Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi

Pamukkale University Journal of Engineering Sciences

Brain-Computer ınterfaces with auditory stimuli: A review study

İşitsel uyaranli beyin-bilgisayar arayüzleri: Bir derleme çalişmasi

Bahar NAZLI[1], Ahmet Reşit KAVSAOĞLU[1]*

[1]Department of Biomedical Engineering, Karabuk University, Karabuk, Turkey
btastan@kastamonu.edu.tr, kavsaoglu@karabuk.edu.tr

Received/Geliş Tarihi: 28.03.2024 Revision/Düzeltme Tarihi: 16.10.2024 doi: 10.5505/pajes.2024.57958
Accepted/Kabul Tarihi: 24.10.2024 Review Article/Derleme Makalesi

Abstract

This study examines auditory stimulus interface studies, an essential development in brain-computer interfaces. Brain-computer interfaces help individuals with limited motor skills communicate without any muscle intervention. Electroencephalogram is the most frequently used method in brain-computer interface studies. Interfaces developed using auditory evoked potentials obtained from electroencephalogram signals are used in fields such as auditory spellers, mood research, and device control. However, compared to visual and tactile stimuli, interfaces with auditory stimuli appear to perform lower in accuracy and information transfer. Interfaces based on auditory stimuli are essential because they allow communication in individuals with no muscle activity and limited vision. This study reviews the design, classification, and evaluation stages of auditory brain-computer interfaces and examines the studies done in the literature.

Keywords: Brain-computer interface, Auditory evoked potential, Electroencephalogram

Öz

Bu çalışma, beyin-bilgisayar arayüzleri alanında önemli bir gelişme olan işitsel uyaranlı arayüz çalışmalarının incelenmesine odaklanmaktadır. Beyin-bilgisayar arayüzleri, motor becerileri kısıtlı bireylerin herhangi bir kas müdahalesi olmadan iletişim kurmalarına yardımcı olmaktadır. Elektroensefalogram beyin bilgisayar aryüzü çalışmalarında en sık kullanılan yöntemdir. Elektroensefalogram sinyallerinden elde edilen işitsel uyarılmış potansiyeller kullanılarak geliştirilen arayüzler işitsel heceleyici, duygu durum araştırmaları ve cihaz kontrolü gibi farklı alanlarda kullanılmaktadır. Ancak görsel ve dokunsal uyaranlarla karşılaştırıldığında, işitsel uyaranlı arayüzlerin doğruluk ve bilgi aktarımı açısından daha düşük performans gösterdiği görülmektedir. Herhangi bir kas aktivitesi bulunmyan ve aynı zamanda görme yetisi de kısıtlanmış bireylerde işitsel uyaranlara dayalı arayüzler, iletişim kurma imkanı sunması nedeniyle önemlidir. Bu çalışma, işitsel uyaranlı beyin-bilgisayar arayüzlerinin tasarım, sınıflandırma ve değerlendirme aşamaları hakkında bir inceleme sunmakta ve literatürde yapılmış olan çalışmaları incelemektedir.

Anahtar kelimeler: Beyin bilgisayar aryüzü, İşitsel uyarılmış potansiyel, Elektroensefalogram

1 Introduction

Brain-computer interface (BCI) is a system that allows people to interact with their environment using control signals generated from electroencephalographic activity without the intervention of peripheral nerves and muscles [1]. BCIs can be used for various purposes in individuals with loss of cognitive or sensorimotor functions [2]. They are used to map and examine cognitive functions while enabling people with loss of motor functions to develop and improve their communication skills. One of the primary purposes of BCIs is for people with severe motor disabilities to create a communication channel with their environment through computers without using their limbs [3]. Conscious but paralyzed patients cannot express themselves. Diseases such as locked-in syndrome also cause vision loss [4]. Brain-computer interface studies use auditory stimulation to improve the quality of life of people with limited or no vision suffering from these and similar diseases and to integrate them into life [5]. Brain signals recorded using auditory stimuli in visually impaired individuals are used in BCI studies. Basically, taking into account the situation of the people to whom the stimulus will be given, people are warned by using one or more visual, tactile, and auditory stimuli in BCI studies. As a result of this stimulation, brain potentials measured using different methods are recorded and processed. Various methods are available to measure brain activity, including electroencephalogram (EEG), electrocorticogram (ECoG), single unit recording (SUR), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), Positron emission tomography (PET), and near-infrared spectroscopy (NIR) [6], [7]. An electrocorticogram (ECoG) is a neurophysiological recording method that directly measures electrical activity from the cerebral cortex via electrodes placed on the brain's surface [8]. The ECoG method is invasive and usually requires a surgical procedure. The skull is opened, and electrodes are placed directly on the brain's surface to detect electrical signals between brain cells [9]. ECoG is mainly used in BCI studies to control prosthetic limbs and communication devices. Single-unit recording is a neurophysiological technique used to directly measure the action potentials of a single neuron using microelectrodes[10]. This technique plays a vital role in brain research and understanding the functioning of the nervous system. Magnetoencephalography is a noninvasive test that measures magnetic fields generated by electrical currents in the brain [11]. The magnetic field effect of the brain is much smaller than the magnetic effect of any metal material in the environment. Therefore, MEG measurements should be performed in isolated rooms to avoid being affected by ambient magnetic fields. MEG has a very high temporal resolution and good spatial resolution with a wide frequency

*Corresponding author/Yazışılan Yazar

range [12]. MEG is often combined with a magnetic resonance imaging (MRI) for an excellent structural perspective. MEG is used in scientific research aimed at determining the functions of specific brain regions, in clinical diagnostics, and as an examination during neurosurgical operations to localize pathological regions. Functional magnetic resonance imaging is a methodology used to observe the working human brain. fMRI measures the increase in blood flow and oxygenation of the active part of the brain [13]. fMRI measurements are a noninvasive technique that does not require the injection of contrast media or radioactive isotopes. Blood oxygenation level dependent response (BOLD) describes the dependence of magnetic resonance signal intensity on blood oxygenation level. The concept of fMRI is based on MRI examination and its extension with observation based on the properties of oxygenated and deoxygenated blood [14].

Positron emission tomography is an imaging technique that studies brain function in vivo and uses radiation emitted during positron annihilation [15]. In a PET scan, an isotope (radioisotope) containing a small amount of radioactive material is injected into the patient. These isotopes emit positively charged particles called positrons. Images are created using gamma rays emitted by the annihilation of positrons [16]. It is used in oncology, neurology, and cardiology. Near Infrared Spectroscopy is a noninvasive imaging technique that uses near-infrared light's absorption and reflection properties between 700 and 2500 nm passing through tissues. It measures physiological parameters such as tissue oxygen content and blood flow [11]. NIRS devices emit light with an LED or laser light source placed on the skin surface. Detectors detect the differences in light absorbed from the tissue. NIRS is used to study brain function, neurological disorders, and brain activity.

EEG is the most frequently used technique among the systems used to measure brain activity [17]. Because EEG-based BCIs are noninvasive and the devices used are portable and affordable, researchers have conducted extensive studies on them. Signals containing information about brain activity are recorded by measuring through electrodes placed on the scalp. EEG signals are non-stationary signals with a low signal-to-noise ratio [18]. EEG signals contain different sub-bands. Potentials are distinguished by separating signals of different frequencies and amplitudes that occur according to the functions of different brain lobes. Neuron activities in each different region of the brain create waves at specific frequencies. These waves are generally classified into four main categories: alpha, beta, theta, and delta [19]. Oscillatory EEG activity refers to these rhythmic electrical fluctuations that occur in different frequency ranges [20]. Alpha waves have a frequency between 8-12 Hz and are commonly seen at rest and with closed eyes. It is considered a transition period between wakefulness and relaxation. Beta waves are between 12-30 Hz. They are concentrated during wakefulness and conscious activity [21]. Mental activity, focus, and attention are associated with an increase in this wave range. Theta waves are between 4-7 Hz and are associated with deep relaxation, light sleep, creative thinking, and imagination. Delta waves are between 0.5-4 Hz and are dominant during deep sleep. It is also associated with some conditions, such as loss of mental function [22].

Various EEG paradigms examined in the literature are given in Figure 1. Event-related potentials are EEGs recorded by measuring the electrical response occurring in the cortex to emotional, sensory, or cognitive events [23]. They are produced in response to peripheral or external stimuli. These are brain activities observed due to visual, auditory, and somatosensory (sensory such as touch, pressure, pain) stimuli. They can also be observed after the expectation of conditioned stimulation or before voluntary movements.

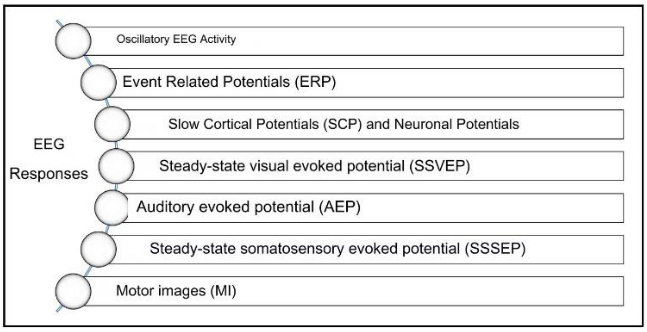


Figure 1. EEG responses used in BCI studies [24].

Slow Cortical Potentials (SCPs) and neuronal potentials are spontaneous signals [25]. These signals occur without any external stimulus affecting the user. SCPs are components of the EEG signal below 1 Hz. Steady state visual evoked potentials (SSVEP) signals are brain signals produced by visually stimulating the person with a periodically produced vibrating image [26]. SSVEP are steady-state evoked potentials that are synchronized to both the frequency and phase of the external stimulus as a result of periodic stimuli [27]. This signal is the result of visual stimulation and has a high signal to noise ratio.

