Spelling with brain-computer interface - current trends and prospects
Hubert Cecotti

To cite this version:

HAL Id: hal-00553422
https://hal.archives-ouvertes.fr/hal-00553422
Submitted on 10 Mar 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Spelling with Brain-Computer Interfaces  
Current trends and prospects

H. Cecotti 1,2  
1 GIPSA-lab CNRS UMR 5216  
38402, Saint Martin d’Hères, France  
2 Institute of Automation (IAT) - University of Bremen  
Otto-Hahn-Allee, NW1  
28359 Bremen, Germany

Abstract  
Brain-Computer Interfaces (BCIs) have become a large research field that include challenges mainly in neuroscience, signal processing, machine learning and user interface. A BCI can allow the direct communication between humans and computers by analyzing brain activity. The main purpose for BCIs is to enable communication for people with severe disabilities. Indeed, spelling was one of the first BCI application. While spelling can be the most basic application, it remains a benchmark for communication application and one major challenge in the BCI community. This paper focuses on the current main strategies for spelling words. It includes recent BCIs based on P300, Steady-State Visual Evoked Potentials (SSVEP) and motor imagery. This work highlights some current challenges in BCI spellers and virtual keyboards.

Keywords
Brain-Computer Interface, Speller, P300, SSVEP, Motor imagery.

1 Introduction

A Brain-Computer Interface (BCI) translates brain activity into computer commands. This can be achieved thanks to the detection of particular brain responses. This type of system allows people to communicate through direct measurements of brain activity, without requiring any movement [4, 6, 24]. The modulation of brain signals can be recorded from the scalp using electroencephalography (EEG), from cortical surface using electrocorticography (ECoG), or from neurons directly within the cortex [44]. BCIs may be the only hope for millions of people with motor disabilities so severe that they cannot communicate with their families. In spite of their motor disabilities, sensory and cognitive functions are usually still enabled. For instance, people with spinal cord injuries or amyotrophic lateral sclerosis (ALS), also called Lou Gehrig’s disease, are good candidates for using daily BCIs [6, 7]. Although, ECoG is an efficient way to record brain activity, non-invasive BCI through EEG recording remain a major research topic as EEG provides a high time resolution in the signals and EEG recording requires relatively inexpensive equipment. In addition, non-invasive BCI could be used by healthy people as a complement to other interfaces. A BCI can be decomposed into four parts [35]. First, the signal is acquired via an amplifier. Then, the signal is processed and assigned to different classes. Finally, the classes are sent to the output device components and the operating protocol links all the previous components. The signal classification component is composed of the brain signal features extraction and the translation of these signals into device commands. The EEG classification strategy depends on the response to detect: event-related potentials (ERP), steady-state evoked potentials (SSVEP), motor imagery or slow cortical potentials. These responses can be provoked by external stimuli, visual or auditory, (P300, SSVEP) or not (motor imagery). The expected EEG drives the classification to some specific feature extraction methods. In this paper, we focus on the non-invasive BCI systems allowing communication through spelling. Indeed, spelling is the main application in interfaces. Although the type of BCI application is extended every year with video games, robotic arm control, wheelchair control,... [14, 28], the most basic application, spelling, is still one active research area. Spelling can indeed enable the physically challenged to perform many activities. It can therefore improve their quality of life. To some extend, it can allow them more independence. This independence can be translated into social cost reduction. Besides, disabled people who can spell could work and get a more rewarding place in society. The rest of the paper is organized as follows: Spellers based on the detection of the P300 wave are first presented. The main spellers based on SSVEP are described in the third section. Then, some spellers based on the detection of motor imagery are presented. Finally, the interest of these spellers are discussed.
2. Spellers based on P300

The detection of event related potentials (ERP) is one way for creating a BCI. A typical ERP based BCI is the P300 speller, which allows people to spell characters. The P300-Speller is one of the first BCIs, it was first introduced in 1988 by Farwell et. al [17]. The P300 wave is an ERP [36]. Its generation is possible thanks to the oddball paradigm. This paradigm provides random visual stimuli that cause a surprise effect to the subject. The classical P300-Speller layout is presented to the user on a computer screen as depicted in Fig. 1. It is composed of a $6 \times 6$ matrix, which contains all the available characters [17, 16]. During the experiments, the user has to focus on the character she/he wants to spell. When the user focuses on a cell of the matrix, it is possible to detect a P300 (a positive deflection in voltage at a latency of about 300 ms relative to the stimuli onset in the EEG) time-locked to the onset of the cell intensification. To generate ERPs, the rows and columns are intensified randomly. Row/column intensifications are block randomized in a number of events equal to the number of rows and columns. The sets of intensifications is repeated $N_{\text{epoch}}$ times for each character.

