Event-Related Potential-Based Brain–Computer Interface Using the Thai Vowels’ and Numerals’ Auditory Stimulus Pattern

Manorot Borirakarawin and Yunyong Punsawad

1 School of Informatics, Walailak University, Nakhon Si Thammarat 80160, Thailand
2 Informatics Innovative Center of Excellence, Walailak University, Nakhon Si Thammarat 80160, Thailand
* Correspondence: yunyong.pu@wu.ac.th; Tel.: +66-86-909-1568

Abstract: Herein, we developed an auditory stimulus pattern for an event-related potential (ERP)-based brain–computer interface (BCI) system to improve control and communication in quadriplegia with visual impairment. Auditory stimulus paradigms for multicommand electroencephalogram (EEG)-based BCIs and audio stimulus patterns were examined. With the proposed auditory stimulation, using the selected Thai vowel, similar to the English vowel, and Thai numeral sounds, as simple target recognition, we explored the ERPs’ response and classification efficiency from the suggested EEG channels. We also investigated the use of single and multi-loudspeakers for auditory stimuli. Four commands were created using the proposed paradigm. The experimental paradigm was designed to observe ERP responses and verify the proposed auditory stimulus pattern. The conventional classification method produced four commands using the proposed auditory stimulus pattern. The results established that the proposed auditory stimulation with 20 to 30 trials of stream stimuli could produce a prominent ERP response from Pz channels. The vowel stimuli could achieve higher accuracy than the proposed numeral stimuli for two auditory stimuli intervals (100 and 250 ms). Additionally, multi-loudspeaker patterns through vowel and numeral sound stimulation provided an accuracy greater than 85% of the average accuracy. Thus, the proposed auditory stimulation patterns can be implemented as a real-time BCI system to aid in the daily activities of quadriplegic patients with visual and tactile impairments. In future, practical use of the auditory ERP-based BCI system will be demonstrated and verified in an actual scenario.

Keywords: brain–computer interface; electroencephalography; event-related potential (ERP); auditory stimulation

1. Introduction

Practical brain–computer interface (BCI) systems for medical applications are complicated. Design and implementation techniques are required to build a communication system between the brain and external devices or systems without muscular activities [1–4]. A BCI-based assistive technology for controlling external devices or machines for mobility enhancement and rehabilitation has been proposed previously [5]. Varying disability levels can pose problems for developing BCI systems. Previous research has focused on user efficiency for evaluating the cognition function in their proposed BCI systems. Electroencephalography (EEG) is a popular technique for acquiring brain data for non-invasive BCI systems [4–6]. Currently, EEG devices are rapidly being developed for clinical [7–9] and non-clinical applications, such as entertainment [10,11], transport [12], and education [13], such as in an examination and comparison of the EEG-based attention test with continuous performance test (CPT) and test of variables of attention (TOVA) to mention the relationship between a cognitive psychological attention test and the attention levels [14]. Moreover, applications of brainwave-based control to examine the effect of different web-based media on human brain waves by BCI systems might increase knowledge transfer efficiency in digital educational materials [15].
Visual evoked potential (VEP) is a widely exploited EEG signal generated by the visual cortex in response to visual stimulation. VEP-based BCI is a high-performance BCI method [16–18]. The method has been widely implemented to enhance communication and control; its applications include BCI spellers based on transient VEP (or P300) [19] and steady-state VEP (SSVEP)-based BCI for machine control [20, 21]. Moreover, a low-cost human–machine interface system can also support some patients via techniques such as eye-tracking [22]. Such systems require normal visual functions. Moreover, somatosensory evoked potential-based BCI based on tactile stimulation has been proposed for patients with impaired vision or eye movements and hearing [23, 24].

In cases of amyotrophic lateral sclerosis (ALS) or locked-in syndrome (LIS) patients with visual and tactile impairment, auditory stimulation is an alternative method of brain activation, and modulation generating evoked potentials (EPs) can be measured using EEG. Event-related potentials (ERPs) are brain responses that involve sensory, cognitive, or motor events. ERP-based BCI usually employs external sound, light, and tactile stimulation. Similar to the VEP-based BCI, an auditory evoked potential (AEP)-based BCI [25] can be categorized according to the stimulus method and EEG changes into two types: (1) auditory ERPs, which are transient responses [26], and (2) steady-state AEP-BCI [27]. AEP-based BCIs require selective auditory attention based on the human hearing process. Neurology can reveal the pathway between both the auditory nerve and cortex in the perception of encoding and auditory processing inputs [28]. Furthermore, auditory recognition and memory processes have been reported previously [29]. Auditory stimuli are popular techniques for evaluation and treatment processes in psychology and neurotherapy [7, 30].

In previous research on AEP-based BCIs, as shown in Table 1, researchers mainly considered protocols and paradigms of auditory stimulation methods and the position of electrodes. Most studies have demonstrated that EEG signals obtained from the frontal, temporal, and parietal areas produce an evoked response (P300) that synchronizes with auditory tasks. For auditory stimulation, Pokorny et al. [31] provided different sound tones (low tone: 396 Hz; high tone: 1900 Hz) for the left and right sides of the stimulus using earphones to create more than a single command in the auditory BCI system. The system achieved up to 90.6% accuracy. Moreover, Cai et al. [32] implemented eight sound stimuli directions with two different short white and pink noise stimuli bursts through eight loudspeakers in an octagonal setting for a multi-command BCI. An offline experiment reported that the proposed auditory stimulus paradigm improved the results by providing higher classification results and an improved information transfer rate than conventional paradigms.