Steady-state somatosensory evoked potentials (SSSEPs) is an EEG paradigm in which brain waves are continuously stimulated at a specific frequency. This method examines brain waves that occur in response to stimuli in the somatosensory system [28]. SSSEPs use a stimulus that is evoked at a fixed frequency instead of evoked potentials that usually occur in response to visual, auditory, or somatosensory stimuli. Therefore, brain waves are also expected to oscillate at a constant frequency. Motor imagery refers to a person mentally performing a movement [29]. The person imaginatively visualizes performing the movement by visualizing that movement in his mind, without making an actual physical movement. Motor imagery is when a person without a limb creates an EEG signal by imagining moving their limb.

Auditory evoked potentials (AEPs) refer to the electrical responses that occur in the brain when a sound stimulus is presented to a person's auditory system (usually through headphones) [30]. This review focuses on studies using AEPs.

2 Auditory evoked potential

After external sounds stimulate auditory receptors, the central nervous system produces AEP. The AEP signal consists of repeatable positive or negative peaks, latency, amplitude, and behavioral correlation, and their amplitudes are more diminutive than EEG signals. Transient and steady-state responses are the two classifications of AEPs [31]. Figure 2 shows the classification of transient responses based on their latency (time between stimulus onset and AEP) and amplitude into fast, middle, and late. Fast responses, named with Roman numerals as shown in Figure 2, occur 1-10 ms after the stimulus. Responses that occur between 10-50 ms are considered middle responses. Late responses occur 50-500 ms after the stimulus. The letter and the second in which it occurs are used to name these responses. P is used to indicate the positive wave, and N is used to indicate the negative wave. P300 (or P3) represents the positive wave occurring after 300 ms. Their phase and amplitude characterize steady-state responses. The steady-state response evoked by stimuli delivered at rates close to 40/second is called 40 Hz auditory steady-state response (40 Hz ASSR) [32].

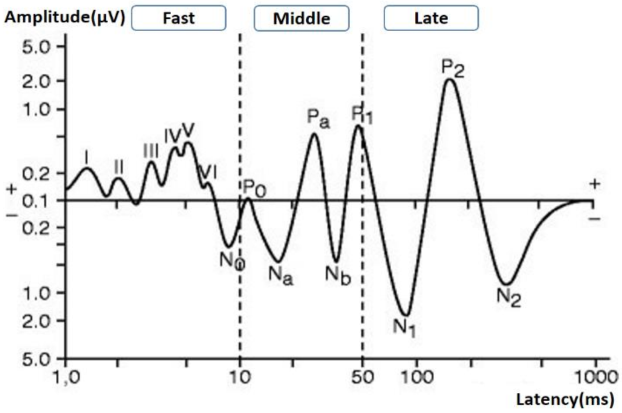


Figure 2. Classification of Transient Auditory Evoked Potentials by Latency and Amplitude [33].

The auditory BCI studies can be classified under the following headings,

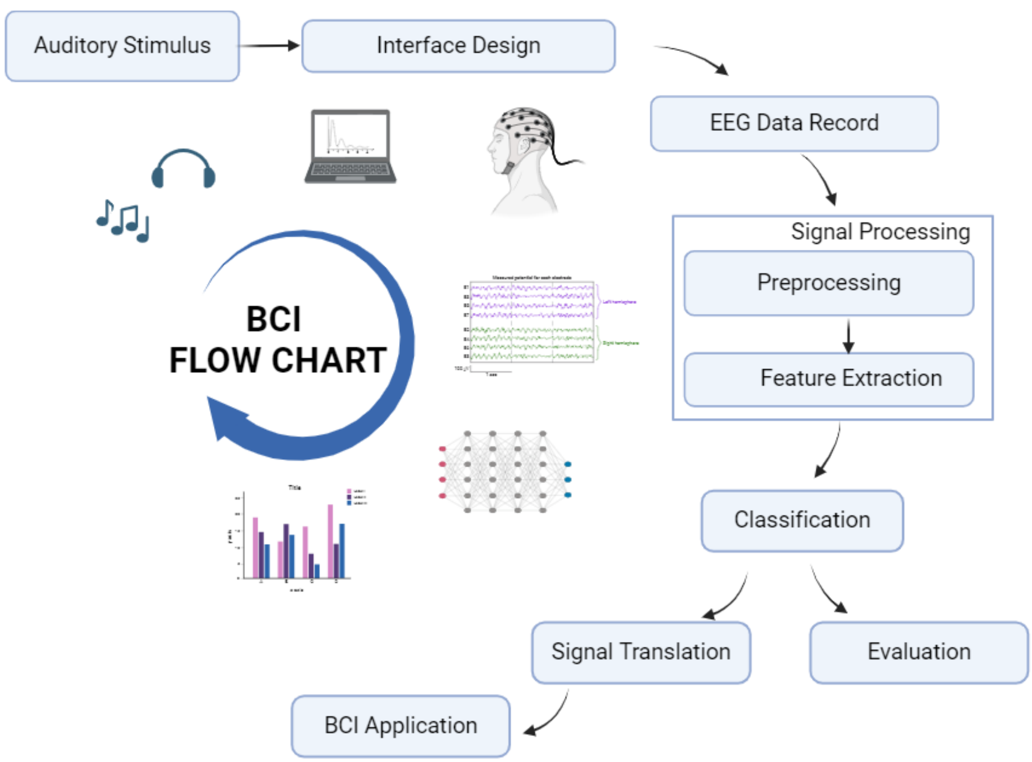
- Auditory speller [34], [35], [36], [37],
- Auditory attention selectivity [38], [39], [40],
- Discrimination of auditory stimuli [41], [42],
- Emotional studies [43], [44],
- Detecting sound direction and source [45], [46], [47], [48], [49], [50],
- Device control and communication [51], [52], [53], [54]

Auditory spellers enable letters in the form of a matrix placed on a screen to be selected using auditory stimuli [55]. It is preferred for motor-disabled individuals who are disturbed by visual stimuli and have poor eyesight to communicate. Auditory attention selectivity deals with correctly selecting the target sound among different auditory stimuli [56]. Discrimination of auditory stimuli is determined by accurately distinguishing auditory stimuli in various styles and frequencies [57]. These studies are essential for analyzing the hearing mechanism and selecting different sounds used in autonomous systems. Emotional research examines the changes in people's emotions and states with auditory stimuli [43]. The mental states of individuals are investigated. Determining the target and direction of the sound source deals with the working mechanism of auditory processes and detecting the direction and source of sound [58]. Device control and communication aims to develop various communication tools and device control systems so that individuals with motor disabilities can be included in social life.

3 Auditory evoked BCI processing steps

BCI studies are conducted by presenting humans with one or more auditory, visual, and tactile stimuli, examining brain responses, and recording potentials. The basic steps of BCI studies consist of interface design, EEG signal recording, signal processing, classification, signal translation, BCI implementation, and evaluation, as shown in Figure 3. Auditory stimuli can be presented to individuals through headphones, speakers, or loudspeakers. Stimuli can be of different types:

Figure 3. Flow chart of the auditory BCI.



short stories, virtual sounds at various frequencies, natural sounds, and loudspeakers. The stimuli are commonly presented to the subjects in a method known as the "Oddball Paradigm." The stimulus presented to the subjects consists of standard stimuli and a different stimulus that is not expected among these stimuli. This stimulus presentation method is called the Oddball paradigm [59]. In studies with evoked potentials, a certain period of time must elapse after the stimulus is presented for the potential to develop and the stimulus to be distinguished. The time between two consecutive stimuli is called inter-stimulus interval (ISI) [60]. The ISI measures the time from one stimulus's presentation to the next stimulus's presentation. ISI is widely applied in psychology, neuroscience, and perception research.

3.1 Interface Design

BCI systems consist of two essential parts: software and hardware. The hardware consists of EEG recorders, EEG caps, electrodes, and other hardware required for signal transmission. The systems developed for this purpose, which we mentioned in the introduction section, are suitable for experimental and professional use. In BCI studies, an interface is designed using different platforms. BCI interfaces can be developed using Python, MATLAB, C/C++, and C# software programs. In BCI studies, interfaces are designed according to the study's primary purpose. In visual stimulus-based BCI studies, visual stimuli are presented through the interface to increase the focus and help people fulfill the target task through the screen to be used. In auditory BCI studies, the interface presents and controls auditory stimuli. In addition, interfaces can include steps such as initiating the experimental procedure, establishing a connection with the recording device for EEG recording, and terminating the experimental procedure according to the designers' wishes. Researchers and scientists can design new interfaces using programming languages, but this is impossible for people without sufficient programming knowledge. Software for BCI studies is sometimes developed by the same companies that developed the hardware used and are also produced as software products for commercial purposes. Some software is made available on free platforms that are accessible to researchers and end-users without any commercial interest. Most of this software is available as open source. BCI2000, OpenViBE, TOBI Common Implementation Platform (CIP), BCILAB, MetaBCI, and BCI++ are examples of such software [61].