Although, the P300-Speller paradigm was initially designed to spell Latin characters, a recent P300-Speller has been proposed for Chinese characters [23]. Most of the improvements in the P300-Speller have been achieved at the signal processing and detection level with techniques like Support Vector Machines [33] or neural networks [11, 26], Bayesian linear discriminant analysis [12, 22]. In contrast, the P300-Speller graphical user interface has not much been evolving for more than two decades. Some improvements have been proposed like in [37], where the color of the flickering matrices should be green/blue. Townsend et al. have presented an alternative paradigm to the classical flashes of the row/column paradigm (RCP). They have proposed the checkerboard paradigm (CBP), which is a combination of two checkerboards to avoid the well known confusion problems in the neighborhood of the target. Using an $8 \times 9$ matrix of alphanumeric characters and keyboard commands, 18 participants used the CPB and RCP paradigms. With approximately 9-12 min of calibration data, they obtain a mean online accuracy of 92% for the CPB, better than the RCP with only 77%. The mean information bit rate was also significantly higher for the CPB with 23 bits per min (bpm), than for the RCP, 17bpm.

In the P300-Speller, research works are carried out in several directions. First, the P300 speller can be improved through methods that increase the reliability over time and across subjects, and the information transfer rate [27]. Second, with a more pessimist point of view about the possible improvement of the P300 detection, the improvement should come with an application oriented P300-Speller, i.e., with word completion, prediction, a knowledge of the vocabulary...

Current P300-Spellers are still far from commercial/clinical applications and some efforts shall be made to let BCIs leave laboratories. The available system lacks robustness over time and across subjects. The performance does not meet user’s requirements due to unadapted end-user interface. However, there exist few exceptions like in [40], where a late stage ALS patient could use at home at P300-Speller developed by the Wolpaw lab. The French National Research Agency (ANR) through the RoBIK (Robust BCI Keyboard) project aims at developing such success story to provide efficient user-dedicated BCI that will be easily daily used by a non-technician staff, e.g., a nurse. The intendiX solution has been proposed by g.tec in 2009. This BCI is designed to be installed and operated by caregivers or the patient’s family at home. The system is based on visually evoked EEG potentials (VEP/P300). It allows the user to sequentially select characters from a keyboard-like matrix on the screen just by paying attention to the target for several seconds. Contrary to the classical P300-Speller, the matrix has a $5 \times 10$ size as depicted in Fig. 2. This BCI requires some training but most subjects can use intendiX after only 10 minutes with a reasonable performance. According to the g.tec company, the performance for the majority of healthy users during their first trial is estimated to a spelling rate of 5 to 10 characters per minute (cpm). It is worth mentioning that the intendiX system is able to detect the idling state. Therefore, the system only selects characters when the user pays attention to it. This system allows also the patient to trigger an alarm, let the computer speak the written text, print out or copy the text into an e-mail or to send commands to external devices. In addition, this BCI is proposed as a whole package (software, amplifier, caps,...), which provides a global solution.

3. Spellers based on SSVEP

In a BCI based on Steady-State Visual Evoked Potentials (SSVEP), the system reflects the user attention to an oscillating visual stimulus [41]. Flicker-
ing lights at different frequencies are usually used as stimuli. Their responses appear in the visual cortex and correspond to SSVEP at the same frequencies and higher harmonics \[31\]. The amplitude and the phase that define an SSVEP response depend on the frequency, intensity and the structure of the repetitive visual pattern \[45\]. SSVEP based BCIs have been used in many types of applications like for neuroprosthetic devices control, for the restoration of the grasp function in spinal cord injured persons \[30\] and video games \[25\]. Indeed, this type of BCI performs very well and is reliable according to previous studies \[5, 10, 19, 18\]. While each cell of the matrix can correspond directly to a command in the P300-Speller, the low number of commands of an SSVEP-BCI involves an adaptive strategy for creating the graphical user interface. Among the different laboratories working on SSVEP-BCI, the Institute of Automation in Bremen, Germany has proposed several efficient SSVEP-Spellers.

