# Exploring motion VEPs for gaze-independent communication

Sulamith Schaeff, Matthias Sebastian Treder, Bastian Venthur & Benjamin Blankertz*

April 25, 2012

Department of Neurotechnology, Berlin Institute of Technology, Germany
Email: benjamin.blankertz@tu-berlin.de

## Abstract

Motion visually evoked potentials (mVEPs) have recently been explored as input features for Brain-Computer Interfaces (BCIs), in particular for the implementation of visual spellers. Due to low contrast and luminance requirements, motion-based intensification is less discomforting to the user than conventional approaches. So far, mVEP spellers were operated in overt attention mode, wherein eye movements were allowed. However, dependence on eye movements limits clinical applicability. Hence, the purpose of this study was to evaluate the suitability of mVEPs for gaze-independent communication. Sixteen healthy volunteers participated in an online study. We used a conventional speller layout wherein the possible selections are presented at different spatial locations both in overt attention mode (fixation of the target) and covert attention mode (central fixation). Additionally, we tested an alternative speller layout wherein all stimuli are sequentially presented at the same spatial location (foveal stimulation), that is, eye movements are not required for selection. As can be expected, classification performance breaks down when switching from overt to covert operation. Despite reduced performance in the covert setting, conventional mVEP spellers are still potentially useful for users with severely impaired eye movements. In particular, they may offer advantages – such as less visual fatigue – over spellers using flashing stimuli. Importantly, the novel mVEP speller presented here recovers good performance in a gaze-independent setting by resorting to foveal stimulation.

# 1 Introduction

Brain- Computer Interfaces (BCIs) provide the user with a communication system that circumvents the brain's normal output pathways of peripheral nerves and muscles and directly translates brain signals into commands to control external devices [1]. BCIs can serve as new output pathways of the brain [2]. One field of BCI applications lies in the development and use of mental typewriters that enable users to spell out letters on a computer screen directly with their neuronal signals [3]. Commonly, speller devices are based on the fact that attention modulates neural processing: Event-related potentials (ERPs) of (attended) target stimuli are different from (unattended) nontarget ERPs. Thus, by assigning the ERP signal input to target/nontarget classes, the users' intentions can be deduced which makes it a suitable paradigm for BCI [4, 5]. ERPs are electrical potentials of the brain that have a constant time relationship to a certain reference event [6]. Most important in the context of BCI spellers are the N200 [7] and the P300 component [8]. The N200 component is evoked in the visual cortex by visual stimuli in the foveal field, and its amplitude is modulated by attention. The P300 has been found to be elicited by task-relevant stimuli. Its amplitude is inversely related to target-to-target interval [9] and it is largest over central and parietal electrode sites [6, 10]. Farwell and Donchin [4] were the first to demonstrate that attention-modulated ERPs serve as a reliable signal for BCI control. They described a Matrix Speller (a.k.a. P300 speller) in which participants focused their attention sequentially on the target characters within a symbol matrix on a screen. Within the sequence of randomly highlighted rows and columns, the ERPs corresponding to the highlighting of the row or the column containing the target symbol are distinguished by an increased amplitude. Thereby, a classifier can identify the symbol that the user is attending. ERP spellers can be of high clinical utility as it was demonstrated to restore communication in patients with severe motor disabilities [11].

It has been shown that the good performances of Matrix Spellers, where symbols are arranged in rows and columns, are affected by eye movements [12, 13, 5]. More precisely, they are only highly efficient in overt attention paradigms where participants are allowed/able to direct their gaze towards the locus of attention. In covert attention mode, where participants have to mentally focus on the stimuli while fixating may be somewhere else, classification performance deteriorates. This is due to a7n attenuation of the P300 as well as to the absence of modulated primary visual components. The attenuation was presumably provoked by low spatial acuity and crowding effects in the visual periphery which hampers peripheral covert attention. Growing receptive field sizes and decreasing cortical magnification with increasing eccentricities have long been recognized to reduce visual acuity in peripheral regions [14, 15]. The dependence of Matrix Spellers on gaze discloses an important limitation; it requires a certain degree of intact oculomotor control. Therefore it appears crucial to develop paradigms that are gaze independent. In the visual domain there exists a hierarchy of typewriters in order to optimally exploit the patient's capabilities. One end of the hierarchy relies on optimal utilization of gaze e.g. electrooculogram (EOG) [16] and code modulation [17]. Whereas the other end is gaze-independent, e.g., rapid serial visual presentation (RSVP) [18], and variants of the Hex-o-Spell [19]. An alternative is the use of other sensory modalities such as auditory [20, 21] and tactile paradigms [22, 23, 24].

