A NEW P300 NO EYE-GAZE BASED INTERFACE: GEOSPELL

F.Aloise[1], P. Aricò[1], F. Schettini[1], A. Riccio[1], M. Risetti[1], S. Salinari[2], D. Mattia[1]
F. Babiloni[1,3] and F. Cincotti[1]
[1]*Neuroelectrical Imaging and BCI Lab, Fondazione Santa Lucia IRCCS, Rome, Italy*
[2]*Dept. of Computer Science, Univ. of Rome "Sapienza", Rome, Italy*
[3]*Department of Human Physiology and Pharmacology, Univ. of Rome "Sapienza", Rome, Italy*

Keywords: Brain Computer Interface, Covert Attention, P300, Eye Tracker, Workload, Electroencephalogram (EEG).

Abstract: Brain Computer Interface (BCI) is an alternative communication system which allows users to send commands and/or messages toward the outside not crossing the normal output channels of the brain, but conveying these outputs from the human brain to a computer (Wolpaw et al., 2002). In an EEG-based BCI messages are obtained from brain activity. This study presents a novel P300 based Brain Computer Interface requiring no eye gaze, and so usable in covert attention status, called GeoSpell (Geometric Speller). GeoSpell performances have been compared with those obtained by the subjects with the standard 6 by 6 P300 Speller (P3Speller) matrix which depends on eye gaze (Farwell and Donchin, 1988). A NASA Task Load Index (TLX) workload assessment was employed to provide a subjective rating about the task's workload and satisfaction with respect to both the interfaces (NASA Human Performance Research Group 1987). Results shown comparable workload values for P3Speller and Geospell; this result has an important impact in term of efficiency and satisfaction for the use of the BCI devices. Geospell interface has shown an accuracy comparable with the P3Speller one but with a lower bit-rate.

1 INTRODUCTION

Brain Computer Interfaces (BCIs) are able to recognize the intention of the subject of completing a particular action and to translate it into control signals for technological devices, through particular transfer algorithms.

The "communicative power" of the BCI systems is very important for people with physical disabilities; e.g. Amyotrophic Lateral Sclerosis (ALS) causes the partial or total loss of the muscles control, while sensory and cognitive functions remain usually intact; this disease, in advanced state, leads to the partial or total loss of the ability to move eyes. A BCI able to translate specific mental trials in a control actions could allow such persons to interact with the surrounding environment improving their autonomy and their quality of life. Different types of brain activity are discernible in EEG signals and are used in EEG based BCIs: e.g., the P300 potential is a positive deflection (ca. 10-20μV) that occurs about 250-400 ms after the presentation of a target stimulus (Fabiani et al., 1987; Polich et al., 1995). This is the case of "Oddball" paradigm during which rare target items are presented within a sequence of frequent No-Target (or "standard") items; in this kind of paradigm the subjects are asked to focus their attention to the Target stimulus (e.g. mentally counting the number of Target occurrences or pushing a button on a keyboard when the user recognizes the Target), and to ignore the other stimuli (No-Target).

1.1 Attention: "Covert" vs "Overt"

An important ability of our cognitive system is the possibility to select, by attentional mechanism, just a part of the big amount of information we are at any time subjected to. Selective attention, is the cognitive process of selectively concentrating on one aspect of the environment while ignoring other things (Anderson, 1999). Fixing an object does not necessarily means to see it and to focus the attention on it: we can focus the attention on a specific target of the visual field directing the eyes towards the stimulus source (overt attention), or mentally

Aloise F., Aricò P., Schettini F., Riccio A., Risetti M., Salinari S., Mattia D., Babiloni F. and Cincotti F..
A NEW P300 NO EYE-GAZE BASED INTERFACE: GEOSPELL.
DOI: 10.5220/0003161202270232
In *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing* (BIOSIGNALS-2011), pages 227-232
ISBN: 978-989-8425-35-5

focusing on one of several possible stimuli, without the necessity of gazing on it (covert attention).