Table 1. Research studies on the auditory stimulus for event related potential (ERP)-based brain–computer interface (BCI).

| Author           | Proposed Method                          | Auditory Stimuli                                                                 | Speakers | Electrodes | Result(s)                                      |
|------------------|------------------------------------------|-------------------------------------------------------------------------------|----------|------------|------------------------------------------------|
| Pokorny et al., 2013 [31] | The auditory P300-based single-switch brain–computer interface | - Low tones (396 Hz) and the deviants (297 Hz) in 300 ms  
                     |                                    | Earphones | F3, Fz, F4, C3, Cz, C4, P3, Pz, P4 | The mean classification accuracies lie between 69% and 83.5% for a stepwise linear discriminant analysis classifier for 12 Minimally conscious state patients |
| Cai et al., 2015 [32]  | Spatial Auditory Brain–Computer Interface | The white- and pink-noise stimuli of 30 ms lengths with 5 ms | Eight loud speakers | P3, P4, P5, P6, Cz, CPz, Pz, POz | The proposed model can boost the results by up to +10.43 bit/min (+44% classification accuracy). |
Table 1. Cont.

| Author                        | Proposed Method                                      | Auditory Stimuli                                                                 | Speakers          | Electrodes                                                       | Result(s)                                |
|-------------------------------|------------------------------------------------------|----------------------------------------------------------------------------------|-------------------|-----------------------------------------------------------------|------------------------------------------|
| Matsumoto et al., 2013 [33]   | P300 Responses to Various Vowel Stimuli for an Auditory BCI | Vowels stimuli “a”, “e”, “i”, “o”, and “u” duration was set to 100 ms               | Inner ear headphones | Classification accuracies between 60% and 100%                   |
| Halder et al., 2016 [34]      | Auditory P300 BCIs via Hiragana syllabary            | 46 Hiragana characters with a two-step procedure: First the consonant (ten choices) and then the vowel (five choices) | Earphones         | The mean classification accuracy is above 70%                   |
| Onishi et al., 2017 [35]      | Affective Stimuli for an Auditory P300 BCI           | Positive and negative affective sounds (PA: a meowing cat sound, NA: a screaming cat sound) | Earphones         | ALS patients achieved 90% online classification accuracy         |
| Simon et al., 2015 [36]       | An auditory multi-class BCI for a speller            | Natural stimuli from five animal voices i.e., duck, bird, frog, seagull, and dove. Each tone had a length of 150 ms and ISI of 250 ms. | Headphones        | Healthy participants achieved an average accuracy of 90%. The average accuracy of ALS patients was 47%.             |
| Kaongoen & Jo, 2017 [37]      | ASSR/P300-hybrid auditory BCI system                 | Two sounds were set to have 1 kHz and 2 kHz pitches with 37 Hz and 43 Hz AM frequencies, respectively | Earphones         | The maximum average classification accuracy for each setting are P300: 98.38%, and ASSR: 85.42%             |

Voice, as auditory stimuli, may be easy to recognize in a multi-command BCI. Matsumoto et al. [33] proposed the use of English vowel sounds such as “a”, “e”, “i”, “o”, and “u” through inner ear headphones for auditory stimulation. The results showed that the accuracy ranged from 60 to 100%. In addition, Halder et al. [34] adopted the Japanese Hiragana syllabary using earphones for a multi-command BCI. Their proposed BCI system demonstrated an average accuracy of over 70%. In addition, Onishi et al. [35] employed a positive affective sound (a meowing cat sound) and negative affective sound (a screaming cat sound) with the answers “yes” and “no”. The results indicated that the affective auditory BCI system yielded a high classification accuracy. Furthermore, Simon et al. [36] demonstrated an auditory BCI for spellers using natural stimuli from five animal voices, i.e., duck, bird, frog, seagull, and dove. The results indicated that the training session increased the efficiency of healthy subjects and patients with ALS.

A steady-state auditory response (ASSR) is usually utilized to enhance the auditory ERP responses in a BCI, and the stimulation can yield high-intensity EEG features in the frequency domain. Kaongoen and Jo [37] demonstrated the combination of ASSR and P300 paradigms for a hybrid-auditory BCI. Different sound modulations for the left and right sides, such as a 1 kHz sound with an amplitude modulation (AM) frequency of 37 Hz and a 2 kHz sound with an AM frequency of 43 Hz, can increase the difference between the sound sources in the auditory P300-based BCI system. The system produces superior classification results.

According to previous research, as shown in Table 1, auditory BCIs can be used as an alternative for developing assistive devices. In this study, we attempted to improve the performance of an AEP-based BCI system for control and communication among patients with severe paralysis. However, the sound-stimulus pattern for the proposed auditory stimulus paradigm is challenging for a practical auditory BCI [38]. Hence, we suggested and observed the auditory ERP stimulation paradigm by employing Thai vowels with a pronunciation similar to that of English vowels for a multicommand auditory...
was used for motion artifact removal via the EEGLAB toolbox [40]. Subsequently, we employed the ERPLAB toolbox to segment EEG signals and detect ERP responses [41].