Recently the feasibility of an online BCI based on motion-onset visual evoked potentials (mVEP) has been demonstrated. A moving cursor that appeared in virtual buttons generated a mVEP that was then used to recognize the user's choice [25]. MVEPs are visual responses from the dorsal stream; hence they rise in the MT/MST (medial temporal/medial superior temporal cortex) region [26, 25]. They are generated in response to the onset of motion stimuli [26]. MVEPs consist of three main parts: P100, N200 and P200 [27]. The most reliable of these is the N200 [7, 27]. It is an negative deflection around 160–200ms post-stimulus and seems to be generated in the region of the temporo-occipito-parietal junction (V5) [28, 29]. MVEPs are advantageous for BCI in that they have large amplitudes, low inter- and intra- subject variability [30], low luminance and contrast requirements [27, 26] and moreover a localized spatial distribution that allows for a sparse EEG channel configuration [25]. The use of mVEPs in BCI paradigms might facilitate the construction of more user-friendly systems since devices devoid of abrupt luminance changes and high contrasts should lead to less fatigue and discomfort of the user. Besides, the reduced variability of mVEPs could improve BCI systems, since inter- and intra- subject variability poses a serious problem for classification [2].

In the present study, an mVEP Speller in an online covert/overt attention paradigm is introduced in order to develop a gaze-independent BCI speller. MVEP spellers are among the most promising approaches to solve the problem of gaze dependency. It has not been investigated to what degree mVEP spellers depend on gaze and if they can be operated in covert attention paradigms. In line with overt attention, covert attention has long been known to alter neuronal responses and improve behavioral performance i.e facilitate detection and discrimination [31]. [32] define it as 'orienting without eye-movements' (p.63) and it has been found to modulate N200 and P300 responses [25, 5]. Top-down control processes on bottom-up sensory processes have been proposed to explain the modified brain activity [32]. Using the Cake Speller [19], a further development of the Hex-o-Spell [5], in combination with a motion stimulus the viability of an online mVEP-governed speller in covert attention modes was tested. To this end three different spellers were developed: Overt Cake, Covert Cake and Motion Center Speller. The Overt Cake Speller was designed to reproduce viability of N200 as input signal [25] in a Cake Speller design. The Covert Cake Speller was employed in order to test gaze dependency of this setting. The Motion Center Speller was adopted as a consequence of [19] where it proved to be best among three spellers that did not rely on eye movements. The configuration of the three spellers was similar, they differed only in the attention mode (covert/overt) and motion stimulus (moving bar/pattern). By using a speller design composed of two levels, i.e. less small elements, problems occurring through reduced visual acuity and spatial crowding in the periphery were alleviated. Thus, reduction or elimination of eye-movements was feasible.

All three spellers were made up of a hexagon with six equal cake-shaped parts. In the Overt and Covert Cake Spellers motion stimulation was foveal: small bars moved within the chunks. In the Motion Center Speller, motion stimulation was entirely foveal elegantly solving the problem of covert attention deployment in the periphery. To further increase speller performance, feature attention was conveyed besides spatial attention. In contrast to spatial attention, non-spatial or feature attention is associated with color, orientation, shape or spatial frequency [33]. In this design it was encouraged by different colors of the motion stimuli. Robustness was further strengthened by combining mVEPs and the P300 as input features in this oddball paradigm. Furthermore, the favorable characteristics of mVEP as described above were expected to reduce ERP variability and thereby assist in creating a robust BCI.

# 2 Methods

## 2.1 Participants

Sixteen volunteers participated in this study. Participants (10 males and 6 females) were aged between 21 and 30 with a mean age of 23.8 years. They had normal or corrected-to-normal vision. Normal color vision in all but two participants (*iac, ibu*) was confirmed using the Nishihara Color Vision Test. Prior to the experiment written consent was obtained from each participant. The study was in accordance with the Declaration of Helsinki.

