A Persian Brain-Computer Interface Speller Based on P300 and Language Model Utilization

Hadi Veisi (✉ h.veisi@ut.ac.ir)

University of Tehran <https://orcid.org/0000-0003-2372-7969>

Hessam Amini

Concordia University

Parisa Daeijavad

University of Tehran

Research

Keywords: Brain-Computer Interface, P300, Speller, EEG, Persian Language Model

Posted Date: August 30th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-809975/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

A Persian Brain-Computer Interface Speller Based on P300 and Language Model Utilization

Hadi Veisi ^{1, *}, Hessam Amini ², Parisa Daeijavad ³

¹ Data and Signal Processing Lab (DSP Lab), Faculty of New Sciences and Technologies (FNST), University of Tehran (UT), Tehran, Iran; h.veisi@ut.ac.ir, orcid.org/0000-0003-2372-7969

² Computational Linguistics at Concordia (CLaC) Laboratory, Montreal Natural Language Processing (MTL-NLP), Concordia University; hessam.amini@concordia.ca

³ Data and Signal Processing Lab (DSP Lab), Faculty of New Sciences and Technologies (FNST), University of Tehran (UT), Tehran, Iran; parisa.daeijavad@ut.ac.ir

* Correspondence: h.veisi@ut.ac.ir; Tel.: +98-218-609-3065

Abstract

Background: Brain-computer interface (BCI) is a system in which complete control is exerted directly through brain signals. One of the most important applications of BCI is in spellers, which are designed to restore communication ability for people with severe speech and motor impairments. These people can communicate using a BCI speller by spelling their desired words. The goal of this paper is to present the achievements of a project whose ultimate purpose is to propose a Persian BCI speller, which utilizes a Persian language model as an assisting tool to improve its performance.

Method: In this study, we first present our method for data collection. To collect the dataset, we designed a screen that is used to display the visual stimuli. Furthermore, we carefully selected the words used in the data collection process. Using the designed stimuli screen and the selected words, the dataset from 5 subjects is collected. This dataset is used in different signal processing and machine learning methods in the designed speller. To classify the P300 signal for spelling purposes, first, the feature vectors of data samples were extracted by the concatenation of the electroencephalography (EEG) channels, which were filtered using a moving average filter. Then, using a support vector machine (SVM), the samples were classified and the results were fused with character-level N-grams to detect the target characters.

Results: In the evaluation of the system, we have achieved an average accuracy of 93.2% in character detection by utilizing a character-level 5-gram language model.

Conclusions: The assessment of the language models in the system demonstrated that using the 5-gram language model resulted in increasing the accuracy by 1.9% compared to when no language models were used.

Keywords: Brain-Computer Interface; P300; Speller; EEG; Persian Language Model

1. Introduction

A brain-computer interface (BCI) enables interaction with surroundings directly by translating brain activities into computer commands without using peripheral nerves and muscles [1]. In fact, BCI creates a non-muscular contact way for interpreting a user's purpose on an external device. Therefore, individuals with severe motor disorders can improve their quality of life through such an interface [2].

The major components of BCI systems are signal acquisition, signal preprocessing, feature extraction, and classification [3]. The signal acquisition step in BCI systems is usually done using invasive, semi-invasive, and non-invasive methods [4]. In an invasive method, such as electrocorticography, signals from the human brain are received by sensors that are surgically inserted into the brain or on the surface of the brain. In the semi-invasive technique, the process of reading the signal is performed inside the skull and outside the brain gray matter [5]. Non-invasive methods, such as electroencephalography (EEG), follow an approach that does not require physical insertion into the brain [3]. EEG is a method for measuring brain electrical activity to understand brain cognitive processes or mental disorders. EEG sensors embedded in wearable devices make it possible to record EEG signals in different living conditions of everyday life [6]. In the non-invasive mode, input signals are always associated with some unwanted data, i.e. noise. Electronic interactions, muscle activity signals, and eye-related noise (e.g., blinking) are also recorded with signals. These unwanted components may negatively affect the performance of non-invasive BCI systems [4].

The main purpose of signal preprocessing is to reduce the artifacts that are recorded with EEG signals. Different filters are used to fulfill this goal [7]. Moving average filters and band-pass filters help to remove noise and artifacts [3]. A moving average filter is optimized to reduce random noise and is an appropriate filter in the time domain for encrypted signals [8]. A band-pass filter passes the cutoff frequencies (e.g., 0.5-65 Hz) and prevents the passage of signals at other frequencies [9].

In the feature extraction step, the features of a signal that encodes the user's purpose are extracted. The extracted electrophysiological features have a strong connection with the user's purposes and they can be in the range of time or frequency or both. One of the most common signal features used in current BCI systems is the latency of event-evoked potentials (such as P300) [10]. The P300 signal occurs in the person's reaction to an unexpected stimulus. The onset of the P300 signal in the individual's brain is between 300 and 500 milliseconds after the stimulus [11]. The P300 signal has been used in a variety of applications, such as lie detections and spellers. The P300-based speller is one of the most commonly used applications in BCI systems, which allows users to select and display different characters on a digital screen by detecting the P300 evoked-potential generated from visual stimuli [12]. BCI-based spellers use language context information as a helpful way to classify brain signals. One of the best context information that can be used in spellers is the language context, which can be used in spellers as a language model [13]. A language model can improve the performance, in terms of accuracy and speed, by predicting the next letter using the previous letters [14].