MetaBCI is an open-source BCI software written in Python that facilitates the development of a BCI system. With Meta BCI, stimulus presentation (Brainstim), data upload and processing (Brainda), and online information flow (Brainflow) can be performed [62]. The Brainstim section contains the main screen and it is through this section that the stimulus is presented to the subjects. In the Brainflow section, communication with the EEG recorder and online prediction steps are performed. With the Brainda section, signal processing steps such as data loading, signal preprocessing, and feature extraction can be performed. BCI2000 is a free software package for non-commercial use [63]. BCI2000 is often used for data access, incentive donation, and brain monitoring. It has been in development since 2000 as part of a research and development project and was first used In 2001. It supports various data access systems, brain signaling, and study/feedback paradigms. During operation, BCI2000 stores data in a standard format (BCI2000 native or GDF) with all valid event tags and information about the system configuration. BCI2000 consists of four modules: operator, source, signal processing, and application. The operator module defines the initialization and configuration. The signal processing module performs feature extraction and feature translation of EEG signals. The application module uses the signals translated into control signals to run an application.

OpenViBE is a free and open-source platform for designing and tuning BCI systems. OpenViBE allows connecting to real and virtual environments [64]. BCILAB and EEGLAB are open-source MATLAB-based toolboxes that help design BCIs. BCILAB includes a graphical user interface (GUI) and a scripting interface [65]. The graphical user interface of BCILAB consists of panels, figures for model visualization, model configuration dialog, evaluation setup dialog, and script editor. EEGLAB is a MATLAB toolbox that can process EEG, MEG, and other electrophysiological data. It allows to perform preprocessing and feature extraction analyses such as noise removal, epoch adjustment, time-frequency analysis, and independent component analysis (ICA) [66].

In the BCI design phase, the platforms listed above provide the environment for creating the interface and performing pre-processing steps. However, these platforms can limit researchers' ability to develop a unique new design. Using these programs with visual support for auditory BCI designs may be useful. Researchers who want to create the whole design as a new study-specific interface can use GUIs and toolboxes to realize their designs.

3.2 EEG Signal Recording

Various devices record EEG signals from humans: Brain Product products, BioSemi Active products, Emotiv EPOC headsets, NeuroSky Body, Neuroscan and Mind [67]. These products have been developed to facilitate the recording of EEG signals according to the 10-20 pattern. As shown in Figure 4, some electrode placement patterns are like a silicone mesh, while others are in the form of a cap that is put on the head. EEG caps are produced in different sizes to be fully compatible with people's heads. Auditory stimuli generate AEP in the occipital, temporal, and parietal lobes of the brain due to auditory stimuli [68]. For this reason, the electrodes used can be placed in different numbers in these head regions. Several electrodes were used in the studies, ranging from 1 to 256 [23], [55], [69].

Figure 4. Examples of EEG Caps from Neuroscan [70] and Brain Product [71]. Sourced from Their Official Websites.

The electrodes record EEG signals according to the internationally accepted standard 10-20 electrode layout shown in Figure 5. In the layout, letters represent the brain lobe where the electrode is located. Odd numbers indicate the left side, and even numbers indicate the right side. The letter z represents the center. EEG recording can be performed using several channels ranging from 1-256. Electrodes are placed on the scalp for EEG measurement. A gel is applied between the scalp and the electrode to increase conductivity and reduce noise while taking measurements from the scalp. The obtained EEG signals are transferred to computers, and BCI studies are performed.

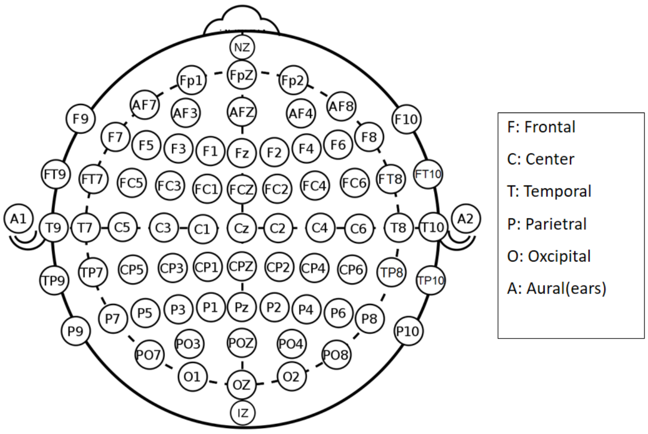


Figure 5. EEG electrode placement according to the 10-20 system [72].

3.3 Signal Processing (Preprocessing, Feature Extraction, and Translation)

The signal processing phase for BCI studies consists of two primary stages: feature extraction and translation [73]. Signal processing is an essential and necessary step for BCI studies to take the signals generated from the brain and convert them into output commands. For the received signals to be recorded and displayed on the computer, the signals must go through a series of processes. Before extracting features from signals and transforming them, steps called pre-processing must be performed [74]. In the preprocessing stage, sampling, amplification, and segmentation processes are applied [75]. These operations aim to make the raw data suitable for analysis. Sampling is the process of converting the analog EEG signal into a digital format [76]. The analog signal is sampled at certain intervals, and each sampling point is transformed into a digital value representing the magnitude of the current signal. The sampling frequency determines the quality of the digital representation of the signal. A higher sampling frequency provides a more accurate digital representation. Amplification is a process that increases the power of the signal [77]. Since the signal is usually measured in a way that is sensitive to noise and EEG signals are low amplitude signals, amplification is necessary to increase the signal-to-attenuation ratio. Segmentation means dividing or separating a long signal into parts based on specific criteria [78]. This process allows the division of EEG data into more manageable parts for analysis. These segments can be of different lengths, ranging from ms to minutes. Further, the signals are filtered using a band-pass filter and a 50/60 Hz Notch filter to clean the signals from ambient and network-related noise [79].

To classify the signals using machine learning and statistical methods, the features of the signal need to be determined. Spectral transform methods, statistical techniques, and filters are used for feature detection [80]. The features given in Table 1 can be defined for EEG signals. During the signal processing stage, researchers examine both the time and frequency domain properties of signals. Short-time Fourier transform and wavelet transform are commonly used methods for time-frequency analysis of EEG signals [81].

During the signal translation stage, the selected features are converted into commands in accordance with the purpose of the designed BBA [82]. For this purpose, signal translation is performed for BCIs designed with auditory stimuli to be applied in areas such as word spelling, wheelchair control, and home device control.

Table 1. Characteristics of the EEG Signal obtained in different domains.

Time Domain	RMS, Peak to Peak, Auto Correlation, Integral of Absolute Value, Zero Crossing and Amplitude distance, Histogram, Collision, Kurtosis, etc.
Frequency Domain	Fast Fourier Transform, Discrete Fourier Transform, etc.
Time-Frequency Domain	Short-time Fourier Transform, Morlet, Wavelet Transform, Wavelet Packet Decomposition, Wavelet Filter Bank etc.
Statistical Features	Averages, Standard Deviation, General characteristics, Minimum or Maximum Value, Normal Distribution, etc.

3.4 Classification

After the features are identified, a classification algorithm is used to assign brain signals to specific classes. These algorithms perform the classification process by analyzing the characteristics of brain signals and the probability of belonging to a particular class. Statistical methods, machine learning, and deep learning studies are used to classify BCI studies [18]. Machine learning is a subdivision of artificial intelligence that acquires knowledge from data and employs this information to make informed decisions [83]. Essentially, machine learning algorithms analyze data sets to execute a task, identify patterns, and make predictions by acquiring knowledge from said patterns. Machine learning works on models developed through experimentation and repetition [84]. Machine learning is generally divided into three main categories: supervised, unsupervised, and reinforcement learning [85]. Machine learning algorithms such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Decision Trees (DT), and Random Forest (RF) are used as supervised learning algorithms in BCI studies [86]. Deep learning is a subband of artificial intelligence that uses multilayer neural networks[87]. These networks have many layers and numerous neurons in each layer. Deep learning performs high-level feature extraction on large and complex data sets. Unlike previous machine learning methods, deep learning provides feature extraction and learning capability within itself [88]. It can automatically recognize and classify complex patterns in data sets. Convolutional Neural Networks (CNN), Long Short Term Memory (LSTM), Multi-Layer Perceptron (MLP), and Recurrent Neural Networks (RNN) are deep learning algorithms used in BCI studies [89]. CNN is a multilayer neural network or deep learning architecture [90]. CNN is suitable for different areas of computer vision, natural language processing, and biomedical studies. Deep learning is often effective in image recognition, natural language processing, speech recognition, and other complex tasks. Deep learning works on big data and computing power, and training processes often take a long time. However, it can deliver highly effective and powerful results when structured correctly. Deep learning methods such as CNN, widely used in image recognition, can be used to detect P300 from scalogram images in BCI tasks involving images [91]. Deep learning methods allow data from brain signals to be converted into images and processed. Khondoker et al. [92] employed deep learning to classify EEG signals that were transformed into scalogram images using continuous wavelet transform. The study highlighted the effectiveness of convolutional neural networks in classifying P300 signals from EEG data.

3.5 Evaluation

Performance metrics such as accuracy, sensitivity, specificity, ROC curve, and information transfer rate (ITR) are used to evaluate BCI's operating performance. Accuracy and ITR are considered two essential criteria for evaluation. Actual and predicted values obtained from the confusion matrix and accuracy, recall, precision, and F1-score values are calculated as given in Table 2. In addition, drawing a ROC chart and calculating the area under the curve is also used to evaluate the classification results using machine and deep learning.

Table 2. Formulas of performance metrics [93]. *TP: a situation that is positive and the model considers positive. FP: the state that is actually positive, but the model considers negative. FN: situations that are actually negative but the model sees as positive. TN: situations that are actually negative and the model views negatively* [94].