## 2.2 Apparatus

EEG recordings were done using a Brain Products (Munich, Germany) actiCap active electrode system with 64 electrodes (Fp1,2, AF3,4,7,8, Fz, F1-10, FCz, FC1-6, FT7,8, T7,8, Cz, C1-6, TP7,8, CPz, CP1-6, Pz, P1-10, POz, PO3,4,7,8, Oz and O1,2) and a BrainAmp EEG amplifier. Electrode were placed according to the international 10-10 system. Right mastoid was chosen as a reference site and a forehead ground electrode. For offline analysis electrodes were re-referenced to linked mastoids. Impedances were kept below 10 k$\Omega$. EEG data were sampled at a rate of 1000Hz and hardware bandpass filtered between 0.016–250Hz. For control of eye movements an Intelligaze IG-30 (Alea Technologies) eyetracker, sampling at 50Hz, was used simultaneously. The control system ceased and restarted a trial immediately whenever the eyes were not on the fixation point. Due to technical problems of the eyetracker, it was only used for seven participants (*ibq, iac, gdf, icv, ibe, gdg, iba*). In any case participants were instructed to strictly remain on the fixation points, this was particularly emphasized for the participants that were not controlled by the eyetracker. Stimuli were presented on a 19" TFT screen with a refresh rate of 60Hz and a resolution of 1280×1024 px$^2$. Participants were seated at 60cm distance from the screen. Furthermore a photodiode was attached to lower left corner of the screen to register stimulus onset (g.trigBox by Guger Technologies) and allow for subsequent correction of TFT delay. Stimulus presentation was synchronized with screen refresh.

## 2.3 Stimuli and Speller Design

Three different Cake Speller modifications were employed in a within-subject design: an Overt Cake Speller, a Covert Cake Speller, and a Motion Center Speller. The basic design and operating mode is the same for all three spellers. It comprises a hexagon split up in six parts. Each piece contains letters or additional characters. In total the participant can choose from 30 different symbols ('ABCDE', 'FGHIJ', 'KLMNO', 'PQRST', 'UVWXY' and 'Z_ . , <'). The speller device is composed of two levels. In the first level (group-level) each chunk features five symbols (e.g. 'ABCDE'). Whereas in the second level (symbol-level) each piece contains only the corresponding single letter (e.g. 'A'). Accordingly, letter selection is a two-stage process. First the desired letter group is selected. Then the speller moves to the second level. In the second level participants choose within this group the single desired letter. As a feedback the detected character appears in grey letters at the top of the screen. The individual letter groups, the single letters and the corresponding motion stimuli are highlighted with a unique color in order to facilitate recognition and enhance attention to the target. So, the speller nurtures two different types of attention. First, spatial attention – the participant can attend to a specific location in space. Secondly feature attention – participants can focus on the color of the stimulus. This facilitates allocation of (covert) attention. Colors of the stimuli include red, green, blue, yellow, purple and turquoise. Corrections of mistakes are implemented by a backspace symbol ('<')in the group-level and an empty disc in the symbol-level, causing a return to the corresponding previous level. Mainly two ERP components are objectives of this stimulus design, P300 and N200. Target presentation among six selectable stimuli constitutes an oddball event thereby eliciting a P300. Induction of N200 responses is implemented differently for the spellers by means of two different visual settings. A moving bar or a moving grid pattern are adopted in order to produce N200 components. To be able to directly investigate gaze (in)dependency, attention is manipulated differently for the spellers (covert/overt). Spellers are illustrated in Figure 1.

**Overt cake speller**. Participants are meant to overtly attend, i.e. shift their attention and gaze towards the piece of cake they are opting for and meanwhile ignoring the others. Within each chunk there is a fixation point in the middle to guard their gaze. In order to increase signal strength, direct attention and most notably keeping it on their destination, participants were asked to silently count the number of recurrent movements of a small bar within their designated piece. This bar moves within the different parts of the hexagon pseudorandomly alternating between the single pieces and eliciting the ERPs in question. It changes color according to the panel it moves in. The participant is instructed to count with her/his inner voice whenever the bar moves in the matching color within the target until the system displays the letter it had selected. Thus, spatial location and color of the bar serve as a hint.

**Covert cake speller**. In the Covert Cake Speller eyes have to strictly fixate a fixation point in the middle of the hexagon. At the meantime, the participant is mentally directing his attention away from the center towards the designated letter respectively the moving bar. Apart from this, its modus operandi equals the one described above.