The final stage of BCI, classification, is the most popular method for identifying brain activity patterns. In this method, the BCI system is considered a pattern recognition system. The performance of a pattern recognition system depends on the extracted features and the classification algorithm [15]. Machine learning techniques such as support vector machine (SVM) are employed for classification. SVM is an algorithm that uses supervised learning to discriminate two different classes of data. In this method, the discriminator is determined to have a maximum distance from the nearest training points [3]. Also, the SVM classifier has better performance than linear discriminant analysis (LDA) in the P300 evoked potentials detection [12].

In this paper, a Persian language model was used to improve the performance of the presented Persian BCI speller using acquired EEG signal datasets of 5 subjects in Iran's National Brain Mapping Lab (NBML)¹. To classify the P300 signal, first, the feature vectors of the data samples were extracted by the concatenation of the EEG channels which were filtered using a moving average filter. Then, the samples were classified using SVM, of which the SVM algorithm is more suitable for the detection of P300 evoked potentials [12], and the results were fused with character-level N-grams to detect the target characters.

This paper is organized as follows: Section 2 describes some works related to BCI spellers, Section 3 presents the details of the materials and methods used in this research, Section 4 analyzes the data and results and, finally, in Section 5 the discussion and conclusion are given.

2. Related Works

The first BCI-based speller was introduced by Farwell and Donchin in 1998 [16]. The speller is composed of a 6x6 matrix of symbols, including letters of the English alphabet, numbers 1 to 9, and a space symbol. This matrix speller uses the P300 signal to identify the symbol choice of the user. Following the introduction of this speller, which is widely used for BCI research, a lot of researches has been improved the accuracy and speed of spellers. Some of these studies, based on different methods of signal processing and machine learning, followed the path of increasing the accuracy and speed of the same matrix speller, while others generally presented different methods for constructing a BCI-based speller.

Some researches aimed at introducing new speller development methods for displaying P300 stimuli, such as the use of audio stimuli [17, 18] and tactile stimuli [19], or visual stimuli other than matrix [20, 21] and rapid serial visual presentation paradigm(RSVP) [22, 23].

One of the most important methods of visual representation of stimuli, presented after the matrix method, is the region-based presentation method, presented by Fazel Rezai and Abhari [20]. In this method, 49 characters, including letters of the English alphabet, numbers from 0 to 9, and the number of syllables (such as dots, commas, etc.) were divided into 7 groups of 7 stimuli, which were located in 6 corners and the center of a hexagon. In this method, character selection is a two-step process. Through the detection of the P300, firstly, the group containing the target character can be found after several intensifications. In the second step, the characters of the detected group are distributed into the other seven regions. This method provides the highest accuracy and speed among P300-based methods [24].

In other researches in the area of BCI-based speller, designers have changed the signal to non-P300 signals. Many methods have used Volitional Cortical Potentials (VCP) as an input. The Hex-o-Spell method developed by the Berlin BCI team is one of the most important methods used to design a VCP signal for a speller. In this method, which uses a two-step selection method similar to the region-based presentation method, character selection at each level is made using a motor imaginary [21].

In addition to VCP, in many visual spelling methods, SSVEP has been used. The fastest BCI provided utilizes the SSVEP signal [25, 26, 27]. Among other BCI studies, other studies have used two or more types of input signals to increase accuracy, speed, or efficiency [28, 29, 30, 31, 32, 33, 34].

In BCI-based spellers, some auxiliary methods have also been used to increase spelling speed and accuracy. The most important of these studies is the RSVP keyboard provided by CSL Laboratory at Northwestern University in the United States. The RSVP keyboard is a P300-based speller that uses a rapid serial visual presentation paradigm to display the stimuli. The RSVP keyboard uses the N-gram language model as an auxiliary tool to detect and classify the P300, which significantly increases the accuracy and speed of the system [14].

¹ The current research is based on the ethics code of IR.TUMS.VCR.REC.1395.1214 dated 22/9/1395 from the Ethics Committee of the Vice-Chancellor of Research of Tehran University of Medical Sciences.

For the Persian language, studies on providing a BCI-based speller are limited. The only we have found is Najafi's master's thesis [35] in which provides a Persian P300-based speller. However, our presented method is different in terms of the language model, the screen design of the stimuli, and the classification method.

3. Materials and Methods

In this section, we present the details of our research in developing a Persian EEG-based speller. To this aim, first, the details of the data collection including the screen design of the stimuli, word selection, and data acquisition are given. Afterward, the proposed speller which is utilized language model is presented.

3.1. Screen design of the stimuli

The screen is designed to enable the user to use the covert attention of the system, that can grasp all the stimuli and recognize their color changes by just staring at the center of the screen without moving the eyes. As a result, the placement of the stimuli (which is following the one in Hex-o-Spell [36]), can also help users who cannot move their eyes (e.g., patients with total Locked-in Syndrome [37]). In this speller, the character selection is done in the form of a two-step process, so that in the first step, the characters are placed into six groups in different parts of the screen as shown in Figure 1. The user chooses a group of characters that contains the target character. Then, in the second step, the user selects the target character among the characters in the first step. Figure 1 shows the stimuli on the screen in each of the two steps, the character selection, and how the color of the stimuli is changing to evoke the P300 signal. The screen of the stimuli is designed by the Psychophysics Toolkit in MATLAB [38].

Figure 1. Stimulus configuration: (a) The first step, group selection; (b) The second step, character selection. The flash pattern mechanisms are composed of random flashings and periodic flickers. Moreover, all of the items within a subarea, flash at the same frequency.

