. Electroencephalogr Clin Neurophysiol 1995;94:26-40.
42. Lauter JL, Herscovitch P, Formby C, Raichle ME. Tonotopic organization in human auditory cortex revealed by positron emission tomography. Hear Res 1985;20:199-205.
43. Strainer JC, Ulmer JL, Yetkin FZ, Haughton VM, Daniels DL, Millen SJ. Functional MR of the primary auditory cortex: An analysis of pure tone activation and tone discrimination. AJNR Am J Neuroradiol 1997;18:601-10.
44. Woods DL, Alain C, Covarrubias D, Zaidel O. Frequency-related differences in the speed of human auditory processing. Hear Res 1993;66:46-52.
45. Ponton CW, Don M, Eggermont JJ, Waring MD, Kwong B, Masuda A. Auditory system plasticity in children after long periods of complete deafness. Neuro Report 1996;8:61-5.
46. Picton TW. The P300 wave of the human event-related potential. J Clin Neurophysiology 1992;9:456-79.
47. Polich J. Updating P300: An integrative theory of P3a and P3b. Clin Neurophysiol 2007;118:2128-48.
48. Yetkin FZ, Roland PS, Christensen WF, Purdy PD. Silent functional magnetic resonance imaging (FMRI) of tonotopicity and stimulus intensity coding in human primary auditory cortex. Laryngoscope 2004;114:512-8.
49. Jacobson GP, Lombardi DM, Gibbens ND, Ahmad B, Newman CW. The effects of stimulus frequency and recording site on the amplitude and latency of multichannel cortical auditory evoked potential (CAEP) component N1. Ear Hear 1992;13:300-6.
50. Luck S. An Introduction to Event-Related Potentials and their Neural Origins. Ch. 1. Cambridge: MIT Press; 2005.
51. Abdulkader SN, Atia A, Mostafa MS. Brain computer interfacing: Applications and challenges. Egypt Informatics J 2015;16:213-30.