2 STATE OF THE ART

2.1 P3Speller Interface

P3Speller (P300 Speller) is an interface developed by Farwell and Donchin (Farwell and Donchin, 1988) (Figure 1a). It allows the subject to select 36 alphanumeric characters positioned in a matrix, using as control feature the P300 event related potential (ERP). Stimuli are presented to the user on a computer screen and randomly intensified at an established frequency. During the stimulation, the user focuses his attention on the character he intends to select and then he mentally counts the number of occurrences while rows and columns are flashing. The flashing of the selected target elicits a P300 potential, while the others (No-Target) do not.

One of the main problems to recognize the P300 is the lowest signal-to-noise-ratio (SNR); for this reason, each character is intensified more than once time (e.g. 8 occurrences for the rows and 8 for the columns for this study) in order to extract the components of interest from background noise averaging the Target and No-Target stimuli.

2.2 P3Speller in Covert Attention Condition

As mentioned above, ALS patients, in advanced stage of the illness, could manifest a paralysis of the ocular muscles losing possibility to freely move the eyes.

A recent study of Brunner et al. showed that the P3Speller performances dramatically decrease when the user is unable to move his eyes (Brunner et al., 2010).

In this regard, Treder and Blankez tested the ERP-based Hex-o-Spell, a two-levels speller consisting of six discs arranged on an invisible hexagon, which does not require eye gaze. They reported classification accuracy about 60% (Treder and Blankez, 2010).

The purpose of this study has been to plan and evaluate a new speller interface P300-based usable in covert attention condition too.

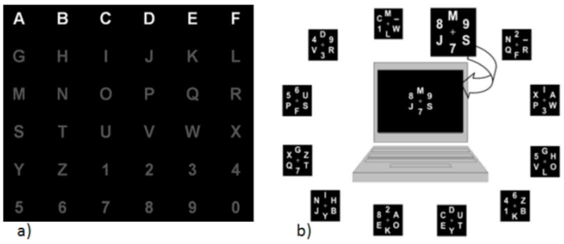


Figure 1: a) Farwell and Donchin speller paradigm P300-based b) The proposed GeoSpell (Geometric Speller) interface; each group contain 6 alphanumeric characters, that are presented in a random sequence to the centre of a screen.

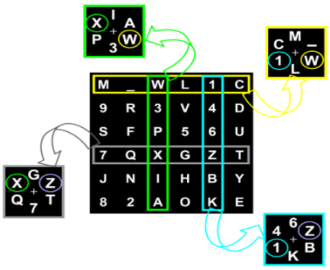
3 METHODS AND MATERIALS

3.1 GeoSpell

In the GeoSpell interface (Figure 1b) characters are organized following the same logic of a N by N matrix: there are a total of N^2 characters and they are organized into 2N groups of N characters each. Characters of the same group are placed at the vertices of a regular geometric figure, and a fixation point is placed at its center. Each character belongs to two groups occupying the same position; the single selected character will be given by the intersection of two groups.

The visual angle subtended by the fixation cross at the center of the visual field and each character doesn’t exceed 1° ; in this way all characters are recognizable by the users (Sutter, 1992). The characters for each group has been chosen so that the number of white pixels between different groups is almost constant (Mean=3274.333pixel; Std=2.93%). This choice allows to avoid the occurrence of no-ERP potentials (e.g. VEPs) that otherwise would be elicited during the stimulation. This eventuality is instead inevitable in the P3Speller interface. Stimulation consists in a random presentation of all the groups, in particular every group was enlighten for 125 ms and a 250 ms lag between the onsets of two consecutive stimuli.

Seven volunteer subjects (4 male, 3 female, Median age=27.75, Std=4.6), with no history of mental or neurological illnesses were involved in this study. Every subject had previous experience with BCI and with the P3Speller interface. Scalp EEG data were acquired using BCI2000 software (Schalk et al., 2004). The EEG was recorded using a cap embedded with 16 Ag/AgCl electrodes covering left, right, and central scalp locations (Fz, FCz, Cz, CPz, Pz, Oz, F3, F4, C3, C4, CP3, CP4, P3, P4,



M	_	W	L	1	C
9	R	3	V	4	D
S	F	P	5	6	U
7	Q	X	G	Z	T
J	N	I	H	B	Y
8	2	A	O	K	E

Figure 2: Each group contains the characters of one row or one column of a matrix.