Changing the color of each stimulus must be in a semi-random order. A very short time between the two color-changing disables the ability of the user to recognize that stimulus, and if this interval is too long, the stimulus loses its surprising ability, and hence, the amplitude. In the presented screen, the duration of each

color-changing is 150 milliseconds, and between the two successive color changes, all the stimuli remain unchanged for 50 milliseconds.

3.2. Target words selection using the language model

Since the language model is an important part of the presented speller, the impact of the language model on Persian phrases should be assessed. For this purpose, the method presented in the work by Amini et al. [13] was used to select target words in the language model. These language models are trained using a corpus with more than two million Persian words collected from the Iranian Student's News Agency (ISNA) website² [13]. For each of the words, the character-level probability was computed using equation (1).

$$P^C(w) = P_1^C(c_1) \times P_2^C(c_2|c_1) \times P_3^C(c_3|c_1c_2) \times P_4^C(c_4|c_1c_2c_3) \times \prod_{i=5}^{length(w)} P_i^C(c_i|c_1c_2c_3 \dots c_{i-1}), \quad (1)$$

In equation (1), $P^C(w)$ is the character-level probability of word w and c_n represents the n -th character in the word w . In this way, among the pairs of words in three easy, medium, and hard categories, 3 words from the easy group, 6 words from the middle group, and 3 words from the hard group were selected as the final candidates, in a way that they could cover the maximum number of Persian characters possible. According to the word-level bigrams with previous words, the word-level probability was computed using (2). As a result, 12 words containing 29 different Persian characters, were used in the data collection and testing system.

$$P^W(w) = P_2^W(w|w_{prev}) \quad (2)$$

At the final stage, for each of these 12 selected words, a sentence was manually generated, in which way these words are put into language context. To test the system, subjects were asked to spell these 12 target words. Tables 1, 2, and 3 show the final target words that were selected. The two bold words represent the selected pairs of words, and the underlined words are the words that were spelled in the EEG data acquisition process using the stimuli screen.

Table 1. Selected easy words and sentences made for each one.

Selected Word	Sentences
خبر [News /xabar/]	مدیر کل نتوانست مانع انتشار <u>خبر</u> اختلاس او شود. [The general manager could not prevent the spreading of his embezzlement news.]
مسکن [Housing /maskan/]	جوانان برای دریافت وام <u>مسکن</u> ثبت نام نمودند. [Youths registered to receive the housing loan.]
موثر [Effective /moaser/]	محیط خانواده <u>نقشی موثر</u> در جامعه ایفا می‌کند. [Family environment plays an effective role in society.]

Table 2. Selected medium words and sentences made for each one.

Selected Word	Sentences
تصویری [Visual /tasviri/]	سخنرانی‌های وی در آرشیو <u>تصویری</u> موجود است. [His lectures are available in the visual archive.]
هموطنان [Countrymen /hamvatanan/]	عید را به <u>عموم هموطنان</u> تبریک عرض می‌کنیم. [We wish all our countrymen a happy Eid]
داغ	در سرمای زمستان یک فنجان <u>چای داغ</u> خیلی می‌چسبد.

² <https://www.isna.ir/>

[Hot /dagh/]	[In the chilly winter air, it feels so good to drink a hot cup of tea.]
پنجشنبه	برای بازدید از موزه، عصر پنجشنبه را انتخاب کرده‌ایم.
[Thursday /panjshanbe/]	[We have chosen a Thursday evening to visit the museum.]
چرب	از مصرف غذای چرب بپرهیزید.
[Fatty /charb/]	[Avoid use of fatty meals.]
اسبق	کریم باقری بازیکن اسبق تیم پرسپولیس است.
[Former /asbagh/]	[Karim Bagheri is a former player of Perspolis f.c.]

Table 3. Selected hard words and sentences made for each one.

Selected Word	Sentences
نفوذی	این منبع از دستگیری عاملان نفوذی خبر داد.
[Intrusive /nofuzi/]	[The source reported the arrest of the intrusive agents.]
حوض	گریهی همسایه همدم تنهایی ماهی حوض بود.
[Pool /howz/]	[The neighbor's cat was the companion of the pool fish's solitude.]
گلزنی	این بازیکن رکورد گلزنی لیگ را شکست.
[Scoring /golzani/]	[This player broke the scoring record of the league.]

3.3. Data Acquisition

At first, the selected target words were displayed to the users. Then for each word, characters were shown to the users once at a time, and users were asked to select these characters using the screen of the stimuli that are in front of them. To this end, users were asked to focus on the stimuli and also count the number of times the color of the stimuli is changed. The focusing of the users on the stimuli leading to enabling SSVEP and counting the color changes resulting in the P300 component excitation.

An EEG recording device made by g.Tec was used to recording EEG data. This device had 63 active electrodes, which were connected to a g.Hlamp amplifier. The ground electrode was placed on the forehead and the reference electrode was placed on the right ear lobe. Followed by the 10-20 standard pattern, figure 2 shows how the electrodes were located during the data acquisition process.

Figure 2. location of EEG recording electrodes

The data was collected in a way that the processing and classification of EEG signals could be done in the future. The data collection process was carried out at the National Brain Mapping Lab, located at the College of Engineering at the University of Tehran. The duration of the presence for each subject in the lab was about 1.5 to 2 hours, depending on the time of the checkup before the recording process, and the time it took to setup the EEG cap. All subjects participated voluntarily in the process of recording the EEG signal. To confirm their consent, the consent form prepared for the subjects was completed. For all stages of this study, including the recording and processing of EEG data, a medical ethics license was received.