The SSVEP based Bremen-BCI speller has been evaluated during the CeBIT fair 2008 in Hannover, Germany and RehaCare 2008 in Düsseldorf, Germany \[3, 13, 39, 43\]. The graphical user interface of Bremen-BCI is presented in Fig. 3. This interface is composed of a virtual keyboard with 32 characters (letters and special symbols), which is located in the middle of the screen. The five white boxes at outer edges and upper left corner of the screen are flickering with different frequencies. These boxes correspond to the commands “left”, “right”, “up”, “down”, and “select”. The subject does not need to shift his gaze too much, because the used stimuli are part of the GUI on the same LCD screen. This setup, as opposed to having an LCD for the GUI and a separate LED board for the visual stimuli, is much more convenient for the user as they do not have to shift their gaze too much.

In the command level, \textit{i.e.}, the five commands, the mean accuracy of the command detection is 92.84\%, with an average information transfer rate of 22.6bpm. In the speller level, the average information transfer rate is 17.4bpm, equivalent to about 3.5cpm.

The SSVEP speller developed by Cecotti \[9\] (CBCI) is a recent SSVEP-Speller that does not need any calibration step. Thus, this speller is ready to work once the subject is prepared. This speller was also developed at the Institute of Automation in Bremen, Germany. The visual stimuli are here fully integrated to the graphical user interface (GUI). Contrary to some other SSVEP-BCIs, the visual stimuli and the commands are merged. This speller allows writing 27 characters: the 26 Latin characters \[A..Z\] and Ś \[ś\] for separating the words. CBCI is depicted in Fig. 4. This interface corresponds to a menu with three possible choices. When a choice is selected, then the content of this choice is splitted into three new choices. Three commands are dedicated to the navigation. They correspond to the three boxes that contain all the possible letters. For writing a letter, the user has to produce three commands. This number of command is fixed and independent of the letter. One command is considered for canceling the previous one. An easy access to the “undo” command must be present for enabling easily a fast correction from the user. An error can come from the user directly or indirectly. This command aims at minimizing the cost of a mistake during spelling tasks. A command is dedicated to the deletion of the last character in the written text. At any moment, the user is able to suppress the last character of the text with only one command.

CBCI was tested on eight healthy subjects. The average accuracy and information transfer rate are 92.25\% and 37.62bpm, which is translated in the speller with an average speed of 5.51cpm. One subject could write with an average speed of 7.34cpm.
4. Spellers based on motor imagery

Like for SSVEP-BCIs, the number of available commands limits the interface: it is not possible to assign an imagery movement to every character. A strategy must be found to combine few basic BCI commands, e.g., thinking to moving the left/right hand. A predictive BCI speller based on motor imagery has been proposed at AIRLab, the Artificial Intelligence and Robotics Laboratory at the Department of Electronics and Information of the Politecnico di Milano, the Technical University of Milan, Italy [15]. The GUI of this speller is presented in Fig. 5. The selection strategy is based on target expansions, like a menu, with 27 available characters, numbers, symbols. This speller possesses a predictive capabilities: it allows word suggestions and disabled improbable symbols.

The achieved performance are relevant. With one subject, they have obtained a high classification accuracy and the overall speller speed is estimated to 3cpm. With two other subjects, they started with lower classification accuracies, but significant improvements have been achieved with more training sessions. These subjects reached a spelling speed of respectively 2 and 2.7cpm [15].