CONFUSION MATRIX		Predicted	
		1	0
Actual	1	TP	FN
	0	FP	TN

Performance Metrics
Accuracy $\frac{TP + TN}{TP + TN + FP + FN}$
Recall $\frac{TP}{TP + FN}$
Precision $\frac{TP}{TP + FP}$
F1 − Score $\frac{2 * TP}{2 * TP + FP + FN}$

ITR, widely used to evaluate BCI performance, indicates the amount of information transmitted from the brain to external devices by BCIs over a certain period and is calculated as given in (1) and (2). For this purpose, BCI researches have focused on increasing ITR by increasing accuracy, increasing the number of target classes, and reducing processing time [76].

$$\mathrm{ITR} = \mathrm{R} * \frac{60}{\mathrm{T}} \quad (1)$$

$$\mathrm{R} = \log_2\mathrm{N} + \mathrm{P} * \log_2\mathrm{P} + (1 - \mathrm{P}) * \log_2\left(\frac{1 - \mathrm{P}}{\mathrm{N} - 1}\right) \quad (2)$$

Where R is bits/election and ITR is bits/minute. N is the number of classes, and P is the classifier accuracy.

4 Studies of Auditory BCI

In this section of the study, studies using auditory evoked potentials are summarized. Table 3 provides information about the systems, feature extraction methods, programs, and classification algorithms used for EEG recording in BCI studies. In the study given in the table, Wang et al. [47] developed a BCI to detect a sound target in a noisy environment. They used a Neuroscan EEG cap to perform their study using ERP obtained from 64-channel EEG signals. Using Python and Matlab, they used Short-Time Fourier Transform as the feature extraction method and SVM as the classification algorithm. Borirakarawin and Punsawad [68] designed an ERP-based BCI to improve communication in visual-impaired individuals. Using 8-channel EEG, they recorded the signals and analyzed them using MATLAB program. They used 12 healthy individuals in the study. They used letter and number sounds from loudspeakers as auditory stimuli and investigated the detection of the target sound. They examined the effect of using single and multiple speakers for auditory stimuli. In the study, vowel letter stimuli showed higher accuracy than two-digit stimuli. In addition, an accuracy higher than 85% of the average accuracy was obtained when multiple speakers were used.

Guo et al. [95] designed a BCI with three auditory stimuli and classified it with a support vector machine. Yes/no binary classification was performed using the oddball paradigm in the study. EEG recordings were taken from 10 people, and the study was conducted using N200 signals. They investigated the effect of tone frequency on the discrimination of auditory stimuli and found a classification accuracy of 87.41 and an ITR of 6.48 bits/min. Kimura et al. [96] designed a BCI to develop an automatic voice control system. Their study used an auditory oddball paradigm by performing two different experiments. In the first experiment, ten subjects without hearing problems were exposed to two sounds at various levels, 60 dB and 70 dB, and were asked to pay attention to the louder sound. They analyzed the P300 potential from the EEG signals recorded with a 64-channel EEG recorder. As a result of the classification using support vector machines, they achieved 90% accuracy in the first experiment. The second experiment used an auditory stimulus at 50 dB as the target sound and 70 dB as the non-target sound. As a result of the study, they achieved a 76% success rate. As a result of the study, it was observed that the higher sound level showed higher success than the lower sound level.

Ogino et al. [97] used two open-access datasets and four different machine learning algorithms combined with semi-supervised learning to compare the performance of auditory ERP-based speller BCIs with machine learning algorithms. The best performances were 96.16% accuracy and 7.15 bit/min ITR. Borirakarawin and Punsawad [98] suggested a hybrid BCI system combining auditory stimulus and speech imagery paradigms in real time to control EEG devices. The study reported an average accuracy rate of 83.3%. Considering the publications in this review, the performance metrics of studies conducted with only auditory stimuli are lower than those of hybrid stimuli. Christodoulides et al. [36] conducted a study using auditory, visual, and hybrid stimuli to classify dyslexic and non-dyslexic university students. They recorded EEG signals using the Emotiv EPOC+ device and headset. They used MATLAB for feature extraction and classification. They used FFT transform, energy, and Shannon entropy for feature extraction. They used the Random Forest algorithm, a machine learning algorithm for the classification algorithm. As a result of the study, they achieved over 95% accuracy for three different conditions. The hybrid stimulus case performed best in the right and left hemispheres.

When the publications given in Table 3 are examined, it is seen that different types of devices ranging from 2-128 channels can be used to record EEG signals. Although EEG containers can have different features, the ones with forms suitable for 10-20 electrode placement are preferred. The caps are placed in a mesh or silicone structure to facilitate the placement of the electrodes on the scalp.

In the classification of EEG signals, studies are carried out using different programs and algorithms. Machine learning algorithms such as SVM are among the frequently preferred algorithms. In addition to these algorithms, ANOVA is another preferred method in BCI studies as given in the Table 3.

Table 3. Overview of EEG Recording Systems, Methodologies, and Software Programs Used in BCI Studies.

Author	EEG Recording Device	Software and Stimulus transmitted	Feature Extraction Method	Program	Classification Algorithm
[47]	Neuroscan EEG cap	Neuroscan acquisition device via PsychoPy	Short-Time Fourier Transform	MATLAB	Support Vector Machine(SVM)
[68]	OpenBCI with an eight-channel EEG amplifier	OpenBCI	Stepwise linear discriminant analysis (SWLDA)	MATLAB, EEGLAB toolbox	Linear regression
[95]	Compumedics Neuroscan, USA	SynAmps	Empirical Mode Decomposition (EMD)	-	SVM, paired t-test
[96]	ActiveTwo; Biosemi, Amsterdam, Netherlands)	-	averaged the time-series	Matlab, Python and MNE-Python	SVM, Wilcoxon signed-rank test
[97]	Brain Products amplifier (Brain Products Co., Munich, Germany) and Fast'n Easy Cap (Easy- Cap GmbH, Inning, Germany)	-	baseline correcting and flattening for the temporal features	MATLAB	step-wise LDA (SWLDA), spatial-temporal discriminant analysis (STDA), Semi-supervised learning (SSL), least squares SVM (LS-SVM), SEMI-SUPERVISED LS-SVM, two-way analysis of variance (ANOVA)
[98]	Cyton board and EEG cap	Cyton board from OpenBCI Company, digital sound level meter (TETSL01)	power spectral density (PSD)	Python, BrainFlow	Proposed algorithm
[23]	EEProbe recording software (ANT, Enschede, The Netherlands)	Eevoke Software for audiovisual presentation	ANOVA/ statistical analysis	ANOVA	ANOVA
[36]	Emotiv EPOC +	-	FFT transformation, Energy, and Shannon entropy	MATLAB and EEGLAB toolbox	Random Forest

The electrodes used are usually Ag/AgCl [99]. The containers used can be produced in different sizes to be compatible with the head structure of the individuals. The use of EEG devices and channels presented in Table 3 directly affects the accuracy rates. Significantly, 64-channel EEG systems increase classification accuracy by increasing the signal's sensitivity.

Table 4 summarizes BCI studies with auditory stimuli conducted in different fields. These studies were investigated in two areas: BCI studies using only auditory stimuli and hybrid BCI studies. The studies in the literature were evaluated in terms of the stimulus used, the method used, and the findings obtained. Proverbio et al. [23] used ERP signals to investigate the psychophysiological markers of imagery processes. In the study with 30 healthy participants, visual and auditory stimuli representing ten different semantic categories were presented. In three categories, 40 speech items, 40 music items, and 40 emotional vocalization items were used as auditory stimuli. In the absence of sensory stimulation, unprecedented electrophysiological signs of imagination were recorded. Peaks were recorded from the scalp during infants' imagination of human faces, animals, music, speech, emotional vocalizations, and sensory (visual and auditory) modalities. Using ANOVA, the results showed that the category-dependent modulation of ERPs during imagery was longer delayed and more anterior than in the perceptual condition. These results suggest that ERP markers may be helpful for BCI systems to be developed for patients affected by disorders of consciousness or for locked patients such as SLA. Halder et al. [100] conducted their study using EEG and fMRI data for auditory BCI studies. Using a 16-channel EEG device, they studied from the EEG signal they recorded. In their study, they designed an auditory speller using five different animal sounds. They used the t-test for signal analysis. They conducted their study in three phases and found that the average online accuracy was 60% in session t2, 73% in session t3, and 65% in session t4. In their brain region imaging study based on the relationship between motor imagery and auditory BCI, it was observed that there was an activation overlap between the active brain regions.

Wang et al. [47] proposed a BCI for audio target detection. The sounds of three different unmanned aerial vehicles were used as target sounds. They presented the recorded sounds to the subjects using in-ear headphones. In the experiment with 8 participants, they recorded the signals using a 64-channel EEG device. Through PsychoPy, a Python library, the target sound moment was transmitted to the Neuroscan recorder via parallel communication. They analyzed the ERP and ERSP signals using SVM and found that the target detection rate was 84.84%, and the detection time was 0.7758 s.

The research conducted by Sun et al. [101] focused on target sound selection and involved the development of an auditory BCI that utilized short stories as stimuli. They obtained P300 and N200 potentials by counting how often the target word was repeated in short stories. They received 78.96% accuracy in their study. Choi et al. [52] used P300-based BCI to control visual stimuli, natural sound and artificial auditory stimuli, and an electric lamp in real-time. In conclusion, natural sounds led to higher online BCI performance and more significant differences in ERP amplitudes between targets and nontargets than artificial sounds. Visual stimuli performed better (77.56% on average) than their auditory counterparts (54.67 % on average). Markovinović et al. [34] designed an auditory stimulus speller BCI and achieved an average spelling accuracy of 30% and ITR of 2.38 bit/min. A hybrid BCI, proposed by Barbosa et al. [102], utilizes both visual and auditory stimulation to detect the P300 signal in their study. The hybrid approach has a significantly higher average online accuracy of 85.3% compared to the visual and audio techniques, which only achieve 53.3%. The highest visual stimulus had an ITR of 5.96 bits per minute. Borirakarawin and Punsawad [98] suggested a hybrid BCI system combining auditory stimulus and speech imagery paradigms to control EEG devices in real-time. The study reported an average accuracy rate of 83.3%. Considering the publications in this review, the performance metrics of studies conducted with only auditory stimuli are lower than those of hybrid stimuli.