Figure 1. Cont.
2.2. Proposed Auditory Stimulation Patterns and Paradigm

For a practical auditory BCI, loudspeakers are used for auditory stimulation, such as inherent hearing, to observe the proposed sound stimulus [32]. Similarly, we also propose an auditory stimulus paradigm for BCI systems. According to previous studies [33–36], this research focuses on applying the auditory BCI to Thai people and then to people belonging to other nationalities. Auditory stimulus intervals (ASI) as tone duration and trials of auditory stimulations using the four vowels “å”, “e”, “i”, and “o”, because their sounds are similar to those of some Thai vowels. Moreover, Thai numeral sounds that were easy to recognize were used on all participants. We performed a pilot study to validate Thai numerals with various durations of ASI through reaction time (RT) testing. Most participants reported unclear hearing for durations < 100 ms, affecting sound recognition. Therefore, we compressed each sound file into 100 ms (minimum duration that can be recognized) and 250 ms (normal duration) durations to observe the ERP responses and features for setting the sound intervals. The audio stimulator included an Arduino Uno microcontroller board that cooperated with a mini MP3 module to generate a random sound pattern stimulation (sound level: 70 dB), and it sent trigger signals to the OpenBCI board, as shown in Figure 1a. Two patterns of the loudspeaker position settings were developed. The loudspeakers were located in front of the participant. The distance between the loudspeaker and the participant was 60 cm for each pattern, following the human spatial auditory performance and the loudspeaker specification (Figure 2). The loudspeakers were placed in the same plane as that of the participant’s ears. The first pattern used only a single loudspeaker, as shown in Figure 2a, for steaming four auditory stimuli, as shown in Table 2. The second pattern was employed apart from sound sources using four loudspeakers set in a semicircle for spatial auditory stimuli [32,42,43], as shown in Figure 2a.

| Sound Stimulus | Vowels | Numerals |
|----------------|--------|----------|
|                | Thai   | English  | Thai     | English |
| S1             | å      | a        | หนึ่ง    | neung   |
| S2             | ฮ       | e        | สอง      | song    |
| S3             | ไอ     | i        | สาม      | sām     |
| S4             | โอ      | o        | สี       | sī      |

Table 2. Vowel and numeral sounds for auditory-evoked potential stimulation.
Figure 2. Two patterns of the loudspeaker position settings for auditory stimulation. (a) Single-loudspeaker position setting. (b) Four loudspeaker position settings in front of the participant.

The experimental setup is illustrated in Figure 3. The paradigm is set to start in a resting state for baseline collection at 2000 ms. A single session consisting of 30 trials of an auditory stream stimulus with two ASI, that is, 100 ms and 250 ms, was used for verification. Subsequently, the sound was muted for 300 ms. Each trial consisted of four random stimulus sounds. The participant performed only the target sounds following the sequence in Table 3 for vowel sounds and Table 4 for numeral sounds as 12 targets (sessions) for three rounds per sound (vowel and numeral) with two proposed auditory stimulation patterns (single and multi-loudspeaker). Before proceeding to the new round, participants were allowed to rest for 5 min.

Table 3. Target of the auditory ERP components using stimuli as vowel sounds with a duration of 100 and 250 ms.

| Session | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 |
|---------|----|----|----|----|----|----|----|----|----|----|----|----|
| Target  | a  | i  | e  | o  | i  | i  | o  | e  | e  | i  | o  | a  |

Table 4. Target of the auditory ERP components using stimuli as numeral sounds with a duration of 100 and 250 ms.

| Session | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 |
|---------|----|----|----|----|----|----|----|----|----|----|----|----|
| Target  | neung | sâm | song | sì | sâm | neung | sì | song | song | sâm | sì | neung |
2.3. Auditory ERP Detections

According to previous studies presented in Table 1, the ERP components, such as N200 (N2) and P300 (P3), were used for EEG feature extraction and as parameters. Note that N200 and P300 exhibit negative voltage peaks and positive voltage values, respectively, with latencies of 150–250 and 250–500 ms, respectively, after the onset of the stimulus. For each EEG channel, the stimulus data were segmented at intervals of 200 ms before the onset and at 600 ms after the onset to collect the baseline and ERP components, respectively. The average was calculated for each trial to predict the target for each stimulation pattern and participant.

Stepwise linear discriminant analysis (SWLDA) employs a combination of forward and backward stepwise analyses of variable features. SWLDA can provide robust feature extraction and a good prediction for P300-based BCIs [44]. The most correlated features are selected to predict the targeted command based on statistical significance [45]. The conventional technique using N200 and P300 features and SWLDA can yield a high efficiency for both offline [31–33] and online testing [35,36]. The training ERP dataset was used to generate the weight of the discriminant model. The SWLDA process starts without feature parameters in the discriminant model. The forward analysis enters an input feature parameter into the model. The model is evaluated by an F-test. If the $p$-value < 0.01, the variable is added. The backward step removes the least significant feature parameter ($p < 0.015$). Thus, linear regression is used to include or exclude feature parameters for prediction. The test ERP dataset was computed with the feature weights to distinguish the target and determine a maximum score for classification [36,45].

2.4. Experiments

Before EEG recording, we had to prepare the participants by placing the cap electrode and checking the position of the electrode by following the 10–20 system (Figure 1b). The experiment was conducted in a quiet room with typical indoor lighting. The participants sat in front of the loudspeaker, as shown in Figure 4. Then, participants received instructions on the auditory stimulus experiment paradigm shown in Figure 3. Each participant attended the target sound, as shown in Tables 3 and 4. The experimental setup is illustrated in Figure 4. The classification accuracies were determined as the number of correct output targets to validate the proposed auditory ERP stimulus pattern.
3. Results

3.1. Observation Corresponding to the Use of Sound of Numerals for Auditory ERP Stimulation

Before using the sound of numerals, we had to observe ERP features by creating the grand-averaged EPR for visualization using the EEGLAB toolbox, as shown in Figure 5. The black lines show the target responses, and the gray lines show the non-target responses. We can demonstrate filtered EEG data obtained from the subject showing a clear P300 in the black plot that occurred approximately 400 ms after the onset of the stimuli. Peaks can be observed at approximately 200 ms and 500 ms for one speaker and multiple speakers, respectively. The results revealed that the ERP components were similar to those of the previous auditory ERP-based BCI.