EEG data were recorded from 5 subjects, including 3 females and 2 male ones. It should be noted that all of these subjects were physically healthy and right-handed. In order to reduce the effect of the sequence of user-spelled words, this sequence was different for each user. Table 4 shows the sequence of these words for each subject.

Table 4. Sequentially arranged words for each subject.

Subject	Word Sequence
1	گلزنی- هموطنان- تصویری- اسبق- خبر- داغ- نفوذی- چرب- مسکن- پنجشنبه- موثر- حوض
2	اسبق- موثر- حوض- تصویری- هموطنان- خبر- گلزنی- پنجشنبه- داغ- نفوذی- چرب- مسکن
3	تصویری- نفوذی- چرب- خبر- هموطنان- پنجشنبه- حوض- موثر- گلزنی- اسبق- مسکن- داغ
4	داغ- تصویری- موثر- پنجشنبه- گلزنی- مسکن- خبر- هموطنان- چرب- اسبق- نفوذی- حوض
5	تصویری- چرب- پنجشنبه- حوض- داغ- گلزنی- اسبق- نفوذی- خبر- مسکن- هموطنان- موثر

For each character, the color of each of the six stimuli on the screen was changed 12 times, for each of the two steps of group selection and the selection of the target character, it was repeated. Between every 4 repetitions of the change in the color of all the stimuli, 2 seconds of the break were considered. To reduce the effect of eye-movement artifacts, subjects were asked to avoid blinking when showing stimuli. Subjects could use the two-second rest period to blink and rest their eyes. Figure 3 shows the process of collecting EEG data from two subjects.

Figure 3. Process of recording EEG data from a subject

3.4. Data processing and classification

The EEG data was stored in the raw format (without filtering) and a sampling rate of 512 Hz was used. To evaluate the available data, it was necessary to preprocess the data. For this purpose, the EEG signal was first filtered using a band-pass filter in the range of 0.5 to 45 Hz. Then, using the ICA method, eye artifacts were manually removed from the filtered signal. All these processes were performed by the EEGLab toolbox in MATLAB environment.

To detect a character using the P300 signal, it is first necessary to classify the EEG samples after a change of color before the P300 signal is detected. For this purpose, an SVM classifier was used. Dimensions of the 600-millisecond-signal for all EEG channels were reduced after each change by using a moving average filter of order 8. Then the range of signal changes across all channels was normalized between zero and one. In the end, by concatenating the outputs for all channels, a feature vector was extracted for each data sample. To do labeling these samples, the ones that were related to the change of color of a target stimulus were labeled as P300, while the others were labeled as non-P300.

Using 9-fold cross-validation, the SVM classifier was trained and tested. When classifying the samples given to the SVM. In addition to determining the class of that sample (P300 or non-P300), its distance from the decision boundary was calculated. This model classifies the samples with a negative distance from the decision boundary as P300 and positive distance samples as non-P300.

By using a certain number of stimuli epochs (a complete change of the color of all the stimuli), the values of the distance between all the samples pertaining to a stimulus were pooled together. In this case, the stimulus, whose sum of the intervals is smallest (with negative and positive values), was chosen as the selected stimulus by the user. Assuming that $dist(s_i)$ is equal to the sample distance of the EEG sample after the stimulus is intensified, the SVM classifier boundary is used, n is the number of epochs used to detect the stimulus, the sum of the sample distance values after the stimulus 's' The classification boundary was calculated using equation (2):

$$dist_t(s) = \sum_{i=1}^n dist(s_i), \quad (2)$$

Assuming that $dist_t(s^j)$ represents the sum of the distance values of the EEG signal samples after the occurrence of the stimulus s^j , the target stimulus was selected using the equation (3):

$$target = \underset{j}{\operatorname{argmin}}(dist_t(s^j)), \quad (3)$$

3.5. Using language model for character recognition

For character detection, the language model was applied to the result of the P300 classification. Character-level N-gram of N values between 2 and 5 was matched with the addition of a smoothing method. The numerical values used to recognize the character using the P300 signal were similar using the same model, and the values of the space used for character recognition using the P300 were converted to probabilities using the softmax function, as shown in equation (4):

$$Prob_{P300}(s^i) = \frac{e^{-dist_t(s^i)}}{\sum e^{-dist_t(s^i)}}, \quad (4)$$

In equation (4), $prob_t(s^j)$ is the probability of the target stimulus s^j and $dist_t(s^j)$ equivalent to the sum of the signal sample distances from the SVM decision boundary.

To calculate the probability of choosing a stimulus using the language model, if the stimulus consists of a group of characters, the probability is calculated by using equation (5) and if the stimuli contain a character, the equation (6) is used to calculate the probability. In these two equations, $Prob_{N-Gram}(s^i)$ indicates the probability of selected target stimulus s^i using the N-gram, and $P_N(c_k^j | c_{k-N+1} \dots c_{k-1})$ is the occurrence

probability of the j -th character at the k -th position of the sentence with the condition of the N-1 previous character(s).

$$Prob_{N-gram}(s^i) = \frac{1}{6} \sum_{j=1}^6 P_N(c_k^j | c_{k-N+1} \dots c_{k-1}), \quad (5)$$

$$Prob_{N-gram}(s^i) = P_N(c_k | c_{k-N+1} \dots c_{k-1}), \quad (6)$$

In equation (7), the selection of a character is obtained by combining the results of the P300 detection and the N-gram language model:

$$target = \underset{i}{\operatorname{argmin}}(Prob_{P_{300}}(s^i) \times [Prob_{N-gram}(s^i)]^\omega), \quad (7)$$

In equation (7), the parameter ω is the weight of the probability of the language model and the probability of the P300 classification. In this study, this number was chosen experimentally by 0.13, which yielded the best result.