**Motion Center Speller**. Here a moving pattern in the middle of the hexagon is meant to generate a mVEP. The grid pattern consisted of arrowheads pointing alternately to one of the pieces. Each arrowhead has a unique color corresponding to the color of the respective letter group. Participants have to fixate on a central fixation point in the middle of the moving pattern.

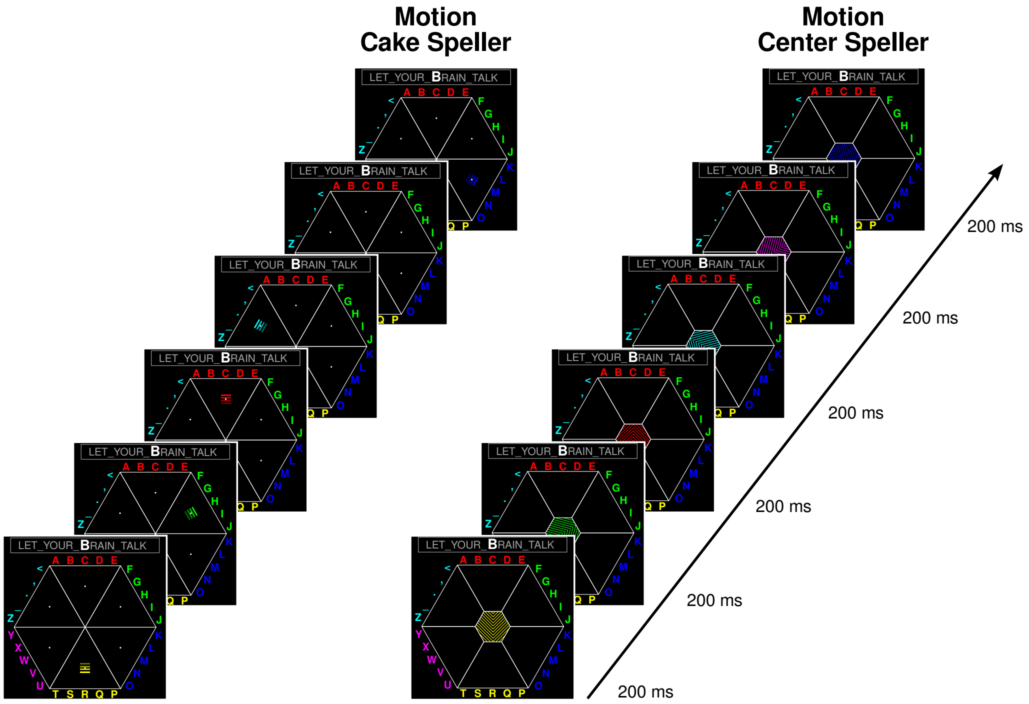

**Figure 1: Left:** Motion Cake Speller: In overt mode, participants had to fixate the fixation point in a hexagon. In covert mode, participants had to fixate a central fixation dot (not shown). In both cases, they have to attend to a moving bar in the respective hexagon. **Right:** Participants focuse on the central grids of arrowheads. Each grid points to one of the six possible selections. The color of the arrowheads corresponds with the color of the respective groups. Participants have to focus on the aorrwhead pointing onto the group that they intend to choose.

The participants' task is to direct their attention towards the moving pattern and silently count whenever the arrow points towards their target piece having the appropriate color. In this design motion stimulation was entirely foveal.

All spellers had a total radius of 300px (8.5°). Bar width in both Cake Spellers was 5px (0.13°) and the bar moved with a speed of 0.51°/frame. During the duration of a stimulus the bar moved a total of 2.55°, starting at an eccentricity of 2.98° and ending at an eccentricity of 5.53°. The grid pattern in the Motion Center Speller was composed of 12 stripes as a whole subtending 64px (3.06°) and it moved 0.43° during the duration of one stimulus.

The moving bar/pattern is presented 2 [levels] × 10 [repetitions] × 6 [groups/symbols] = 120 times for each selection. The order is semi-randomized with at least two intermittent intensifications before a certain group is intensified again. The duration of a single movement is 100ms with 100ms between movements (inter-stimulus duration). Therefore total stimulus-onset-asynchrony (SOA) amounts to 200ms. The duration of one sequence (one stimulus repetition) adds up to 1.2sec. Total trial duration is about 30 seconds. Due to a technical problem, the SOA for the Motion Center Speller was 266ms rather than 200ms.