**Figure 5.** Example of grand–averaged ERPs obtained from all subjects over the auditory stimulus interval 250 ms. (a) Grand–averaged ERPs from the auditory stimulus based on the sounds of numerals produced via a single loudspeaker. (b) Grand–averaged ERPs from the auditory stimulus based on the sound of numerals produced via multiple loudspeakers.

3.2. Observation Corresponding to the Use of EEG Electrodes for ERP Detection

Before using the sound of numerals, we had to observe ERP features by creating the grand-averaged ERP for visualization using the EEGLAB toolbox, as shown in Figure 5. The black lines show the target responses, and the gray lines show the non-target responses. We can demonstrate filtered EEG data obtained from the subject showing a clear P300 in the black plot that occurred approximately 400 ms after the onset of the stimuli. Peaks can be observed at approximately 200 ms and 500 ms for one speaker and multiple speakers, respectively. The results revealed that the ERP components were similar to those of the previous auditory ERP-based BCI.

**Figure 4.** Experimental setup of the auditory stimuli evoked EEG potentials. (a) Auditory stimulus with a single loudspeaker. (b) Auditory stimulus with multi loudspeakers.
3.2. Observation Corresponding to the Use of EEG Electrodes for ERP Detection

According to Figure 6, we selected the Cz, Fz, and Pz electrodes to obtain additional responses from parietal cortices known to generate ERPs related to spatial and P300 responses. This figure presents the verification of the electrodes for ERP detection. The maximum average accuracy of the electrode was 83.62% for Pz. We observed that the EEG from channel Pz exhibited a positive peak for the target sound. We can observe a grand−averaged ERP in Figure 5. This channel was used to create parameters for classification using a minimum N200 response combined with a maximum P300 response.

![Figure 6. Average classification accuracy corresponding to different positions of electrodes of the proposed auditory ERP stimulation using SWLDA. Confidence interval (Alpha: 0.05).](image)

3.3. Verification of the Proposed Auditory ERP Stimulus Pattern

3.3.1. Auditory Stimulus Trials and Intervals

The trial number and stimulus intervals are essential parameters for the efficiency of auditory BCIs, such as the information transfer rate (ITR). Therefore, we also validated the number of trials for each ASI for the real-time auditory ERP BCI. The EEG recorded from 30 trials was divided into five ranges for ERP detection. Figure 7 shows that the average classification accuracy ranged from 65.5% to 89.6%, and the average classification accuracy in the number of trials was 20, 25, and 30 at an ASI of 100 ms, and for an ASI of 250 ms, the average classification accuracy was more than 80%. The ASI of 250 ms achieved more than 85% of the average classification accuracy. We observed a maximum classification accuracy of 89.6% for the ASI at 250 ms in 30 trials. The ASI at 100 ms with 20 trials reached an average classification accuracy of 83.8% and 32 s for generating an output faster than the ASI at 250 ms, which used 44 s. Moreover, the ASI at 250 ms, with 15 trials yielded approximately 80% of the average classification accuracy using 33 s for output creation. We employed 20 and 25 trials to observe an ASI with different sounds, such as vowels and numerals.

According to Table 5, the average classification accuracy of the auditory stimulus with an ASI at 100 ms ranged from 69.4% to 92.5%. The average classification accuracy for the ASI at 250 ms ranged from 77.5% to 90.5%. The maximum classification accuracy of the sound stimulus of numerals was 88.5%, whereas the maximum classification accuracy of the sound stimulus of vowels was 92.5% for the ASI at 100 ms. In addition, the average accuracy of the numeral stimuli with the ASI at 250 ms was 83.8%, which was greater than the average accuracy of the ASI at 100 ms. The results suggested that the sound stimulus of vowels with the ASI at 100 ms can produce a high average classification accuracy for all participants, as opposed to the sound stimulus of numerals.
the number of trials for each ASI for the real-time auditory ERP BCI. The EEG recorded from 30 trials was divided into five ranges for ERP detection. Figure 7 shows that the average classification accuracy ranged from 65.5% to 89.6%, and the average classification accuracy in the number of trials was 20, 25, and 30 at an ASI of 100 ms, and for an ASI of 250 ms, the average classification accuracy was more than 80%. The ASI of 250 ms achieved more than 85% of the average classification accuracy. We observed a maximum classification accuracy of 89.6% for the ASI at 250 ms in 30 trials. The ASI at 100 ms with 20 trials reached an average classification accuracy of 83.8% and 32 s for generating an output faster than the ASI at 250 ms, which used 44 s. Moreover, the ASI at 250 ms, with 15 trials yielded approximately 80% of the average classification accuracy using 33 s for output creation. We employed 20 and 25 trials to observe an ASI with different sounds, such as vowels and numerals.

Figure 7. Average classification accuracy between 100 ms and 250 ms in the auditory stimulus intervals. Confidence interval (Alpha: 0.01).