With BCI design for device and object control, studies on television, air conditioners, and lamp control are carried out in home applications. Velasco-Álvarez et al. [54] designed a spelling BCI using P300 potential to control televisions, air conditioners, smart bulbs, smart sockets, and WhatsApp and Spotify applications with voice commands. They classified using stepwise linear discriminant analysis (SWLDA); the accuracy was 80.68%, and the ITR was 25.9 bits/min. Shivappa et al. [53] designed BCI with auditory stimulation to develop a home automation system based on ASSR. In the study, fans and bulbs were checked. For the smart bulb, four states were tested: on, off, 25% illumination, and 25% dimming, and the smart socket was tested for on and off states. In the smart bulb, the average system response time for on, off, bright, and dim states was 17, 20.4, 16, and 17.6 seconds, respectively, and the accuracy was 92%, 50%, 59%, and 67%, respectively. The average system response time for the smart plug was 31 and 22 seconds for the on and off states, respectively, and 100% and 92% accuracy was achieved, respectively. Hybrid studies are being conducted to improve performance metrics in brain-computer interfaces. A combined visual stimulus BCI was proposed by Edlinger et al. [51] to control a smart home environment, incorporating P300 and SSVEP. Their study yielded a remarkable 100% accuracy rate.

Ogino et al. [103] proposed a portable auditory BCI system with two-channel (one reference) EEG signals recorded from the anterior frontal region. Their study used natural sounds consisting of five different animal sounds as auditory stimuli. They performed their study using ERP from the EEG signals recorded from a subject group consisting of one ALS patient and nine healthy individuals. In their study, they performed the experimental procedure with the oddball paradigm. They performed their studies in two steps, online and offline. They selected and classified features using SWLDA. In their two-class study as target and non-target, the average accuracy obtained from cross-validation in offline analysis was 70% and 80% in online analysis. The ITR value was 1.29 for offline analysis and 1.16 for online analysis.

Halder et al. [55] designed a speller with auditory stimuli. Their study used 5 subjects, and they used five different animal sounds as stimuli in their study. Using a 16-channel g.USBamp amplifier, they recorded EEG signals and analyzed them from the P300 potential. As a result of the classification they performed using SWLDA, they obtained 92% accuracy and found the ITR to be 5.78 Bit/min. Borirakarawin et al. [68] designed a BCI with ERP signals for target sound detection. As auditory stimuli, they played different auditory stimuli, including one and multiple speakers, to the subjects through loudspeakers. They used SWLDA for feature extraction and prediction. As a result of the study, they obtained an accuracy of 86% for single speakers and 87.2% for multiple speakers. Jijomon et al. [104] proposed a BBA to develop a person identification system by making people listen to familiar names. Their study with six subjects analyzed AEP signals recorded with a 64-channel EEG device. They performed classifications using the CNN algorithm and achieved 99% accuracy as a result of the study.

Table 4. Detailed Summary of Auditory BCI Studies Including Stimuli, Methods, and Results.										
Author	Year	Subject	Person	Stimulus	Repetition	Time	Signal	Channel	Acc	ITR
[55]	2016	Speller with auditory stimulus	5	Five different animal sounds	10	150 ms stimulus+287.5 ms ISI. 2s pause between rows and columns, 12 s pause between letters	P300	16	92%	5.78
[34]	2022	Speller with auditory stimulus	10	Letter sounds	2	22 letter sounds at 1s intervals	P300	16	89.5%	9.2
[52]	2022	Lamp control	30	Three types of Auditory stimuli	6-8	Each stimulus is 275 ms. Four different voices for four missions. Three different stimulus types: animal, word, and beep. 250ms wait. 30 blocks of training, 15 blocks of testing	P300	31	54.67%	4.47
[100]	2019	brain region imaging with motor imagery and auditory.	8	Five different animal sounds	10	Animal stimulation duration is 150 ms and inter-stimulus interval of 287.5 ms. 6000 and 2000 ms sn cue	EEG and FMRI	16	73%	-
[68]	2022	detection of the target sound.	12	single and multiple speakers	30	100 ms and 250 ms auditory stream stimulus and 300 ms ISI. Each trial consists of four random stimulus sounds.	ERP	8	85%	-
[95]	2015	discrimination of auditory stimuli	10	Four different auditory stimuli: three pure tones at 100 Hz, 1000 Hz, 4000 Hz, and white noise.	20	Simple yes/no question. Each stimulus lasted 100 ms and ISI 400-800 ms.	N200	32	87.41%	6.48
[96]	2024	automatic voice control system	10	50 dB, 60 dB, and 70 dB sound	20	100 ms auditory stimulus, 900 ms rest.	P300	64	90%	1,46
[23]	2023	investigate the psychophysiological markers of imagery processes	30	40 speech items, 40 music items, and 40 emotional vocalization	-	1500 ms stimulus, 500 ms ISI. Auditory stimulus 2000 ms perceptual and imagery condition. Inter+trial interval 900 ms.	ERP	128	-	-
[47]	2023	detection of the target sound	8	unmanned aerial vehicles sound	10	5 s preparation. 30 s wing noise and target sound.	ERP	64	84.84%	-

[101]	2023	Device control	24	Six women, six men, 12 short stories	20	45-60 s story. Seven target words in each story. 5-8 s between each target word.	P300 N200	32	78.96%	0.21
[104]	2021	EEG-based biometric identificatio n	20	Name-saying auditory stimulus	4	Twenty-nine different names, 10 of which are tatak. 90 seconds for each trial	AEP	64	99%	-
[102]	2016	word selection	10	Word photo and sound	15	450 ms stimulus, 100 ms pause for each word. 8 s wait between each trial	P300	16	85.3%	5.92
[98]	2023	Auditory stimulus attention test	6	4400 Hz 80 dB, three different female voices.	10	432 ms stimulus + 500 ms pause seven stimuli (1 target and six non-targets) in each trial	P300	8	74.4%	-
[105]	2021	Music attention selectivity	9	Two types of music Three different musical instrume nts	28	1 s pause, then 8 s stimulus twice. 1. stimulus is standard, 2. stimulus is changing. 4 pieces of each stimulus	EEG	11	71.23%	1.01
[103]	2019	Portable auditory BCI	10	Five animal sounds (duck, singing bird, frog, seagull, and dove)	30	150 ms sound and 150 ms ISI	ERP	2	70%	1,16
[54]	2022	Device control with Speller	15	Letter on a screen and sound. Voice command for control device	12	192 ms stimulus, 32 ms ISI. One sequence 14 stimulus.	P300	8	80.68%,	25.9
[53]	2018	Device control	4	Voice command for the control of device	3	37 Hz and 43 Hz wireless stereo speakers. 30 s stimulus	ASSR	8	92%	31

An et al. [105] designed a BCI using EEG signals to investigate the selective attention of music listeners to musical instruments. In the experiment with 9 participants, stimuli were presented to the subjects using an oddball paradigm. Three different musical instruments were played from three different directions: left, right, and center. Before the auditory stimulus, a visual cue was presented on a screen before the subjects. The features obtained by the auditory attention decoding (AAD) method were classified using a linear support vector machine. The study achieved 71.23% accuracy and 1.01 bit/min ITR.

The performance metrics presented in Table 4 show the overall success of auditory stimulus BCI systems. While studies based on P300 signals have shown high accuracy rates, researchers have noted that auditory stimulus systems can reach an accuracy of up to 90%. The reviewed studies reveal different degrees of success in BCI implementations using auditory stimuli. In general, auditory BCIs using only auditory stimuli show promising findings, but their performance metrics are generally lower compared to hybrid BCIs. Hybrid BCIs exhibit higher accuracy and reliability by combining auditory and visual stimuli. When looking at BCI studies using auditory stimuli only, [23] Investigated psychophysiological markers using ERP signals with 30 participants. The study used various auditory stimuli, such as speech, music, and emotional vocalizations. This study found unique electrophysiological markers during imagination processes and suggests potential BCI applications for patients with impaired consciousness. [100] Designed an auditory printer with animal sounds using EEG and fMRI and achieved varying accuracy rates (60-73%) in different sessions. Their study highlighted the overlap between motor imagery and auditory BCI active brain regions, indicating that complex interaction may be beneficial for designing more effective BCIs. [99]Achieved 84.84% detection rate using drone sounds for target sound detection. Analysis of ERP and ERSP signals with SVM demonstrated the effectiveness of specific auditory stimuli in improving BCI performance. [101] developed an auditory BCI that acquires P300 and N200 potentials using short stories and achieved 78.96% accuracy. This approach demonstrated the usability of complex auditory stimuli for BCI target detection. [52] Designed a P300-based BCI to control devices using natural and artificial sounds and found that natural sounds provided higher performance (54.67%) than artificial sounds. Visual stimuli outperformed auditory stimuli, indicating areas for improvement in auditory BCI design. [34] designed an auditory printer with 30% accuracy and 2.38 bit/min ITR. This study highlights the challenges of accurate typing tasks using auditory stimuli. [98] They achieved 83.3% accuracy by combining auditory stimuli and speech images. This study showed that including multiple sensory modalities can improve BCI performance.