PO7, PO8) referenced to both earlobes, and grounded to the right mastoid, based on the 10-10 standard of the International Federation. The electrode impedance did not exceed 10 kΩ. The EEG was acquired with a g.Tec gUSBamp acquisition device (Austria, sampling rate 256Hz). In order to demonstrate that the GeoSpell interface does not require eye gaze we have performed recordings using an eye tracker system (spatial resolution of 0.5°) constituted by an infrared-light webcam "Genius iSlim 320" managed by the open source software "ITU GazeTracker" (San Augstin et al, 2010). After a phase of calibration, the software, in START mode returns X and Y screen coordinates of the eye gaze; through UDP communication, In this way we synchronized the BCI2000 and the ITU gaze tracker, and calculated, through offline analysis, the number of ocular movements and blinks made by subject during the stimulation, and Target/s on which he/she has moved the eyes.

3.2 Usability

The introduction of the BCI technology into the home-place would have a great impact on opportunities available to severe motor disable people. From this point of view approaching in the assessment of BCI technology development in terms of usability is a first important stage. In this study we focused on the evaluation of user's mental workload in operating the two different interfaces. In order to compare the workload of the GeoSpell and the P3speller, we used a subjective workload rating scale called NASA-tlx (Hart and Staveland, 1988). The NASA-tlx assesses the workload by considering six different factors: Mental, Physical and Temporal Demands, Frustration, Effort and Performance.

The workload has a direct bearing on the usability of a software interface. If fewer mental resources are used, then the efficiency, and also the effectiveness and satisfaction associated with the interface can be increased.

3.3 Experimental Protocol

As mentioned earlier each participant to the protocol had previous experience with the interface P3Speller, instead the GeoSpell was presented them for the first time; this could represent an intrinsic bias between the two interfaces. Furthermore recent work refers that the mental training can significantly affect attentive stability, brain function, and selectively reduced cognitive effort (Lutz et al. 2009). So each subject was asked to take part in 4 training sessions without EEG acquisition with the GeoSpell before to start with the effective data collecting protocol. The purpose of this sessions was to get used the subjects to the new interface; in each training session subjects were asked to attend to some letters and push a button when they occur; every letter on the interface was presented as a Target with the same incidence and we monitored the number of lost Targets. Every session consists of 9 runs of 6 trials each, where a trial denotes a fixed number of stimulation sequences during which the target is the same. We set 8 stimulation sequences for trials and then each letter is presented 16 times. Proceeding with the training, we noticed for all the subjects a diminution of the number of lost Targets and an arrangement in the reaction times.