Table 5. Results of the average classification accuracy of different stimuli intervals (tone duration) of vowels and numerals for each participant.

| Auditory Stimuli | Average Classification Accuracy (%) | Vowels | Numerals |
|------------------|-------------------------------------|--------|---------|
|                   | ASI: 100 ms | ASI: 250 ms | ASI: 100 ms | ASI: 250 ms |
| 1                | 88.8        | 87.5       | 80.0      | 88.5       |
| 2                | 90.3        | 86.3       | 84.5      | 88.0       |
| 3                | 90.0        | 84.8       | 79.5      | 84.5       |
| 4                | 90.8        | 86.0       | 72.5      | 82.5       |
| 5                | 85.0        | 83.8       | 73.5      | 81.7       |
| 6                | 92.5        | 86.5       | 77.5      | 85.0       |
| 7                | 90.0        | 90.5       | 77.5      | 87.5       |
| 8                | 87.5        | 87.5       | 72.5      | 77.5       |
| 9                | 91.5        | 82.5       | 80.0      | 82.8       |
| 10               | 85.4        | 88.9       | 69.4      | 89.0       |
| 11               | 83.5        | 83.5       | 82.5      | 78.5       |
| 12               | 86.5        | 82.5       | 80.3      | 80.0       |

Mean ± SD. 88.5 ± 2.86 85.8 ± 2.53 77.5 ± 4.57 83.8 ± 3.95

We observed a significant difference between 100 ms and 250 ms in the stimulus intervals of vowel and numeral (acquainted sound) stimuli. We used the paired t-test to analyze statistically significant differences between the stimulus intervals and sound stimuli groups. The results are shown in Figure 8. We found that the paired t-test (n = 12) indicated that there was a significant difference in the average classification accuracy between the stimulus intervals of 100 and 250 ms for the vowels (p = 0.024; p < 0.05). Figure 8 also indicates a significant difference between the accuracies obtained for the stimulus intervals of 100 and 250 ms for numeral stimuli (p = 0.004; p < 0.005). Moreover, a statistically significant difference could be identified between the average classification accuracy obtained from the vowel and numeral stimuli, irrespective of the stimulus intervals (p = 0.000; p < 0.005).
We observed a significant difference between 100 ms and 250 ms in the stimulus intervals of vowel and numeral (acquainted sound) stimuli. We used the paired \( t \)-test to analyze statistically significant differences between the stimulus intervals and sound stimuli groups. The results are shown in Figure 8. We found that the paired \( t \)-test (\( n = 12 \)) indicated that there was a significant difference in the average classification accuracy between the stimulus intervals of 100 and 250 ms for the vowels (\( p = 0.024; p < 0.05 \)). Figure 8 also indicates a significant difference between the accuracies obtained for the stimulus intervals of 100 and 250 ms for numeral stimuli (\( p = 0.004; p < 0.005 \)). Moreover, a statistically significant difference could be identified between the average classification accuracy obtained from the vowel and numeral stimuli, irrespective of the stimulus intervals (\( p = 0.000; p < 0.005 \)).

**Figure 8.** Average classification accuracy between vowels and numeral sounds in auditory stimulus intervals of 100 ms and 250 ms.

### 3.3.2. Auditory Stimulation Patterns

According to the results in Table 6, the average accuracy of the vowel auditory stimuli ranged from 71.0% to 97.5% for all subjects and that for the numerical stimuli ranged from 67.5% to 95.0%. We found that both sounds provided similar efficiencies. The second test is an evaluation of the number of speakers between single and multiple (four) loudspeakers for comparison between vowel and numeral sound stimuli. The average accuracy of using the vowel sound was 86.0% and the numerical sound was 75.1% for a single loudspeaker. The average accuracy of the vowel sound was 87.2% and that of numerals was 86.2% for multiple loudspeakers, which was greater than the average accuracy of the single speaker. These results verified and supported the proposed auditory stimulus pattern. The vowel and numerical patterns obtained with the multi-loudspeaker method can generate the average classification accuracy of all subjects. For some participants, both stimulus patterns yielded similar classification accuracies for each auditory stimulation-detection method.

**Table 6.** Results of the average classification accuracy under different loudspeaker patterns of vowel and numeral stimuli for each participant.

| Loudspeaker Patterns | Average Classification Accuracy (%) |
|----------------------|-------------------------------------|
|                      | Single | Multi |
| Participants         | Vowels | Numerals | Vowels | Numerals |
| 1                    | 87.5   | 75.0     | 88.8   | 93.5     |
| 2                    | 82.8   | 84.5     | 93.8   | 88.0     |
| 3                    | 93.5   | 80.0     | 81.3   | 84.0     |
| 4                    | 82.3   | 72.5     | 94.5   | 82.5     |
| 5                    | 85.0   | 71.7     | 83.8   | 83.5     |
| 6                    | 91.5   | 80.0     | 87.5   | 82.5     |
| 7                    | 97.5   | 70.0     | 90.0   | 95.0     |
| 8                    | 87.5   | 67.5     | 87.5   | 82.5     |
| 9                    | 86.5   | 77.8     | 87.5   | 85.0     |
| 10                   | 85.5   | 71.5     | 88.8   | 86.9     |
| 11                   | 71.0   | 77.5     | 86.0   | 83.5     |
| 12                   | 81.5   | 72.8     | 77.5   | 87.5     |
| Mean ± SD.           | 86.0 ± 6.71      | 75.1 ± 4.95 | 87.2 ± 4.79 | 86.2 ± 4.24 |