The BCI research group from the Fraunhofer FIRST (IDA), Berlin, Germany has proposed the Berlin BCI (BBCI) called Hex-o-Spell [8, 29]. This asynchronous BCI speller allows to write 29 different characters and the backspace command. The speller is controlled by two mental states: imagined right hand movement and imagined right foot movement. Six hexagonal fields are surrounding a circle. In each field, five characters or other symbols like backspace are arranged. An arrow is placed in the center of the circle for the selection of a character. When the subject imagines a right hand movement the arrow turns clockwise. With an imagined foot movement, the rotation stops and the arrow starts extending to the desired field. Once the field is selected, the six fields are arranged with the content of the selected field. The BBCI has been tested in real condition on two volunteer and healthy subjects during the CeBIT fair 2006 in Hannover, Germany. The speed of the hex-o-spell BCI was between 2.3 and 5cpm for one subject and between 4.6 and 7.6cpm for the other one. This speed was measured for error-free, completed sentences, i.e., all typing errors that have been committed had to be corrected by using the backspace of the mental typewriter. This protocol was also used to evaluate CBCI. The original and efficient Hex-o-spell interface has also recently been tested with the visual oddball paradigm [38].

5. Discussion

In spite of the different results reported in the literature, it is not possible to have an objective comparison between the different available BCI spellers due to the inter-subject variabilities and the conditions of the experiments. For instance, the experimental conditions are very different between a dedicated EEG room in a laboratory and a booth at an international fair with all the surrounding noise. However, each BCI paradigm possesses its advantage and drawbacks. BCI s based on the detection of the P300 or SSVEP require external visual stimuli. For spelling applications, the visual stimuli are not really a disadvantage. Indeed, spellers based on motor imagery consider also a graphical user interface. With the P300 speller, each symbol is usually available on the screen, like a classical virtual keyboard. Contrary the P300 speller, an SSVEP speller must take into account several constraints based on the visual stimuli. With LEDs, it is possible to produce a large number of visual stimuli with different frequencies [19, 46]. However, such solution requires an external device; the application and the visual stimuli are not located at the same place. With visual stimuli on an LCD screen, the size, the low luminosity, the vertical refresh rate of the screen are some parameters that limit the number of simultaneous visual stimuli on the screen. For this reason, it is not possible to propose to the user a virtual keyboard with a direct access to the letters. Other BCI commands shall be used to navigate on the virtual keyboard. For an efficient BCI, the performance
shall be reliable over time and across subjects. BCIs based on motor imagery, like the Hex-o-spell, can be efficient. However, BCIs based on motor imagery or P300 requires a training session for the calibration of the system. In addition, BCIs based on motor imagery suffer of BCI illiteracy; the performance is highly dependent of the subject. On the other hand, SSVEP-BCIs do not require a training session and possess a high transfer rate [5, 9].

A low BCI performance can be due to a lack of attention, to the disrespect of what should be written. It is possible that the user wants to produce a command but the signal processing module delivers the wrong command. In this case, the error is not voluntary and shall be corrected easily. At any moment, the user should be able to cancel the previous command with only one command.

6. Conclusion

Communication through spelling is still one of the main challenge in BCI applications. Writing a simple message, an e-mail,... remains a difficult task to achieve for people with severe disabilities. The BCI literature has expontially increased in the past few years. Whereas recent BCI competitions have allowed to compare different machine learning methods, these benchmarks are limited to one aspect of a BCI. The graphical user interface should actually benefit the comparison between spellers [35, 34]. Such solution could provide faster and more robust spellers. They could solve to some extent the BCI illiteracy. This problem can be determinant for the choice of a specific BCI. Recent works have been conducted to address this problem: for P300 [21, 20], SSVEP [3], and sensorimotor rhythms [42]. Further work should be carried out in the comparison of different well known and proven BCI systems with the same set of subjects. The different materials (amplifiers, caps,...) could be an obstacle for comparing and sharing BCIs. Hopefully, well established and promising BCI frameworks like BCI2000 and OpenViBE could allow a better comparison between spellers [35, 34].

Acknowledgment

This work has been supported by a Marie Curie European Transfer of Knowledge grant BrainRobot, MTKD-CT-2004-014211, within the 6th European Community Framework Program, and by French National Research Agency (ANR) through TecSan program (project RoBIK ANR-09-TECS-013).

References