Data collecting protocol consists of 5 sessions during which we compared the 2 different interfaces, using a visual oddball paradigm as a baseline both for reaction times and for waveforms features. During the first 4 sessions we asked the subject to perform 3 runs with each different stimulation interface. The system suggested to the user the letter that he had to concentrate on before that the stimulation began. We used 6 different words of 6 different characters per word (for a total of 36 different characters) as a text to spell: "AX6L1O", "TVM3CH", "2EWY_8", "BJZN7G", "DR5K9Q", "FU4SPI". The characters of the same word were chosen to occupy all possible positions within a group. As mentioned earlier, every stimulus was intensified for 125ms, with an inter stimulus interval (ISI) of 125 ms, so 250 ms lag between the onset of two different stimuli. Also we have set of pseudorandom stimulation sequences to ensure that at least 500 ms elapsed between two target stimuli. This avoids the "Attentional Blink" phenomenon that occurs when the Target to Target Interval (TTI) is shorter than 500ms (Raymond et al. 1992). We provided a 2 seconds pre-trial presentation, during which the target appeared in the its group position; in this way the subject knew the Target position before stimulation started. The first 2 sessions were about response times and EEG data acquisition was not required; the subject had to keep his eyes fixed on the cross at the center of the interface and push a button every time a Target stimulus appeared. The last session aimed at a direct comparison of the online performances of GeoSpell and P3Speller. Data of third and fourth sessions were used to extract the control features for each participant; in particular we used a Stepwise Linear Discriminant Analysis (SWLDA) to select the most relevant features that allowed to discriminate Target Stimuli from NoTarget one (Krusienski et al. 2006). The two interfaces were put on in the same operational conditions; particularly, for the P3Speller online classification, 8 stimulation sequences per trial were used, and before the beginning of every trial, subject had 4 seconds of "Presentation" during which the stimulation was off and he could look for the Target of interest. For GeoSpell we provided 10 stimulation sequences: the first 2 sequences (Presentation), allowed the subjects to find the wanted Target; during these 2 sequences, each letter was presented 4 times. Rather the following 8 sequences of stimulation were used for online classification. In both interface a feedback on classification result was given at the end of each trial. For text to spell we select two made sense Italian word that move on all the different positions in the GeoSpell's group (as it happened for the words in previous sessions).

4 RESULTS

4.1 Reaction Time

We used the 2 sessions without EEG acquisition to compare the reaction times of the 2 text writing interface with the visual oddball paradigm. Figure 3 shows the mean of reaction time for each stimulation interface relating to the 2 different sessions.

Geospell interface exhibited an averaged reaction times statistically different ($p < .05$) from each other; such result was expected, because the covert attention condition increases the difficulty of the discrimination task with respect to overt attention condition; the number of missed Targets confirms this results, in fact GeoSpell interface produced a greater number of lost Targets with respect to other interfaces.

4.2 Offline Counting Accuracy

The data collected during third and fourth sessions were used to determinate counting accuracy. In particular we performed a cross-validation exploring all the possible combinations of training and testing data set from the initial data set. For each participant, counting accuracy was determined depending on the number of stimulation sequences mediated during the trial.

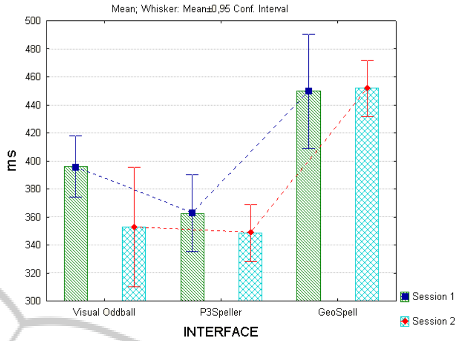


Figure 3: Mean Error (0.95 CI) of reaction times for each session and the 3 interfaces.

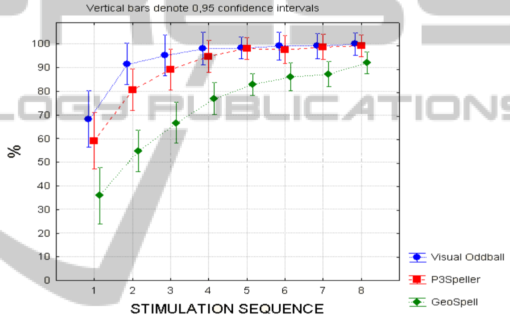


Figure 4: ANOVA test for the accuracy of the 3 different interfaces depending on the number of stimulation sequences.

Then, we analyzed accuracy values using two-way repeated measures ANOVA, using *Interface* and *Number of Stimulation Sequence* as factors (figure 4). After that, we performed two way t-test ($\alpha = .05$) between P3Speller and GeoSpell for each *Number of Stimulation Sequence*. Results are summarized in Table 1: GeoSpell reached comparable performances with P3Speller after a high number of stimulation sequences.

Table 1: *t* and *p* values of t-test for each stimulation sequence.