According to Figure 9, two main issues were observed. The first issue concerns the efficiency of each loudspeaker pattern. The paired \( t \)-test (\( n = 24 \)) showed a statistically
significant difference between the single- and multiple-loudspeaker patterns ($p = 0.003$; $p < 0.005$). In addition, using four loudspeakers could yield a higher efficiency than using a single loudspeaker. The second issue is the efficiency of each auditory stimulus. The paired $t$-test ($n = 12$) indicated a significant difference between the vowel and numeral auditory stimuli using a single loudspeaker ($p = 0.001$; $p < 0.005$). Furthermore, the paired $t$-test indicated no significant difference between the vowel and numeral patterns using the multi-loudspeaker method ($p = 0.559$; $p > 0.05$). Even though the proposed auditory stimulus via a numeral pattern can achieve a higher average classification accuracy than the vowel pattern for all subjects, fatigue with concentration on sound must be observed between the vowel and numeral patterns for practical purposes.

According to Figure 9, two main issues were observed. The first issue concerns the average classification accuracy between vowel and numeral sounds of the auditory ERP stimulus of single and multi-loudspeaker patterns.

4. Discussion

Observing the proposed auditory ERP stimuli via a numeral stimulus (in Thai) for the auditory BCIs, ERP responses of the numeral stimulus pattern showed obvious N200 and P300 responses, refs. [31,36] which were mainly produced by the mid-frontal and mid-parietal areas (Fz and Pz). Moreover, we verified auditory stimulus intervals and trials for ERP detection using the SWLDA method. The experiment reported that 100 ms of ASI with more than 20 stimulus trials and 250 ms of ASI with more than 15 stimulus trials achieved acceptable accuracies. In addition, the ASI at 100 ms of the vowel stimulus provided a better ERP response than the numeral stimulus. The 250 ms numeral (word) stimulation yielded acceptable accuracies. Furthermore, we investigated the efficiency of single- and multi-loudspeakers position patterns of auditory stimuli. The experiment reported that spatial auditory stimuli using multi-loudspeaker yielded better classification accuracy than using only a single loudspeaker for vowel and numeral sound streaming.

Nevertheless, some limitations of the proposed auditory ERP stimulus patterns for BCI are as follows. (1) Users require a training session for recognition, and they need to be guided on how to pay attention to target sounds. (2) The proposed method requires a long time to create each command, which results in a low ITR. This disadvantage should be addressed for real-time BCI systems. The addition of a command to enable/disable an auditory stimulator is recommended for practical use. In addition, the consideration of further validation can be listed as follows: (1) the sound level and distance between loudspeakers require a recommendation protocol; (2) longer intervals can be applied after successive stimuli for studying the cognitive components; (3) a usability test can be

![Figure 9. Average classification accuracy between vowel and numeral sounds of the auditory ERP stimulus of single and multi-loudspeaker patterns.](image-url)
performed to compare related BCI techniques such as tactile BCI [24]; (4) a greater number of targets can be employed to verify the system efficiency and user workload.

5. Conclusions

This paper proposes an auditory ERP pattern for a practical BCI system for quadriplegia with visual impairments. We verified the auditory stimulus pattern to observe the ERP responses to Thai vowel and numeral sounds with different stimulus intervals and a number of trials for command classification. We found that a vowel auditory stimulus interval of 100 ms for 20 trials (32 s per command) and a numeral auditory stimulus interval of 250 ms for 15 trials (33 s per command) can be used for an auditory BCI. In addition, the loudspeaker positions and patterns were investigated. As reported previously [40–42], the proposed multi-loudspeaker pattern via a spatial auditory stimulation can yield a higher classification accuracy than a single loudspeaker for both vowel and numeral stimuli. Finally, the proposed auditory stimulation can be used in real-time BCI systems for quadriplegic patients with visual and tactile impairments. We suggested that the multi-loudspeaker patterns for auditory stimuli via vowel and numeral sounds can be implemented in auditory ERP-based BCI systems for control and communication to aid in their daily activities. In future studies, we will verify the practical use of the auditory ERP stimulation with an independent loudspeaker position in actual situations with natural noise.

Author Contributions: Conceptualization, M.B. and Y.P.; methodology, M.B. and Y.P.; software, M.B. and Y.P.; validation, Y.P.; formal analysis, Y.P.; investigation, Y.P.; resources, Y.P.; data curation, M.B. and Y.P.; writing—original draft preparation, M.B. and Y.P.; writing—review and editing, M.B. and Y.P.; visualization, Y.P.; supervision, Y.P.; project administration, Y.P.; funding acquisition, Y.P. All authors have read and agreed to the published version of the manuscript.

Funding: This study was partially supported by Walailak University, Thailand.

Institutional Review Board Statement: The study was conducted in accordance with the guidelines of the Declaration of Helsinki and approved by the Office of the Human Research Ethics Committee of Walailak University (protocol code: WU-EC-EN-2-140-64, approval date: 14 June 2021).

Informed Consent Statement: The informed consent of all of the subjects participating in the study has been obtained, and the written informed consent of the participants has been obtained to publish this article.