4. Evaluation Results

In this section, the results of our evaluation for the given speller are presented. In the evaluations, we have used well-known evaluation criteria, such as accuracy, precision, and recall [39]. The primary results on the classification of the P300 signals are presented in Table 5.

Table 5. Results for P300 classification.

Subject	Accuracy	Precision	Recall
1	86.9	77.6	29.9
2	88.1	83.5	36.0
3	90.0	81.3	51.6
4	93.6	91.6	67.9
5	89.9	89.8	44.4
Average	89.7	84.8	46.0

Despite the high accuracy and precision, the results in Table 5 show that the SVM classifier was not capable of achieving a good recall. This means that most of the P300 samples were mistakenly classified as non-P300. To improve the performance, the language model score is integrated with the P300 classification score to classify the characters. The results of character recognition are presented in Table 6 for different levels of the N-gram language model. In this table, the results are given for different numbers of epochs, as well. The last column of this table, i.e., Avg, denotes the average accuracy of the corresponding row.

Table 6. Accuracy (%) results for P300 classification on character recognition using the language model[illegible]

P300 with 1_gram language model	1	37.0	57.4	79.6	87.0	92.6	94.4	94.4	92.6	96.3	100	100	100	86.0
	2	59.3	83.3	92.6	96.3	98.1	98.1	98.1	100	100	100	100	100	93.8
	3	27.8	68.5	83.3	87.0	90.7	94.4	94.4	98.1	98.1	98.1	98.1	98.1	86.4
	4	75.9	90.7	94.4	98.1	98.1	98.1	98.1	100	100	100	100	100	96.1
	5	75.9	96.3	98.1	98.1	100	100	100	100	100	100	100	100	97.4
P300 with 2_gram language model	1	40.7	59.3	79.6	88.9	88.9	94.4	92.6	92.6	96.3	100	100	100	86.1
	2	55.6	87.0	90.7	94.4	96.3	98.1	98.1	100	100	100	100	100	93.4
	3	29.6	68.5	79.6	90.7	92.6	96.3	98.1	98.1	98.1	98.1	98.1	98.1	87.2
	4	74.1	92.6	94.4	100	100	100	98.1	100	100	100	100	100	96.6
	5	72.2	94.4	98.1	98.1	100	100	100	100	100	100	100	100	96.9
P300 with 3_gram language model	1	48.1	59.3	79.6	90.7	88.9	96.3	96.3	96.3	100	100	100	100	88.0
	2	63.0	87.0	90.7	98.1	98.1	98.1	100	100	100	100	100	100	94.6
	3	31.5	72.2	77.8	92.6	90.7	96.3	98.1	98.1	98.1	98.1	98.1	98.1	87.5
	4	75.9	90.7	94.4	100	100	100	98.1	100	100	100	100	100	96.6
	5	70.4	92.6	100	100	100	100	100	100	100	100	100	100	96.9
P300 with 4_gram language model	1	37.0	64.8	83.3	90.7	88.9	94.4	98.1	98.1	100	100	100	100	88.0
	2	61.1	79.6	90.7	98.1	100	100	100	100	100	100	100	100	94.1
	3	31.5	70.4	75.9	92.6	90.7	96.3	98.1	98.1	98.1	98.1	98.1	98.1	87.2
	4	83.3	92.6	98.1	100	100	100	100	100	100	100	100	100	97.8
	5	66.7	96.3	100	100	100	100	100	100	100	100	100	100	96.9
P300 with 5_gram language model	1	40.7	68.5	83.3	92.6	90.7	94.4	96.3	98.1	100	100	100	100	88.7
	2	55.6	83.3	92.6	98.1	100	100	100	100	100	100	100	100	94.1
	3	29.6	72.2	81.5	88.9	90.7	96.3	98.1	98.1	98.1	98.1	98.1	98.1	87.3
	4	87.0	94.4	98.1	100	100	100	100	100	100	100	100	100	98.3
	5	75.9	96.3	100	100	100	100	100	100	100	100	100	100	97.7

The results in Table 6 show that the method used to recognize characters using the P300 achieved high accuracy, and with an average of approximately 93%, with four display stimulus epochs. By applying the language model in the P300 classification, the character recognition for all subjects, except subject number 3, the accuracy eventually reached 100% with the increasing number of epochs. We believe that the reason for not having 100% accuracy of the classification of P300 in subject number 3 is due to a mistake by the subject; in another word, subject 3 probably selected characters other than the target character in one of the steps.

To more precisely study the impact of the language model on the P300 classification of characters, the performance was calculated for each difficulty level, separately. Table 7 shows the results of character recognition in each of the easy, medium, and hard groups. Each accuracy score of this table denotes the average accuracy for all 5 subjects.