In hybrid BCI studies, [102] detected P300 signals using visual and auditory stimuli and obtained an average accuracy rate of 85.3%. This hybrid approach reveals that combining sensory modalities is helpful in improving BCI accuracy. [54] Developed a P300-based printer for home automation and achieved 80.68% accuracy and 25.9 bit/min ITR. This study demonstrates the practical application of BCIs in smart home environments. [53] Designed an ASSR-based BCI for home automation and achieved high accuracy in controlling devices such as smart light bulbs and sockets. This work demonstrated the potential of BCIs to facilitate complex home automation tasks with high reliability. [51] Proposed a BCI combining P300 and SSVEP and achieved 100% accuracy rate in smart home control. This study provides reliable results for using hybrid BCIs in practical applications.

5 Conclusion

This review examines the structure of auditory BCI systems and the studies conducted in this field. Auditory BCIs are important in facilitating communication and device control for individuals with visual and tactile sensory loss and those with motor impairments. This review illuminates the various methodologies and technologies used to design and develop auditory BCIs by analyzing the studies and developed systems in the literature. EEG stands out in its non-invasive nature, ease of use, portability, and cost-effectiveness. In this study, we focus on BCI studies developed with EEG signals. In auditory BCI studies, ERPs and AEPs generated as a result of a response to an auditory stimulus are of critical importance. The properties and usage areas of these potentials are analyzed in this study. The general steps of BCI design, such as interface design, signal recording, signal processing, signal conversion, classification, and evaluation, are explained. While platforms such as BCI2000, MetaBCI, and OpenVIBE can be used as interface development tools, people with programming knowledge can also use software languages such as Python, MATLAB, and C++. Recently, machine learning and deep learning algorithms have been used to classify the signals, and their effects on the performance of the study have been examined. In addition, a comparison between hybrid BCI studies and auditory BCI studies is also made in this study. Although there have been studies with high accuracies in auditory BCIs, hybrid BCIs generally have higher accuracy rates. Achieving high information transfer rates without too many repetitions is an important point emphasized in the studies.

Future studies in these areas will increase the effectiveness of auditory BCIs. It is important to continue studies in this field and to develop auditory BCI systems that reach end users for individuals who only have the ability to hear and can communicate with their environment in this way. Future studies could focus on improving the effectiveness of hybrid BCI systems. In particular, further research could be conducted on the effect of combining visual and tactile stimuli with auditory stimuli. It is worth noting that hybrid systems show improved accuracy and speed in information transfer compared to auditory-only based systems. Further research is needed to examine the effect of auditory stimulus complexity on BCI performance. An example is the comparison between the cognitive processing required for complex auditory stimuli, such as short stories or natural sounds, and the processing required for simple tones or animal sounds and their impact on performance. Developing new methods for more precise recording and processing of EEG signals could improve the accuracy of auditory BCIs. In particular, comparing existing signal processing algorithms and evaluating their performance may be advantageous.

6 Author contribution statement

Author 1 created the article's first draft, conceptualized the ideas, and contributed to the publication. He put a lot of effort into writing the publication by conducting a literature review and analyzing the publications in detail. He also took an active role in the visualization process and integrated visual elements to make the text more understandable.

Author 2 contributed significantly to the revision and editing of the manuscript. At the same time, he undertook supervision and project management tasks to ensure the integrity of the article and manage the publication effectively.

7 Ethics committee approval and conflict of interest declaration

"There is no need to obtain ethics committee permission for the article prepared."

“There is no conflict of interest with any person/institution in the article prepared.”

8 References

[1] Sengar P, Bardhan S. *Brain–Computer Interface*. Jena O P, Bhushan B, Kose U Machine Learning and Deep

Learning in Medical Data Analytics and Healthcare Applications, 23–38, Boca Raton, Florida, USA, CRC Press, 2022.

[2] Gao S, Wang Y, Gao X. "Visual and Auditory Brain – Computer Interfaces". *IEEE Transaction of Biomedical Eng*ineering, 61(5),1436–47, 2014.

[3] Mori F, Sugino M, Huang Y, Kotani K, Jimbo Y. "Control of Electric Wheelchair by Brain-Computer Interface Using Mixed Reality and Virtual Sound Source". *IEEJ Transactions on Electrical and Electronic Engineering*, 2024.

[4] Allison BZ, Kübler A, Jin J. "30+ years of P300 brain–computer interfaces". *Psychophysiology*, 57(7),1–18, 2020.

[5] Saha S, Mamun KA, Ahmed K, Mostafa R, Naik GR, Darvishi S, et al. "Progress in Brain Computer Interface, Challenges and Opportunities". *Frontiers in Systems Neuroscience,* 15(February),1–20, 2021.

[6] Miltiadous A, Aspiotis V, Peschos D, Tzimourta KD, Abosaleh AHS, Giannakeas N, et al. "An Ensemble Method for EEG-based Texture Discrimination during Open Eyes Active Touch". Engineering, Technology & Applied Science Research, 14(1),12676–87, 2024.

[7] Liu Z, Shore J, Wang M, Yuan F, Buss A, Zhao X. "A systematic review on hybrid EEG/fNIRS in brain-computer interface". *BiomedicalSignal Process Control,* 68(March),102595, 2021. Available from, https,//doi.org/10.1016/j.bspc.2021.102595

[8] Moon H, Kwon J, Eun J, Chung CK, Kim JS, Chou N, et al. "Electrocorticogram (ECoG), Engineering Approaches and Clinical Challenges for Translational Medicine". *Advanced Materials Technologies,* 2301692,1–18, 2024.

[9] Liu C. "Application of ECoG and Electrode in BCI". *Theoretical and Natural Science,* 3(1),410–6, 2023.

[10] Kubska ZR, Kamiński J. "How human single-neuron recordings can help us understand cognition, Insights from memory studies". *Brain Science*, 11(4), 2021.

[11] I. U. "Analysis and classification of EEG signals for brain computer interfaces". *Frontiers in Human Neuroscience*, Vol. 2, 2008.

[12] Ioannides A. *Magnetoencephalography (MEG).*Editor: Hyder F. Dynamic Brain Imaging. Methods In Molecular Bıology, 167–188, Yale University, New Haven, CT, USA, Humana Press, 2009. https://doi.org/10.1007/978-1-59745-543-5_8.

[13] Buxton RB. "The physics of functional magnetic resonance imaging (fMRI)". *Reports Prog Phys*, 76(9), 2013.

[14] Cox DD, Savoy RL. "Functional magnetic resonance imaging (fMRI) "brain reading", Detecting and classifying distributed patterns of fMRI activity in human visual cortex". *Neuroimage*, 19(2),261–70, 2003.

[15] Alavi A, Huang SS. "Positron Emission Tomography in Medicine, An Overview". *Cancer Imaging*, 31(1),39–44, 2007.

[16] Zhu Y, Xu K, Xu C, Zhang J, Ji J, Zheng X, et al. "PET mapping for brain-computer interface stimulation of the ventroposterior medial nucleus of the thalamus in rats with implanted electrodes". *Journal of Nuclear Medicine*, 57(7),1141–5, 2016.

[17] Padilla-Buritica JI, Hurtado J V., Castellanos-Dominguez G. "Supervised piecewise network connectivity analysis for enhanced confidence of auditory oddball tasks". *BiomedicalSignal Process Contro*l, 52,341–6, 2019. Available from, https,//doi.org/10.1016/j.bspc.2019.04.020

[18] Alsuwaiket MA. "Feature Extraction of EEG Signals for Seizure Detection Using Machine Learning Algorthims". *Engineering, Technology and Applied Science Research,* 12(5),9247–51, 2022.

[19] Fernández-Sotos P, García-Martínez B, Ricarte JJ, Latorre JM, Sánchez-Morla EM, Fernández-Caballero A, et al. "Electroencephalographic spectral analysis from a wireless low-cost brain-computer interface for symptom capture of auditory verbal hallucinations in schizophrenia". *Schizophrenia Research,* 220,297–9, 2020. Available from, https,//doi.org/10.1016/j.schres.2020.04.011

[20] Bera TK. "Noninvasive electromagnetic methods for brain monitoring, A technical review"., *Intelligent Systems Reference Library*. Vol. 7451–95 p, 2015.

[21] Belwafi K, Ghaffari F, Djemal R, Romain O. "A Hardware/Software Prototype of EEG-based BCI System for Home Device Control". *Journal of Signal Processing Systems*, 89(2),263–79, 2017.

[22] Zhuang T, Zhao H, Tang Z. "A Study of Brainwave Entrainment Based on EEG Brain Dynamics". *Computer and Information Science*, 2(2),80–6, 2009.

[23] Proverbio AM, Tacchini M, Jiang K. "What do you have in mind? ERP markers of visual and auditory imagery". *Brain Cognitive*, 166(January),105954, 2023. Available from, https,//doi.org/10.1016/j.bandc.2023.105954

[24] Mudgal SK, Sharma SK, Chaturvedi J, Sharma A. "Brain computer interface advancement in neurosciences, Applications and issues". *Interdisciplinary Neurosurgery: Advanced Techniques and Case Management,* 20(April),100694. 2020. Available from, https,//doi.org/10.1016/j.inat.2020.100694

[25] Hinterberger T, Veit R, Wilhelm B, Weiskopf N, Vatine JJ, Birbaumer N. "Neuronal mechanisms underlying control of a brain-computer interface". *European Journal of Neuroscience*, 21(11),3169–81, 2005.