All spellers are implemented in the open-source framework Pyff [34] using VisionEgg [35] and remote controlled via MATLAB (The MathWorks, Natick, MA, USA).

## 2.4 Procedure

Participants received verbal and written instructions about the procedure and the purpose of the experiment. In order to reduce artifacts they were instructed to sit still, relax their muscles and avoid eye blinks/movements. The order of the three spellers was counterbalanced across participants. Each speller traversed through the same four phases: a practice phase, a calibration phase and two online phases. In a short *Practice* phase, participants imitated the writing process to become familiar with the setting. During *Calibration* participants copy-spelled three default words ('WINKT_QUARZ_FJORD') and their EEG was recorded. Participants were instructed to attend to the respective letter they had to select. They received no online feedback. Data were subsequently used as training set for the classifier. In the *Copy Spelling* run participants had to write a given sentence online ('LET_YOUR_BRAIN_TALK'). They received online feedback of the selected symbols but they did not have to correct wrong selections. During *Free Spelling* participants spelled a sentence (roughly 20 characters) that they invented in the previous break. Here, erroneously selected symbols had to be erased using the backspace symbol as explained above (and erroneously erased symbols had to be reselected).

### 2.5 Data analysis

For offline ERP analysis the data were downsampled to 200Hz and lowpass filtered using a Chebyshev filter with a 42Hz passband and a 49Hz stopband. The continuous EEG was epoched with epochs spanning 200ms pre-stimulus to 1000ms post-stimulus. Baseline correction was done on the 200ms pre-stimulus interval. Physiological artifacts (e.g. muscular activity, eye movements) were rejected using a min-max criterion (difference min-max voltage $\geq 70\mu V$) and a variance criterion. In stimulus designs with short SOAs, ERPs of successive stimuli often affect each other. Response durations overlap in time and thereby distort the ERP under investigation. To reduce the effect of targets on nontarget epochs, only those nontarget epochs were considered whose 3 preceding and 4 following stimuli were also nontargets. Statistical significance was tested using a repeated measures ANOVA.

For online and offline classification, the data were downsampled to 100Hz. All nontarget trials were included and no artifact rejection was performed. A binary linear classifier using linear discriminant analysis (LDA) with shrinkage of covariance matrix was implemented [36]. A separate classifier was trained for each of the spellers with data from the corresponding calibration phase and subsequently tested on the free and copy spelling phases. As spatial features all electrodes except for frontal electrodes close to the eyes (i.e., *Fp1,2* and *AF3,4,7,8*) were included. The discriminability index for target/nontarget classes was based on signed square values of point-biserial correlation coefficients ($sign\, r^2$). A heuristic search was used to automatically select temporal features, i.e. optimal peaks in the $sign\, r^2$ values of targets/nontargets in the 100–800ms post-stimulus interval. LDA acts as spatial filter, that enhances the target signal and suppresses non-discriminative sources (Blankertz et al., 2011). For online classification, timing and number of classification windows could be manually adjusted by the experimenter, with an initial default of 5 temporal intervals. This yielded a feature vector with 58 spatial features times typically 5 temporal features, hence a total of 290 spatio-temporal features.

## 3 Results

### 3.1 Event-Related Potentials

Grand average ERPs for the three spellers are illustrated in Figure 2. To increase statistical power we restricted the analyses to a subset of electrodes. As indicated in Figure 2, N200 was most distinct over parieto-occipital sites ('P7' 'P3' 'PO7'), whereas P300 had a central focus ('FCz' 'Cz' 'Pz'). Statistical analysis was done using a three-way repeated-measures analysis of variance (RM-ANOVA) on *Speller* (Overt Cake/Covert Cake/Motion Center Speller) x *Status* (target/nontarget) x *Electrode* (N200: 'P7' 'P3' 'PO7'; P300: 'FCz' 'Cz' 'Pz'). Peak amplitudes and latencies were determined within the 100–250ms (N200) respectively 300–500ms (P300) post-stimulus interval.