	SeqSt 1	***SeqSt 2***	***SeqSt 3***	***SeqSt 4***	***SeqSt 5***	***SeqSt 6***	***SeqSt 7***	***SeqSt 8***
p	0,034	0,001	0,006	0,063	0,006	0,050	0,031	0,078
t	2,52	4,13	3,63	2,16	3,52	2,22	2,74	2,12

In conclusion, the GeoSpell interface exhibits a lower bit-rate than P3Speller, but the performances in terms of accuracy are comparable, since the differences among the performances using GeoSpell and the P3Speller decrease when the number of stimulation sequence increases.

4.3 Event Related Potentials

Some analyses have been performed on the amplitudes and latencies of the P300 and N200 ERPs (elicited by the Target stimuli). In particular we compared the P3Speller and GeoSpell with two-way repeated measures ANOVA, with *Interface* and *Amplitude/Latency* as factors.

Peak amplitude and peak latency were determined for each subject by picking the largest positive or negative peak for all the sites within particular intervals, these were selected through a Grand-Average on the EEG signal of all the subjects, for the two interfaces (Table 2).

Table 2: Interest intervals for ERP's amplitude and latency.

LATENCY [ms]	*P3SPELLER*	*GEOSPELL*
P300	[220 : 400]ms	[400 : 600]ms
N200	[150 : 250]ms	[250 : 400]ms

There were no statistically differences between N2 and P3 amplitudes for the two different spellers ([N2] Interface: $F = 0.38462$, $p = 0.55239$; [P3] Interface: $F = 2.0602$, $p = 0.18911$). Instead, N2 and P3 latencies were longer for GeoSpell than for P3Speller ([N2] Interface: $F = 64.624$, $p = 0.00004$; [P3] Interface, $F = 54.862$ $p = .00008$). The increase in the N2 and P3 latencies using the GeoSpell interface, was caused by the increase of the task difficulty. E.g., changes in ERPs component latency between different groups and conditions can be assumed to reflect changes in stimulus processing; P300 latency is often correlated with task difficulty: P300 peak latency is longer for the more difficult compared to the easier tasks (Allison et al., 2003).

4.4 Online Counting Accuracy

We determined the online accuracy basing on the results of the fifth session. Accuracy per subject and mean accuracy (AVG) were depicted in figure 5 for each interface.

The results of the online session, confirm those of the copy-mode sessions: the use of the P3Speller allows performances (AVG = 95.75%, Std = 2.45) more elevated than using the GeoSpell interface (AVG = 83.18%, Std = 8.29); the Std value in the GeoSpell performances, shows a great variability in the performances among the subjects respect to P3Speller. However it must be remembered that the two interfaces have been used in different attentional conditions; previous studies (Brunner et al., 2010) shown that the use of P3Speller in covert attention condition causes a significant performances decrease that doesn't allow to use P3Speller as "Communicative Mean".

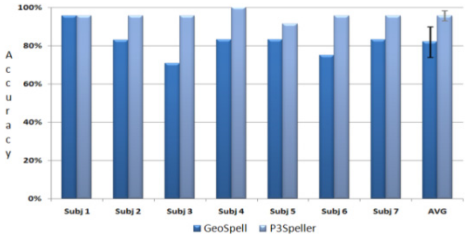


Figure 5: Online classification accuracy for each subject.

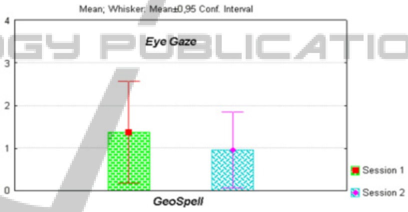


Figure 6: Mean error (0.95 CI) related to subjects eye movements during the stimulation for the 2 copy mode sessions.

On the contrary the performances achieved with the GeoSpell interface far exceed 70%, value that represents the threshold above which an interface can be defined efficient in communication terms (Kübler et al. 2006).

The small number of eye movements recorded during the third and fourth capture sessions confirmed the hypothesis of the covert attention condition. Figure 6 depicts the mean error (0.95 CI) referred to the number of ocular movements performed by the subjects during the 2 copy-mode sessions. In each run, the number of the stimuli presented during the last 8 sequences was 96.