[26] Bilal Aygün A, Reşit Kavsaoğlu A. "An innovative P300 speller brain–computer interface design, Easy screen". *BiomedicalSignal Process Control.*, 75(February), 2022.

[27] Xiao W, Manyi G, Khaleghi A. "Deficits in auditory and visual steady-state responses in adolescents with bipolar disorder". *Journal of Psychiatric Research*, 151(April),368–76, 2022. Available from, https,//doi.org/10.1016/j.jpsychires.2022.04.041

[28] Müller-Putz GR, Scherer R, Neuper C, Pfurtscheller G. "Steady-state somatosensory evoked potentials, Suitable brain signals for brain-computer interfaces". *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(1),30–7, 2006.

[29] Ergün E, Aydemir Ö. "Improving Classification

Accuracy of Motor Imagery EEG Signals via Effective Epochs". *Pamukkale University Journal of Engineering Sciences,* 24(5), 817–23, 2018.

[30] Nijboer F, Furdea A, Gunst I, Mellinger J, McFarland DJ, Birbaumer N, et al. "An auditory brain-computer interface (BCI)". *Journal of Neuroscience Methods*, 167(1),43–50, 2008.

[31] Paulraj MP, Subramaniam K, Yaccob S Bin, Adom AH Bin, Hema CR. "Auditory Evoked Potential Response and Hearing Loss, A Review". *The Open Biomedical Engineering Journal*, 9(1),17–24, 2015.

[32] Bohórquez J, Özdamar Ö. "Generation of the 40-Hz auditory steady-state response (ASSR) explained using convolution". *Clinical Neurophysiology,* 119(11), 2598–607, 2008. Available from, http,//dx.doi.org/10.1016/j.clinph.2008.08.002

[33] Pérez-González D, Malmierca MS. "Adaptation in the auditory system, An overview". *Frontiers in Integrative Neuroscience*, 8(FEB), 2014.

[34] Markovinović I, Vrankić M, Vlahinić S, Šverko Z. "Design considerations for the auditory brain computer interface speller". *BiomedicalSignal Process Control*, 75(January), 2022.

[35] Baykara E, Ruf CA, Fioravanti C, Käthner I, Simon N, Kleih SC, et al. "Effects of training and motivation on auditory P300 brain-computer interface performance". *Clinical Neurophysiology,*. 127(1),379–87, 2016. Available from, http,//dx.doi.org/10.1016/j.clinph.2015.04.054

[36] Christodoulides P, Miltiadous A, Tzimourta KD, Peschos D, Ntritsos G, Zakopoulou V, et al. "Classification of EEG signals from young adults with dyslexia combining a Brain Computer Interface device and an Interactive Linguistic Software Tool". *Biomedical Signal Process Control*, 76(February),103646, 2022. Available from, https,//doi.org/10.1016/j.bspc.2022.103646

[37] Zhang B, Zhou Z, Jiang J. "A 36-class bimodal erp brain-computer interface using location-congruent auditory-tactile stimuli". Brain Science, 10(8),1–18, 2020.

[38] Klatt LI, Getzmann S, Wascher E, Schneider D. "The contribution of selective spatial attention to sound detection and sound localization, Evidence from event-related potentials and lateralized alpha oscillations". *Biol Psycho*logy, 138(August),133–45, 2018. Available from, https,//doi.org/10.1016/j.biopsycho.2018.08.019

[39] Jijomon CM, Vinod AP. "Person-identification using familiar-name auditory evoked potentials from frontal EEG electrodes". *Biomedical Signal Process Control,* 68(March),102739, 2021. Available from, https,//doi.org/10.1016/j.bspc.2021.102739

[40] Lu Y, Wang M, Yao L, Shen H, Wu W, Zhang Q, et al. "Auditory attention decoding from electroencephalography based on long short-term memory networks". *Biomedical Signal Process Control,* 70(October 2020),102966, 2021. Available from, https,//doi.org/10.1016/j.bspc.2021.102966

[41] Heo J, Baek HJ, Hong S, Chang MH, Lee JS, Park KS. "Music and natural sounds in an auditory steady-state response based brain–computer interface to increase user acceptance". *Comput Biol Medicine,* 84(March),45–52, 2017. Available from, http,//dx.doi.org/10.1016/j.compbiomed.2017.03.011

[42] Kaongoen N, Jo S. "A novel hybrid auditory BCI paradigm combining ASSR and P300". *J Neurosci Methods*, 279,44–51, 2017. Available from, http,//dx.doi.org/10.1016/j.jneumeth.2017.01.011

[43] Onishi A, Takano K, Kawase T, Ora H, Kansaku K. "Affective Stimuli for an Auditory P300 Brain-Computer Interface". *Frontiers in Neuroscience,* 11(September),1–9, 2017.

[44] Iacoviello D, Petracca A, Spezialetti M, Placidi G. "A real-time classification algorithm for EEG-based BCI driven by self-induced emotions". C*omput Methods Programs Biomed*, 122(3),293–303, 2015. Available from, http,//dx.doi.org/10.1016/j.cmpb.2015.08.011

[45] Nambu I, Ebisawa M, Kogure M, Yano S, Hokari H, Wada Y. "Estimating the Intended Sound Direction of the User, Toward an Auditory Brain-Computer Interface Using Out-of-Head Sound Localization". *PLoS One* , 8(2), 2013.

[46] Schreuder M, Blankertz B, Tangermann M. "A new auditory multi-class brain-computer interface paradigm, Spatial hearing as an informative cue". *PLoS One*, 5(4), 2010.

[47] Wang R, Liu Y, Shi J, Peng B, Fei W, Bi L. "Sound Target Detection Under Noisy Environment Using Brain-Computer Interface". *IEEE Trans Neural Syst Rehabil Eng*ineering, 31,229–37, 2023.

[48] Manian V. "Auditory Source Localization by Time Frequency Analysis and Classification of Electroencephalogram Signals". *BiomedicalJ Sci Tech Res*, 19(5),14595–606, 2019.

[49] Cai Z, Makino S, Rutkowski TM. "Brain Evoked Potential Latencies Optimization for Spatial Auditory Brain–Computer Interface". *Cognit Comput*..7(1),34–43, 2015.

[50] Kiroy VN, Bakhtin OM, Krivko EM, Lazurenko DM, Aslanyan E V., Shaposhnikov DG, et al. "Spoken and Inner Speech-related EEG Connectivity in Different Spatial Direction". *Biomedical Signal Process Control*, 71(PB),103224, 2022. Available from, https,//doi.org/10.1016/j.bspc.2021.103224

[51] Edlinger G, Holzner C, Guger C. "A hybrid brain-computer interface for smart home control". *Lect Notes Computer Science,* 6762,417–26, 2011.

[52] Choi YJ, Kwon OS, Kim SP. "Design of auditory P300-based brain-computer interfaces with a single auditory channel and no visual support". *Cognitive Neurodynamics,* 3, 2022. Available from, https,//doi.org/10.1007/s11571-022-09901-3

[53] Shivappa VKK, Luu B, Solis M, George K. "Home automation system using brain computer interface paradigm based on auditory selection attention". *2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Houston, TX, USA, 2018. doi: 10.1109/I2MTC.2018.8409863.*.

[54] Velasco-Álvarez F, Fernández-Rodríguez Á, Ron-

Angevin R. "Brain-computer interface (BCI)-generated speech to control domotic devices". *Neurocomputing,* 509,121–36, 2022.

[55] Halder S, Käthner I, Kübler A. "Training leads to increased auditory brain-computer interface performance of end-users with motor impairments". *Clinical Neurophysiology*, 127(2),1288–96, 2016. Available from, http,//dx.doi.org/10.1016/j.clinph.2015.08.007

[56] de Souza AP, Soares QB, Mendes EMAM, Felix LB. "Brain–computer interface in an inter-individual approach using spatial coherence, Identification of better channels and tests repetition using auditory selective attention". *Biomedical Signal Process Control*, 83(January),104573, 2023. Available from, https,//doi.org/10.1016/j.bspc.2023.104573

[57] C.M. J, A.P. V. "Detection and classification of long-latency own-name auditory evoked potential from electroencephalogram". *Biomedical Signal Process Control*, 68(March),102724, 2021. Available from, https,//doi.org/10.1016/j.bspc.2021.102724

[58] Sugi M, Hagimoto Y, Nambu I, Gonzalez A, Takei Y, Yano S, et al. "Improving the performance of an auditory brain-computer interface using virtual sound sources by shortening stimulus onset asynchrony". *Frontiers in Neuroscience*, 12(February),1–11, 2018.

[59] Squires, Nancy K, Squires, Kenneth C, Hillyard, Stephen A. "Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man". *Electroencephalography and Clinical Neurophysiology,* 38(4),387–401, 1975.

[60] Pereira DR, Cardoso S, Ferreira-Santos F, Fernandes C, Cunha-Reis C, Paiva TO, et al. "Effects of inter-stimulus interval (ISI) duration on the N1 and P2 components of the auditory event-related potential". I*nt J Psychophysiol*, 94(3),311–8, 2014. Available from, http,//dx.doi.org/10.1016/j.ijpsycho.2014.09.012

[61] Stegman P, Crawford CS, Andujar M, Nijholt A, Gilbert JE. "Brain-computer interface software, A review and discussion". *IEEE Trans Human-Machine Syst*ems, 50(2),101–15, 2020.

