Clinical Neurophysiology 127 (2016) 1288–1296

Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

Training leads to increased auditory brain–computer interface performance of end-users with motor impairments

S. Halder [a,b,*], I. Käthner [a], A. Kübler [a]

[a] Institute of Psychology, University of Würzburg, Marcusstr. 9-11, 97070 Würzburg, Germany
[b] Department of Rehabilitation for Brain Functions, Research Institute of National Rehabilitation Center for Persons with Disabilities, Tokorozawa, 359-8555 Saitama, Japan

ARTICLE INFO

Article history:
Accepted 5 August 2015
Available online 29 August 2015

Keywords:
Assistive technology
Communication
Brain–computer interfaces
End-user evaluation

HIGHLIGHTS

- Motor impaired end-users can communicate with an auditory brain–computer interface.
- End-users' performance increases with training.
- To our knowledge, this is the first time end-users controlled an auditory brain–computer interface speller with such a high accuracies and information transfer rates.

ABSTRACT

Objective: Auditory brain–computer interfaces are an assistive technology that can restore communication for motor impaired end-users. Such non-visual brain–computer interface paradigms are of particular importance for end-users that may lose or have lost gaze control. We attempted to show that motor impaired end-users can learn to control an auditory speller on the basis of event-related potentials.
Methods: Five end-users with motor impairments, two of whom with additional visual impairments, participated in five sessions. We applied a newly developed auditory brain–computer interface paradigm with natural sounds and directional cues.
Results: Three of five end-users learned to select symbols using this method. Averaged over all five end-users the information transfer rate increased by more than 1800% from the first session (0.17 bits/min) to the last session (3.08 bits/min). The two best end-users achieved information transfer rates of 5.78 bits/min and accuracies of 92%.
Conclusions: Our results show that an auditory BCI with a combination of natural sounds and directional cues, can be controlled by end-users with motor impairment. Training improves the performance of end-users to the level of healthy controls.
Significance: To our knowledge, this is the first time end-users with motor impairments controlled an auditory brain–computer interface speller with such high accuracy and information transfer rates. Further, our results demonstrate that operating a BCI with event-related potentials benefits from training and specifically end-users may require more than one session to develop their full potential.

1. Introduction

Brain–computer interfaces (BCIs) can provide a muscle independent communication channel for people with neurodegenerative diseases, such as amyotrophic lateral sclerosis (ALS), or acquired brain injuries (Vidal, 1973; Kübler et al., 2001). All BCIs require that a control signal is recorded, that can be modulated (either by attention or another mental task), if the system is intended for communication and control purposes. The methods for recording these signals range from electrophysiological (electroencephalogram (EEG) and magnetoencephalography (MEG)), metabolic (functional magnetic resonance imaging (fMRI) and functional near infrared red spectroscopy (fNIRS)) to invasive (electrocorticography (ECoG), microarray) recordings (Birbaumer et al.,

* Corresponding author at: Department of Rehabilitation for Brain Functions, Research Institute of National Rehabilitation Center for Persons with Disabilities, Tokorozawa, 359-8555 Saitama, Japan. Tel.: +81 4 2995 3100; fax: +81 4 2995 3132.
E-mail address: sebastian.halder@uni-wuerzburg.de (S. Halder).

http://dx.doi.org/10.1016/j.clinph.2015.08.007

1999; Weiskopf et al., 2004; Hill et al., 2006; Mellinger et al., 2007; Sitaram et al., 2007). In the EEG components that can be directly modulated are e.g., the sensorimotorrhythm (SMR) and slow cortical potentials (SCPs). Common components that are modulated by directing attention toward specific stimuli are steady state evoked potentials (SSEPs) and event-related potentials (ERPs). Of these components the P300 is often used for communication with motor impaired end-users (Nijboer et al., 2008).