4.5 Workload Results

Two repeated measures ANOVAs were conducted separately for the workload scores of the online and the offline sessions, with the *GeoSpell task* and *P3speller task* entered as the independent factors. Although the workload scores of the GeoSpell tasks (offline: mean=37.2 std=16.21; online: mean=42.4 std=18.4) were higher than those in the P3Speller task (offline: mean=26 std=17.6; online: mean=33.1 std=21.7) we didn't find any significant difference between them both in the offline condition (p=0.19) both in online condition (p=0.4).

5 CONCLUSIONS AND FUTURE DEVELOPMENTS

The eye tracker systems as communicative means represent the ideal solution for the ALS subjects that are able to move the eyes, compared to the BCI P300-based text writers, because the detection of eye movements is quicker, easier, and more accurate than the detection of ERPs. A BCI system operable without the necessity to move the eyes is the only way to communicate for the ALS subjects, completely "locked-in."

In this study was shown a new P300-based BCI system, useable in covert attention status. The performances using the GeoSpell interface (> 70%), allow defining it as "Communicative Mean".

In a future study, it will be tried to bring some changes to the GeoSpell interface, that allow to improve the usability and accuracy, giving particular relief to the training, that could improve the performances.

ACKNOWLEDGEMENTS

This work is partly supported by the EU grant FP7-224332 "SM4ALL" project, and FP7-224631 "TOBI" project. This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

REFERENCES

Allison, B. Z. & Pineda, J. A., 2003. ERPs evoked by different matrix sizes: implications for a brain computer interface (BCI) system. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2), pp.110-113.

Anderson, J., 1999. *Cognitive Psychology and its Implications* Fifth Edition., Worth Publishers.

Brunner, P. et al., 2010. Does the 'P300' speller depend on eye-gaze? *Journal of Neural Engineering*, 7(5), p.056013.

Fabiani, M. et al., 1987. Definition, identification and reliability of measurement of the P300 component of the event-related brain potential. , 2, pp.1-78.

Farwell, L. A. & Donchin, E., 1988. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, 70(6), pp.510-523.

Hart, S. G. & Staveland, L. E., 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*. North-Holland, pp. 139-183.

Krusienski, D. J. et al., 2006. A comparison of classification techniques for the P300 Speller. *Journal of Neural Engineering*, 3(4), pp.299-305.

Kübler, A. et al., 2006. BCI Meeting 2005--workshop on clinical issues and applications. IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the *IEEE Engineering in Medicine and Biology Society*, 14(2), pp.131-134.

Lutz, A. et al., 2009. Mental training enhances attentional stability: neural and behavioral evidence. The Journal of Neuroscience: *The Official Journal of the Society for Neuroscience*, 29(42), pp.13418-13427.

Polich, J. & Kok, A., 1995. Cognitive and biological determinants of P300: an integrative review. *Biological Psychology*, 41(2), pp.103-146.

Raymond, J. E., Shapiro, K. L. & Arnell, K. M., 1992. Temporary suppression of visual processing in an RSVP task: an attentional blink? *Journal of Experimental Psychology. Human Perception and Performance*, 18(3), pp.849-860.

San Agustin, J. et al., 2010. Evaluation of a low-cost open-source gaze tracker. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications - ETRA '10*. the 2010 Symposium.Austin, Texas, p. 77.

Schalk, G. et al., 2004. BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Transactions on Bio-Medical Engineering*, 51(6), pp.1034-1043.

Sutter, E., 1992.The brain response interface: communication through visually-induced electrical brain responses. *J. Microcomput.* Appl., 15(1), pp.31-45.

Treder, M. S. & Blankertz, B., 2010. (C)overt attention and visual speller design in an ERP-based brain-computer interface. *Behavioral and Brain Functions*: BBF, 6, p.28.

Wolpaw, J. R. et al., 2002. Brain-computer interfaces for communication and control. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 113(6), pp.767-91.