Table 7. The accuracy of the character recognition in each easy, medium, and hard word-level using the P300 and N-gram language model in the character level

Type of Classification	Word- Level Difficulty	Number of epochs												Avg
		1	2	3	4	5	6	7	8	9	10	11	12	
P300 without a language model	Easy	58.1	74.5	89.0	94.5	96.3	94.5	98.1	98.1	98.1	98.1	98.1	98.1	91.3
	Medium	50.6	81.3	88.6	94.6	95.3	97.3	96.0	97.3	98.6	100	100	100	91.6
	Hard	47.6	75.3	84.6	87.6	93.8	98.4	96.9	98.4	100	100	100	100	90.2

P300 with 1_gram language model	Easy	60.0	76.3	89.0	94.5	96.3	96.3	98.1	98.1	98.1	98.1	98.1	98.1	91.8
	Medium	55.9	82.0	91.3	94.0	95.3	96.6	96.0	98.6	99.3	100	100	100	92.4
	Hard	49.2	75.3	86.1	90.7	96.9	98.4	98.4	96.9	98.4	100	100	100	90.8
P300 with 2_gram language model	Easy	65.4	76.3	89.0	96.3	96.3	96.3	98.1	98.1	98.1	98.1	98.1	98.1	92.4
	Medium	53.3	82.6	91.3	94.6	96.0	98.0	97.3	98.6	99.3	100	100	100	92.6
	Hard	47.6	78.4	81.5	92.3	93.8	98.4	96.9	96.9	98.4	100	100	100	90.3
P300 with 3_gram language model	Easy	70.9	80.0	87.2	96.3	96.3	96.3	98.1	98.1	98.1	98.1	98.1	98.1	93.0
	Medium	56.6	83.3	91.3	97.3	96.6	98.6	98.6	99.3	100	100	100	100	93.4
	Hard	49.2	73.8	83.0	93.8	92.3	98.4	98.4	98.4	100	100	100	100	90.6
P300 with 4_gram language model	Easy	67.2	80.0	89.0	96.3	96.3	96.3	98.1	98.1	98.1	98.1	98.1	98.1	92.8
	Medium	54.6	84.6	91.3	97.3	97.3	98.6	100	100	100	100	100	100	93.6
	Hard	49.2	72.3	86.1	93.8	92.3	98.4	98.4	98.4	100	100	100	100	90.7
P300 with 5_gram language model	Easy	70.9	78.1	90.9	92.7	96.3	96.3	98.1	98.1	98.1	98.1	98.1	98.1	92.8
	Medium	56.6	87.3	92.6	98.0	98.0	98.6	99.3	100	100	100	100	100	94.2
	Hard	49.2	76.9	87.6	93.8	92.3	98.4	98.4	98.4	100	100	100	100	91.2

The results in Table 7 show that increasing the amount of N in the N-gram language models (i.e. increasing the context window) in the words of the medium group has the most positive effect and even more than the easy group. In the hard group, the 1-gram language model is better than the 2-gram, 3-gram, and 4-gram. Also, the 5-gram language model is better than other models.

5. Discussion and Conclusions

In this work, a stimuli screen was designed based on the Hex-O-Spell model. Then, using the language model in this speller, target words with three different levels of difficulty were selected. After collecting the EEG data from 5 subjects, the P300 classification and the character detection were tested. In the character recognition using the P300 signal, an average accuracy of 91.3 % was obtained. The used language model was a combination of character-level N-grams with values of N from 1 to 5 and a word-level bigram. The comparison of the average accuracy of the P300 classification for character recognition by different methods used in this study is summarized in Figure 4. The results of these N-grams showed that the use of the language model generally increased the precision of the P300 classification in character recognition. It was also concluded that by increasing the number of N in the N-gram language models, the character recognition accuracy was increased, and as a result, when using the 5-gram language model, the maximum average accuracy of 93.2% was achieved (a 1.9% improvement from when only P300 is used).

Figure 4. Comparison of average accuracy of proposed methods for character recognition

As we had no access to any similar data to compare our method with, to better understand the results of this study, these results are compared with the results presented in the study by Cecotti and Graser [40]. The average accuracy values provided by various methods in the classification of P300 (Hofftam, Zongton, Yandong, mLVQ, LDA, ESVM CNN-1, MCNN-1) are given in the character detection of the research [40], based on the number of epochs which are given in Table 8. The dataset collected in the study [40] contained 29 different characters from the 36 characters shown on the screen, which has a similar approach to this study. The results of the comparison of these two studies show that this study with 12 epochs even when not using the language model, has had better results than the 5 and 10 epochs in the methods mentioned in the study [40]. Although the results are not quantitatively comparable due to differences in data, they can give the insight to understand the position of our work besides similar research.

Table 8. Comparison of the average accuracy for different P300 classification methods modified from [40]

Method	Hofftam	Zongton	Yandong	mLVQ	LDA	ESVM	CNN-1	MCNN-1
No. of Epochs								
5	53	59.5	55	59.5	60.5	73.5	70	69
10	-	-	-	84	85	87	88.5	87
15	89.5	90.5	90.5	91.5	92.7	96.5	94.5	95.5

Our method can be improved in various directions including the improvement of language models and the smoothing methods, enriching the dataset, and also benefiting from hybrid BCI techniques. Hybrid BCIs combine two or more different types of BCIs and also could improve universality because a second approach might convey the user's intent if another approach cannot. Also, the use of multiple types of brain signals as inputs to a BCI system can increase the accuracy and speed of the system. In the process of collecting EEG data, the stimuli screen was designed to induce SSVEP in the subjects' brain signals for the Persian speller, but it was not modeled in this work.

Declarations

Ethics Approval and Consent to Participate: This study was approved by the Tehran University of Medical Sciences (approval no. IR.TUMS.VCR.REC.1395.1214).

Consent for Publication: Written informed consent was obtained from the users for the publication of their data.

Availability of Data and Material: Data and material of this research are available upon request.

Competing Interests: The authors declare that there is no conflict of interest regarding the publication of this paper.