[62] Mei J, Luo R, Xu L, Zhao W, Wen S, Wang K, et al. MetaBCI, "An open-source platform for brain–computer interfaces". *Computers in Biology and Medicine,* 168(November 2023),107806, 2024. Available from, https,//doi.org/10.1016/j.compbiomed.2023.107806

[63] Schalk G, Mcfarland DJ, Hinterberger T, Birbaumer N, Wolpaw JR, "BCI2000 , A General-Purpose Brain-Computer Interface (BCI) System". 1034 *IEEE Transactıons On Bıomedıcal Engıneerıng,* 51(6),1034–43, 2004.

[64] Renard Y, Lotte F, Gibert G, Congedo M, Maby E, Delannoy V, et al. OpenViBE, "An open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments". *Presence Teleoperators Virtual Environ*. 19(1),35–53., 2010.

[65] Kothe CA, Makeig S. "BCILAB, A platform for brain-computer interface development". *Journal of Neural Engineering,* 10(5), 2013.

[66] Delorme A, Makeig S. "EEGLAB, An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis". *Journal of Neuroscience Methods*, 134(1),9–21, 2004.

[67] Erat K, Onay Durdu P. "User experience evaluation of low cost EEG headsets". *Pamukkale University Journal of Engineering Sciences*,.27(5),646–59, 2021

[68] Borirakarawin M, Punsawad Y. "Event-Related Potential-Based Brain–Computer Interface Using the Thai Vowels' and Numerals' Auditory Stimulus Pattern". S*ensors,* 22(15), 2022.

[69] Akira D. "An Information Theoretic Analysis of 256-Channel EEG Recordings: Mutual Information and Measurement Selection Problem," International Conference on Independent Component Analysis and Blind Signal Separation, San Diego, Calif, USA, January 2001.

[70] Compumedics Neuroscan, "Compumedics Neuroscan/ Caps," *Caps*. https://compumedicsneuroscan.com/products/caps/ (21.05.2024).

[71] Brain Products, "Electrodes & caps," *Electrodes & caps*. https://www.brainproducts.com/solutions/#electrodes-caps (21.05.2024).

[72] Klem GH, Luders HO, Jasper HH. "The ten-twenty electrode system of the International Federation". *International Federation of Clinical Neurophysiology*. p. 371–375, 1998.

[73] McFarland DJ, Anderson CW, Müller KR, Schlögl A, Krusienski DJ. "BCI Meeting 2005 - Workshop on BCI signal processing, Feature extraction and translation". *IEEE Trans Neural Syst Rehabil Eng,* 14(2),135–8, 2006.

[74] Oralhan Z. "A New Paradigm for Region-Based P300 Speller in Brain Computer Interface". *IEEE Access*. 7,106618–27, 2019.

[75] Tzimourta KD, Astrakas LG, Gianni AM, Tzallas AT, Giannakeas N, Paliokas I, et al. "Evaluation of window size in classification of epileptic short-term EEG signals using a Brain Computer Interface software". *Engineering, Technology & Applied Science Research,* 8(4),3093–7, 2018.

[76] Lee T, Kim M, Kim S. "Improvement of P300-Based Brain – Computer Balancing Techniques". *Sensors*, 2020.

[77] Phanikrishna BV, Pławiak PJ, Prakash A. "A Brief Review on EEG Signal Pre-processing Techniques for Real-Time Brain-Computer Interface Applications," TechRxiv. September 30, 2021. doi: 10.36227/techrxiv.16691605.

[78] Xu Z *et al.*, "Temporal segmentation of EEG based on functional connectivity network structure," *Sci. Rep.*,13(1), 1–19, 2023. doi: 10.1038/s41598-023-49891-8.

[79] Olgun N, Turkoglu I. "Brain-Computer Interface Systems". 8th International Advanced Technologies Symposium (IATS'17), 19-22 October 2017.

[80] Pahuja SK, Veer K." Recent Approaches on Classification and Feature Extraction of EEG Signal , A

Review". *Robotica,*(May),1–25, 2021.

[81] Huang Z and Wang M. “A review of electroencephalogram signal processing methods for brain-controlled robots,” *Cognitive Robotics*, 1(July), 111–124, 2021. doi: 10.1016/j.cogr.2021.07.001.

[82] Turner MR, Wolpaw JR. "EEG-Based Brain-Computer Interfaces". *Physiology & behavior,* 176(1),139–48, 2016.

[83] Vanneschi L, Silva S." Introduction to Machine Learning." *Natural Computing Series,* (February 2010),115–48, 2023.

[84] Kononenko I, Kukar M. "Machine Learning Basics". *Machine Learning and Data Mining,* 59–105, 2007

[85] Oladipupo T. "Machine Learning Overview". *New Advances in Machine Learning,* ,8–18, 2010.

[86] Mahesh B. "Machine Learning Algorithms - A Review". *International Journal of Science and Research*. 9(1),381–6, 2020.

[87] RK, Reddy GYS, Pathak H." The Understanding of Deep Learning, A Comprehensive Review". *Mathematical Problems in Engineering,* 2021.

[88] Sengupta S, Basak S, Saikia P, Paul S, Tsalavoutis V, Atiah F, et al. "A review of deep learning with special emphasis on architectures, applications and recent trends". Knowledge-Based System, 194,105596, 2020. Available from, https,//doi.org/10.1016/j.knosys.2020.105596

[89] Al-Canaan A, Chakib H, Uzair M, Toor S, Al-Khatib A. "BCI-control and monitoring system for smart home automation using wavelet". *IET Signal Processing,* 16,141–56, 2022.

[90] Ghosh A, Sufian A, Sultana F, Chakrabarti A, De D. "Fundamental concepts of convolutional neural network". *Intelligent Systems Reference Library.* Vol. 172,519–567 p, 2019.

[91] Singh SA, Meitei TG, Devi MD, MajumderS. “A deep neural network approach for P300 detection-based BCI using single-channel EEG scalogram images,” *Physical and Engineering Sciences in Medicine*, 44(4), 1221–1230, 2021. doi: 10.1007/s13246-021-01057-4.

[92] Hossain KM, Islam MA, Hossain S, Nijholt A, Ahad MAR. “Status of deep learning for EEG-based brain–computer interface applications,” *Frontiers in Computational Neuroscience,* 16(1), 2023. doi: 10.3389/fncom.2022.1006763.

[93] Alanazi M, Aldahr RS, Ilyas M. "Human Activity Recognition through Smartphone Inertial Sensors with ML Approach". *Engineering, Technology and Applied Science Research*,14(1),12780–7, 2024.

[94] Küçükakarsu M, Kavsaoğlu AR, Alenezi F, Alhudhaif A, Alwadie R, Polat K. "A Novel Automatic Audiometric System Design Based on Machine Learning Methods Using the Brain’s Electrical Activity Signals". *Diagnostics,* 13(3), 2023.

[95] Guo M, Xu G, Wang L, Masters M, Milsap G, Thakor N, et al. "The anterior contralateral response improves performance in a single trial auditory oddball BMI". *BiomedicalSignal Process Control*, 22,74–84, 2015. Available from, http,//dx.doi.org/10.1016/j.bspc.2015.06.014

[96] Kimura R, Nambu I, Fujitsuka R, Maruyama Y, Yano S, Wada Y. "An auditory brain-computer interface to detect changes in sound pressure level for automatic volume control". *Heliyon,* 10(1),e23948, 2024. Available from, https,//doi.org/10.1016/j.heliyon.2023.e23948

[97] Ogino M, Kanoga S, Ito SI, Mitsukura Y. "Semi-Supervised Learning for Auditory Event-Related Potential-Based Brain-Computer Interface". *IEEE Access,*.9,47008–23, 2021.

[98] Borirakarawin M, Punsawad Y. "Hybrid Brain-Computer Interface System Using Auditory Stimulation and Speech Imagination Paradigms for Assistive Technology". *IEEE Access,* 11(May),53079–90, 2023.

[99] Wang J, Wang T, Liu H, Wang K, Moses K, Feng Z, et al. "Flexible Electrodes for Brain–Computer Interface System". *Advanced Materials,* 35(47), 2023.

[100] Halder S, Leinfelder T, Schulz SM, Kübler A. "Neural mechanisms of training an auditory event-related potential task in a brain–computer interfa ce context". *Human Brain Mapping*, 40(8),2399–412, 2019.

[101] Sun K, Hsieh K, Lee S. "Using Mental Shadowing Tasks to Improve the Sound-Evoked Potential of EEG in the Design of an Auditory Brain – Computer Interface". *Applied Science*, 2023.

[102] Barbosa S, Pires G, Nunes U. "Toward a reliable gaze-independent hybrid BCI combining visual and natural auditory stimuli". *J Neurosci Methods,* 261,47–61, 2016. Available from, http,//dx.doi.org/10.1016/j.jneumeth.2015.11.026

[103] Ogino M, Kanoga S, Muto M, Mitsukura Y. "Analysis of Prefrontal Single-Channel EEG Data for Portable Auditory ERP-Based Brain–Computer Interfaces". *Frontiers in Human Neuroscience,* 13(July),1–14, 2019.

[104] Jijomon CM, Vinod AP. "Person-identification using familiar-name auditory evoked potentials from frontal EEG electrodes". *BiomedicalSignal Process Control*, 68(April),102739, 2021. Available from, https,//doi.org/10.1016/j.bspc.2021.102739

[105] An WW, Shinn-Cunningham B, Gamper H, Emmanouilidou D, Johnston D, Jalobeanu M, et al. " "Decoding music attention from “EEG headphones”: A user-friendly auditory brain-computer interface ". I *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing, Toronto, ON, Canada, 6-* 11 June 2021.