Data Availability Statement: The data presented in this study are available upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wolpaw, J.R.; Birbaumer, N.; McFarland, D.J.; Pfurtscheller, G.; Vaughan, T.M. Brain-computer interface for communication and control. Clin. Neurophysiol. 2002, 113, 767–791. [CrossRef]
2. Abdulkader, S.N.; Atia, A.; Mostafa, M.S.M. Brain computer interfacing: Applications and challenges. Egypt. Inform. J. 2015, 16, 213–230. [CrossRef]
3. Mridha, M.F.; Das, S.C.; Kabir, M.M.; Lima, A.A.; Islam, M.R.; Watanobe, Y. Brain-computer interface: Advancement and challenges. Sensors 2021, 21, 5746. [CrossRef] [PubMed]
4. Nicolas-Alonso, L.F.; Gomez-Gil, J. Brain computer interfaces: A review. Sensors 2012, 12, 1211–1279. [CrossRef] [PubMed]
5. Jamil, N.; Belkacem, A.N.; Ouhbi, S.; Lakas, A. Noninvasive electroencephalography equipment for assistive, adaptive, and rehabilitative brain-computer interfaces: A systematic literature review. Sensors 2021, 21, 4754. [CrossRef] [PubMed]
6. Morshed, B.I.; Khan, A. A brief review of brain signal monitoring technologies for BCI applications: Challenges and prospects. J. Bioeng. Biomed. Sci. 2014, 4, 1. [CrossRef]
7. Birbaumer, N.; Murguialday, A.R.; Weber, C.; Montoya, P. Neurofeedback and brain–computer interface: Clinical applications. Int. Rev. Neurobiol. 2009, 86, 107–117.
8. Alkawadri, R. Brain–computer interface (BCI) applications in mapping of epileptic brain networks based on intracranial-EEG: An update. Front. Neurosci. 2019, 13, 191. [CrossRef]
9. McFarland, D.J.; Daly, J.; Boulay, C.; Parvaz, M.A. Therapeutic applications of BCI technologies. Brain Comput. Interfaces 2017, 4, 37–52. [CrossRef]
10. Ahn, M.; Lee, M.; Choi, J.; Jun, S.C. A review of brain-computer interface games and an opinion survey from researchers, developers and users. Sensors 2014, 14, 14601–14633. [CrossRef]
11. Marshall, D.; Coyle, D.; Wilson, S.; Callaghan, M. Games, gameplay, and BCI: The state of the art. IEEE Trans. Comput. Intell. AI Games. 2013, 5, 82–99. [CrossRef]
12. He, Z.; Li, Z.; Yang, F.; Wang, L.; Li, J.; Zhou, C.; Pan, J. Advances in multimodal emotion recognition based on brain–computer interfaces. Brain Sci. 2020, 10, 687. [CrossRef]
13. Spüler, M.; Krumpe, T.; Walter, C.; Scharinger, C.; Rosenstiel, W.; Gerjets, P. Brain-computer interfaces for educational applications. In Informational Environments, 1st ed.; Buder, J., Hesse, F.W., Eds.; Springer: Cham, Switzerland, 2017; pp. 177–201.
14. Katona, J. Examination and comparison of the EEG based attention test with CPT and T.O.V.A. In Proceedings of the 2014 IEEE 15th International Symposium on Computational Intelligence and Informatics (CINTI), Budapest, Hungary, 19–21 November 2014.
15. Katona, J.; Ujbanyi, T.; Sziládi, G.; Kovári, A. Examine the effect of different web-based media on human brain waves. In Proceedings of the 2017 8th IEEE International Conference on Cognitive Informatics (CogInfoCom), Debrecen, Hungary, 11–14 September 2017.
16. Wang, Y.; Wang, R.; Gao, X.; Hong, B.; Gao, S. A practical VEP-based brain–computer interface. IEEE Trans. Neural Syst. Rehabil. Eng. 2006, 14, 234–240. [CrossRef]
17. Fazel-Rezai, R.; Allison, B.Z.; Guger, C.; Sellers, E.W.; Kleih, S.C.; Kübler, A. P300 brain computer interface: Current challenges and emerging trends. Front. Neuroeng. 2012, 5, 14. [CrossRef]
18. Amari, S.; Ahmed, R.; Azinfar, L.; Fazel-Rezai, R. A Review of P300, SSVEP, and hybrid P300/SSVEP brain-computer interface systems. In Brain-Computer Interface Systems-Recent Progress and Future Prospects; Reiza, R.-F., Ed.; IntechOpen: London, UK, 2013.
19. Guy, V.; Soriani, M.-H.; Bruno, M.; Papadopoulos, T.; Desnuelle, C.; Clerc, M. Brain computer interface with the P300 speller: Usability for disabled people with amyotrophic lateral sclerosis. Ann. Phys. Rehabil. Med. 2018, 61, 5–11. [CrossRef]
20. Na, R.; Hu, C.; Sun, Y.; Wang, S.; Zhang, S.; Han, M.; Yin, W.; Zhang, J.; Chen, X.; Zheng, D. An embedded lightweight SSVEP-BCI electric wheelchair with hybrid stimulator. Digit. Signal Process. 2021, 116, 103101. [CrossRef]
21. Quiles, E.; Dadone, J.; Chio, N.; Garcia, E. Cross-platform implementation of an SSVEP-based BCI for the control of a 6-DOF robotic arm. Sensors 2022, 22, 5000. [CrossRef]
22. Bissoli, A.; Lavino-Junior, D.; Sime, M.; Encarnação, L.; Bastos-Filho, T. A human–machine interface based on eye tracking for controlling and monitoring a smart home using the Internet of Things. Sensors 2019, 19, 859. [CrossRef]
23. Brouwer, A.-M. A tactile P300 brain-computer interface. Front. Neurosci. 2010, 4, 19. [CrossRef]
24. Jin, J.; Chen, Z.; Xu, R.; Miao, Y.; Jung, I.P. Developing a novel tactile P300 brain-computer interface with a Cheeks-Stim paradigm. IEEE Trans. Biomed. Eng. 2020, 67, 2585–2593. [CrossRef]
25. Klobassa, D.S.; Vaughan, T.M.; Brunner, P.; Schwartz, N.E.; Wolpaw, J.R.; Neuper, C.; Sellers, E.W. Toward a high-throughput auditory P300-based brain–computer interface. Clin. Neuropsychol. 2009, 120, 1252–1261. [CrossRef]
26. Cai, Z.; Makino, S.; Yamada, T.; Rutkowski, T.M. Spatial auditory BCI paradigm utilizing N200 and P300 responses. In Proceedings of the 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference, Hollywood, CA, USA, 3–6 December 2012.
27. Punswad, Y.; Wongswat, Y. Multi-command SSAEP-based BCI system with training sessions for SSVEP during an eye fatigue state. IEEE Trans. Electron. Eng. 2017, 12, 572–578. [CrossRef]
28. Price, C.N.; Moncrieff, D. Defining the role of attention in hierarchical auditory processing. Audiol. Res. 2021, 11, 112–128. [CrossRef]
29. Kars, C.M.; Isbell, E.; Giuliano, R.J.; Neville, H.J. Auditory attention in childhood and adolescence: An event-related potential study of spatial selective attention to one of two simultaneous stories. Dev. Cogn. Neurosci. 2015, 13, 53–67. [CrossRef]
30. Legatt, A.D. Electrophysiologic auditory tests. In Handbook of Clinical Neurology; Aminoff, M.J., Boller, F., Swaab, D.F., Eds.; Elsevier: Amsterdam, The Netherlands, 2015; Volume 129, pp. 289–311.
31. Pokorny, C.; Klobassa, D.S.; Pichler, G.; Erlbeck, H.; Real, R.G.; Kübler, A.; Müller-Putz, G.R. The auditory P300-based single-switch brain–computer interface: Paradigm transition from healthy subjects to minimally conscious patients. Artif. Intell. Med. 2013, 59, 81–90. [CrossRef]
32. Cai, Z.; Makino, S.; Rutkowski, T.M. Brain evoked potential latencies optimization for spatial auditory brain–computer interface. Cogn. Comput. 2015, 7, 34–43. [CrossRef]
33. Matsumoto, Y.; Makino, S.; Mori, K.; Rutkowski, T.M. Classifying P300 responses to vowel stimuli for auditory brain-computer interface. In Proceedings of the 2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, Kaohsiung, Taiwan, 29 October–1 November 2013.
34. Halder, S.; Takano, K.; Ora, H.; Onishi, A.; Utsumi, K.; Kansaku, K. An evaluation of training with an auditory P300 brain–computer interface for the Japanese hiragana syllabary. Front. Neurosci. 2016, 10, 446. [CrossRef] [PubMed]
35. Onishi, A.; Takano, K.; Kawase, T.; Ora, H.; Kansaku, K. Affective stimuli for an auditory P300 brain-computer interface. Front. Neurosci. 2017, 11, 522. [CrossRef] [PubMed]
36. Simon, N.; Käthner, I.; Ruf, C.A.; Pasqualotto, E.; Kübler, A.; Halder, S. An auditory multiclass brain-computer interface with natural stimuli: Usability evaluation with healthy participants and a motor impaired end user. Front. Hum. Neurosci. 2015, 8, 1039. [CrossRef] [PubMed]
37. Kaongoen, N.; Jo, S. A novel hybrid auditory BCI paradigm combining ASSR and P300. *J. Neurosci. Method* 2017, 279, 44–51. [CrossRef]