Funding: This project was funded by the Iran National Science Foundation under project number 95844527. However, the responsibility for the content of this study lies with the authors alone.

Authors' Contributions: Hadi supervised the project, advised in developing the concept and methods, advised in designing the experiments, edited the manuscript, and is the corresponding author of the paper. Hessam developed the concept, implemented the algorithms, did the simulations designed data collection, collected and prepared the data, prepared the first draft of the paper. Parisa did some implementation, did several evaluations, prepared some parts in the first draft of the paper. All authors discussed the final results.

Acknowledgment: The authors would like to thanks the Iran National Science Foundation which supports this research under project number 95844527.

References

1. Mora-Cortes, A.; Manyakov, N.; Chumerin, N.; Van Hulle, M. Language model applications to spelling with brain-computer interfaces. *Sensors* 2014, 14(4), 5967-93.
2. Nicolas-Alonso, L.F.; Gomez-Gil, J. Brain computer interfaces, a review. *Sensors* 2012, 12(2), 1211-79.
3. Abdulkader, S.N.; Atia, A.; Mostafa, M.S. Brain computer interfacing: Applications and challenges. *Egyptian Informatics Journal* 2015, 16(2), 213-30.
4. Prashant, P.; Joshi, A.; Gandhi, V. Brain computer interface: A review. In Engineering (NUICONE), 2015 5th Nirma University International Conference, Ahmedabad, India, 2015 Nov 26-28, (pp. 1-6); IEEE: 2016 April 11.
5. Arafat, I. Brain-computer interface: past, present & future. *International Islamic University Chittagong (IIUC)*, Chittagong, Bangladesh. 2013.
6. Grosselin, F.; Navarro-Sune, X.; Vozzi, A.; Pandremmenou, K.; De Vico Fallani, F.; Attal, Y.; Chavez, M. Quality Assessment of Single-Channel EEG for Wearable Devices. *Sensors* 2019, 19(3), 601.
7. Maiorana, E.; Solé-Casals, J.; Campisi, P. EEG signal preprocessing for biometric recognition. *Machine Vision and Applications* 2016, 27(8), 1351-60.
8. Smith, S.W. Moving Average Filters. In *The scientist and engineer's guide to digital signal processing*, California Technical Publishing: San Diego, CA, 1997; pp. 277-284.

9. Jebelli, H.; Hwang, S.; Lee, S. EEG signal-processing framework to obtain high-quality brain waves from an off-the-shelf wearable EEG device. *Journal of Computing in Civil Engineering* 2017, 32(1), 04017070.
10. Mak, J.N.; Wolpaw, J.R. Clinical applications of brain-computer interfaces: current state and future prospects. *IEEE reviews in biomedical engineering* 2009, 2, 187-99.
11. Fazel-Rezai, R.; Ahmad, W. P300-based brain-computer interface paradigm design. In *Recent Advances in Brain-Computer Interface Systems*, Fazel-Rezai, R., Eds.; InTech: London, United Kingdom, 2011; pp. 83-98.
12. Pérez-Vidal, A.; Garcia-Beltran, C.; Martínez-Sibaja, A.; Posada-Gómez, R. Use of the Stockwell Transform in the Detection of P300 Evoked Potentials with Low-Cost Brain Sensors. *Sensors* 2018, 18(5), 1483.
13. Amini, H.; Veisi, H.; Mohammadi, E. Target words selection for a Persian brain-computer-interface-based speller using language model. In 2016 Eighth International Conference on Information and Knowledge Technology (IKT), Hamedan, Iran, 2016 Sep 7, (pp. 216-220); IEEE: 2016 Dec 12.
14. Orhan, U.; Erdogmus, D.; Roark, B.; Purwar, S.; Hild, K.E.; Oken, B.; Nezamfar, H.; Fried-Oken, M. Fusion with language models improves spelling accuracy for ERP-based brain computer interface spellers. In 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA, 2011 Aug 30, (pp. 5774-5777); IEEE: 2011 Dec 1.
15. Lotte, F.; Congedo, M.; Lécuyer, A.; Lamarche, F.; Arnaldi, B. A review of classification algorithms for EEG-based brain-computer interfaces. *Journal of neural engineering* 2007, 4(2), R1-R13.
16. L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and clinical Neurophysiology*, vol. 70, no. 6, pp. 510-523, 1988.
17. A. Furdea, S. Halder, D. Krusienski, D. Bross, F. Nijboer, N. Birbaumer, and A. Kübler, "An auditory oddball (P300) spelling system for brain-computer interfaces," *Psychophysiology*, vol. 46, no. 3, pp. 617-625, 2009.
18. M. Chang and T. M. Rutkowski, "Two-Step Input Spatial Auditory BCI for Japanese Kana Characters," in *Advances in Cognitive Neurodynamics (V)*: Springer, 2016, pp. 383-389.
19. A.-M. Brouwer and J. B. Van Erp, "A tactile P300 brain-computer interface," *Frontiers in neuroscience*, vol. 4, 2010.
20. R. Fazel-Rezai and K. Abhari, "A region-based P300 speller for brain-computer interface," *Canadian Journal of Electrical and Computer Engineering*, vol. 34, no. 3, 2009.
21. B. Blankertz, G. Dornhege, M. Krauledat, M. Schröder, J. Williamson, R. Murray-Smith, and K.-R. Müller, "The Berlin Brain-Computer Interface presents the novel mental typewriter Hex-o-Spell," 2006.
22. Orhan, U.; Hild, K.E.; Erdogmus, D.; Roark, B.; Oken, B.; Fried-Oken, M. RSVP keyboard: An EEG based typing interface. In 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, 2012 Mar 25, (pp. 645-648); IEEE: 2012 Aug 31.
23. L. Acqualagna, M. S. Treder, M. Schreuder, and B. Blankertz, "A novel brain-computer interface based on the rapid serial visual presentation paradigm," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, 2010, pp. 2686-2689: IEEE.
24. R. Fazel-Rezai and K. Abhari, "A comparison between a matrix-based and a region-based P300 speller paradigms for brain-computer interface," in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, 2008, pp. 1147-1150: IEEE.
25. X. Chen, Y. Wang, M. Nakanishi, T.-P. Jung, and X. Gao, "Hybrid frequency and phase coding for a high-speed SSVEP-based BCI speller," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, 2014, pp. 3993-3996: IEEE.