The P300 ERP component of the EEG is elicited by a rare stimulus in a series of frequent stimuli. To utilize this potential for BCI control with visual stimuli the user attends to a matrix on a computer screen, often with 6×6 symbols (Farwell and Donchin, 1988). These symbols are then flashed in a random pattern. Visual P300 BCIs were used by end-users also with complex applications such as painting and web browsing (Mugler et al., 2010; Zickler et al., 2013; Halder et al., 2015) and can be used to achieve high communication speeds (Kaufmann and Kübler, 2014; Käthner et al., 2015). In some cases, if gaze or vision are not impaired, eye trackers may be used to accomplish the same task with higher speed and lower workload than visual P300 BCIs (Pasqualotto et al., 2015). It must be considered though, that some neurological diseases or acquired brain injuries can lead to impaired gaze control or vision which prevents focusing visual attention on a particular stimulus or direction. The solution to this problem may be eye-gaze independent P300 BCIs (Riccio et al., 2012). In principle any modality of human perception can be used to deliver stimuli. Somatosensory and auditory domains have proven to be the most practical (Kaufmann et al., 2013). Tactile stimuli have been used for applications such as communication and simulated wheelchair control (Brouwer and van Erp, 2010; Kaufmann et al., 2014). Two distinct branches of auditory P300 based BCIs have evolved. The first focuses on basic communication, mostly binary choice, with low complexity for users with severe brain injuries and short attention span that have no other method of communication (Halder et al., 2010; Hill and Schölkopf, 2012; Pokorny et al., 2013). These paradigms are also robust enough to be used by healthy controls in a mobile setting (De Vos et al., 2014). The second branch that has developed is that of multi-choice auditory P300 BCIs. This type of BCI is desirable for users with longer attention spans who wish to control more complex applications. Initial implementations of auditory multi-choice BCIs assigned words or stimuli based on naturally occurring sounds to rows and columns of the matrix (Furdea et al., 2009; Klobassa et al., 2009). Despite some success the requirements on sustained attention due to long selection times led to little success with end-users (Kübler et al., 2009). Continued development, in particular of the stimuli by including spatial cues, led to substantial increases in performance (Schreuder et al., 2010; Höhne et al., 2012).

In an attempt to combine spatial cues with the portability of stereo headphones we implemented an auditory P300 BCI utilizing interaural time difference (ITD) and interaural level difference (ILD) to create the impression of sounds originating from five different directions (Käthner et al., 2013). This concept was further improved by using animal sounds as stimuli (Simon et al., 2014). These stimuli require only a short duration to be recognized by the user and are easily discriminated. Additionally, we found that there is a strong training effect in healthy participants (Baykara et al., 2016). In this paper we proceed to show that this auditory BCI can also be applied to end-users with motor impairments.

2. Materials and methods

Five end-users were trained with an auditory P300 BCI using animal sounds as stimuli for five sessions on different days.

2.1. Participants

Five end-users agreed to participate in the study. See Table 1 for details. The ALS functional rating scale revised(ALS FRS-R) was administered in the first session (0 locked-in state, 48 no impairment). None of the end-users had previous experience with the auditory BCI paradigm.

2.2. Procedure

To investigate the effects of training with the auditory BCI the end-users participated in five sessions on separate days. On the basis of the findings in (Baykara et al., 2016; Simon et al., 2014) we retrained the classifier with two five-letter words at the beginning of every session. The words consisted of the letters on the diagonal of the 5×5 letter matrix used in this study (AGMSY; see Fig. 1(A)). After calibration the participants wrote five words comprising five letters each (VARIO, GRUEN, RUBIO, TUMBI, PHLEX). The words were chosen such that the stimuli needed to select the target symbols were equally distributed across the five different rows and columns.

The stimuli used in this study were modified to create an impression of directionality using ITD and ILD as described in Käthner et al. (2013). We used the stimuli found to provide the best results in Simon et al. (2014). Duck, bird, frog, gull and pigeon were arranged on positions of a circle from left, middle left, front, middle right to right and presented using Sennheiser HD280 Pro stereo headphones (see Fig. 1(B) for the spatial arrangements and (C) for the spectrograms).

None of the end-users except end-user one were provided with the static visual support matrix in Fig. 1(A). All other end-users received only auditory cues as described in Fig. 1.

End-user two completed four sessions only. In session three the quality of the calibration data was not sufficient to train the classifier for feedback. End-user three requested to discontinue participation in the study after session three.

2.3. Calibration

For calibration each stimulus was presented ten times. Thus, for selection of one letter the participants had to attend to twenty stimuli (ten for the row, ten for the column). On the basis of the calibration data the number of repetitions was reduced to avoid a ceiling effect. We set the number of repetitions such that 70% accuracy was reached (based on offline analysis of the calibration data) plus three additional repetitions. This procedure was repeated in each daily session. Previous calibration data was not used. Timings were identical to the copy spelling task (except number of repetitions). Collecting the calibration data required approximately ten minutes. In all sessions all calibration data was used to calibrate the classifier.

2.4. Copy spelling

In each session the end users wrote five words comprising five letters (see above). The system paused for twelve seconds between letter selections to give the user time to focus on the stimulus that needed to be attended, to select the next letter (see Fig. 1). There was a two seconds pause between the stimuli for the rows and those for the columns (also Fig. 1). The stimuli lasted for 150 ms with an inter stimulus interval (ISI) of 287.5 ms. These parameters (437.5 ms stimulus presentation, ten stimuli, ten repetitions, two seconds between rows and columns, twelve seconds between letters) lead to a maximum time of 57.75 s per letter. These parameters are identical to the ones used in Baykara et al. (2016). Note, that in most cases the time per letter was reduced