38. Hsieh, K.L.; Sun, K.T.; Yeh, J.K.; Pan, Y.U. Home care by auditory Brain computer interface for the blind with severe physical disabilities. In Proceedings of the 2017 International Conference on Applied System Innovation (ICASI), Sapporo, Japan, 13–17 May 2017.

39. Borirakarawin, M.; Punsawad, Y. Multicommand auditory ERP-based BCI system. In Proceedings of the 2021 13th Biomedical Engineering International Conference (BMEiCON), Ayutthaya, Thailand, 19–21 November 2021.

40. Delorme, A.; Makeig, S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 2004, 134, 9–21. [CrossRef]

41. Lopez-Calderon, J.; Luck, S.J. ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Front. Hum. Neurosci.* 2014, 8, 213. [CrossRef]

42. Schreuder, M.; Blankertz, B.; Tangermann, M. A new auditory multi-class brain-computer interface paradigm: Spatial hearing as an informative cue. *PLoS ONE* 2010, 5, e9813. [CrossRef]

43. Chang, M.; Makino, S.; Rutkowski, T.M. Classification improvement of P300 response based auditory spatial speller brain-computer interface paradigm. In Proceedings of the 2013 IEEE International Conference of IEEE Region 10 (TENCON 2013), Xi’an, China, 22–25 October 2013.

44. Xiao, X.; Xu, M.; Jin, J.; Wang, Y.; Jung, T.P.; Ming, D. Discriminative canonical pattern matching for single-trial classification of ERP components. *IEEE Trans. Biomed. Eng.* 2019, 67, 2266–2275. [CrossRef]

45. Ogino, M.; Kanoga, S.; Ito, S.I.; Mitsukura, Y. Semi-supervised learning for auditory event-related potential-based brain–computer interface. *IEEE Access* 2021, 9, 47008–47023. [CrossRef]