26. M. Nakanishi, Y. Wang, X. Chen, Y.-T. Wang, X. Gao, and T.-P. Jung, "Enhancing Detection of SSVEPs for a High-Speed Brain Speller Using Task-Related Component Analysis," *IEEE Transactions on Biomedical Engineering*, 2017.
27. M. Nakanishi, Y. Wang, Y.-T. Wang, Y. Mitsukura, and T.-P. Jung, "A high-speed brain speller using steady-state visual evoked potentials," *International journal of neural systems*, vol. 24, no. 06, p. 1450019, 2014.
28. B. Z. Allison, C. Brunner, V. Kaiser, G. R. Müller-Putz, C. Neuper, and G. Pfurtscheller, "Toward a hybrid brain-computer interface based on imagined movement and visual attention," *Journal of neural engineering*, vol. 7, no. 2, p. 026007, 2010.
29. M. Wang, I. Daly, B. Z. Allison, J. Jin, Y. Zhang, L. Chen, and X. Wang, "A new hybrid BCI paradigm based on P300 and SSVEP," *Journal of Neuroscience Methods*, vol. 244, pp. 16-25, 2015.
30. M. Xu, H. Qi, B. Wan, T. Yin, Z. Liu, and D. Ming, "A hybrid BCI speller paradigm combining P300 potential and the SSVEP blocking feature," *Journal of neural engineering*, vol. 10, no. 2, p. 026001, 2013.
31. B. Z. Allison, J. Jin, Y. Zhang, and X. Wang, "A four-choice hybrid P300/SSVEP BCI for improved accuracy," *Brain-Computer Interfaces*, vol. 1, no. 1, pp. 17-26, 2014.
32. S. Amiri, R. Fazel-Rezai, and V. Asadpour, "A review of hybrid brain-computer interface systems," *Advances in Human-Computer Interaction*, vol. 2013, p. 1, 2013.
33. Y. Yu, Z. Zhou, J. Jiang, E. Yin, K. Liu, J. Wang, Y. Liu, and D. Hu, "Toward a Hybrid BCI: Self-Paced Operation of a P300-based Speller by Merging a Motor Imagery-Based "Brain Switch" into a P300 Spelling Approach," *International Journal of Human-Computer Interaction*, pp. 1-10, 2017.
34. Hosni, Sarah M., Howida A. Shedeed, Mai S. Mabrouk, and Mohamed F. Tolba. "EEG-EOG based virtual keyboard: Toward hybrid brain computer interface." *Neuroinformatics*, 2018, 1-19.
35. F. Najafi, "Improvement of Stimulation Method in BCI Speller Systems with P300 and Presentation of New Protocol for Persian Typing" M.Sc., Faculty of Engineering, Shahed University, 2014.
36. Treder, M.S.; Blankertz, B. (C) overt attention and visual speller design in an ERP-based brain-computer interface. *Behavioral and brain functions* 2010, 6(1), 28.
37. Fried-Oken, M.; Mooney, A.; Peters, B.; Oken, B. A clinical screening protocol for the RSVP keyboard brain-computer interface. *Disability and Rehabilitation: Assistive Technology* 2015, 10(1), 11-8.
38. Brainard, D. H., & Vision, S., The psychophysics toolbox. *Spatial vision*, 1997, 10, 433-436.
39. Bishop, C. M., Pattern recognition and machine learning. Springer, 2006.
40. Cecotti, H.; Graser, A. Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE transactions on pattern analysis and machine intelligence* 2011, 33(3), 433-45.

Hadi Veisi received his Ph.D. in Artificial Intelligence from the Sharif University of Technology in 2011. He joined the University of Tehran, Faculty of New Sciences and Technologies (FNST) in 2012, and established the Data and Signal Processing (DSP) lab. The main research interests of Hadi are artificial neural network and deep learning, natural language processing, and signal processing. Email adresse: h.veisi@ut.ac.ir, ORCID: 0000-0003-2372-7969.

Hesam Amini graduated with a master's degree in Information Technology at the University of Tehran in 2017. He received his bachelor of science degree from the Ferdowsi University of Mashhad in 2015. He is a member of Computational Linguistics at Concordia (CLaC) Laboratory, Montreal Natural Language Processing (MTL-NLP), Concordia University, and pursuing his Ph.D. degree in that department. His research interests are Artificial Intelligence (AI) and Machine Learning (ML) techniques in Natural Language Processing (NLP) applications.. Email address: hessam.amini@concordia.ca.

Parisa Daeijavad graduated with a master's degree in Computer Engineering at the University of Tehran in 2019. She received her bachelor of science degree from the University of Isfahan in 2016. Her research interests are cognitive science and machine learning. Email address: parisa.daeijavad@ut.ac.ir.