**N200** N200 were most pronounced in the Overt Cake Speller: mean amplitude over the three electrodes was $-2.58\mu V$ for target presentations and $-.63\mu V$ for nontarget presentations. The Covert Cake Speller had mean voltages of $-1.57\mu V$ for targets and $-0.95\mu V$ for nontargets. Overall

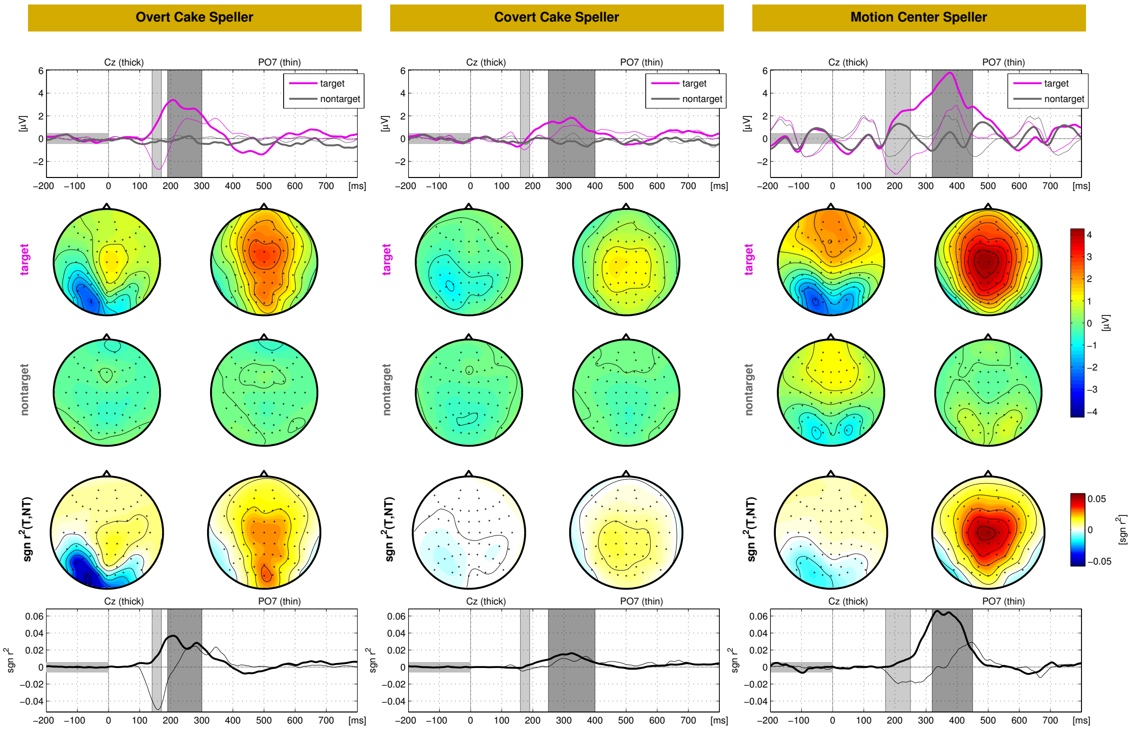

**Figure 2:** Grand average ERPs for Motion Center, Covert Cake and Overt Cake Speller. **Top:** Event-related voltage changes in $\mu V$ plotted against time in ms following the presentation of the stimulus. Target responses are depicted in magenta, nontargets in grey. Thick lines represent average responses at electrode Cz and thin lines at PO7. **Middle:** Scalp topographies were generated by averaging the voltages in the two shaded intervals shown in the graphs. The first two rows depict the response to targets and nontargets, the third row gives the $sign\, r^2$ difference between targets and nontargets. **Bottom:** $sign\, r^2$ as a function of time for electrodes Cz and PO7.

amplitude was highest in the Motion Center Speller setting, with means of $-3.15\mu V$ for targets and $-2.17\mu V$ for nontargets. Mean N200 amplitudes for factor *Speller*: Covert Cake ($-1.26\mu V$), Overt Cake ($-1.60\mu V$) and Motion Center Speller ($-2.77\mu V$). Statistical analysis of N200 amplitude showed a significant effect of *Speller* ($F = 35.23, p < 0.001$), *Status* ($F = 69.24, p < 0.001$) and *Electrode* ($F = 4.73, p < 0.01$). The two-way interactions of *Speller* x *Status* ($F = 7.85, p < 0.001$) and *Speller* x *Electrode* ($F = 2.6, p < 0.05$) were significant, but *Status* x *Electrode* ($F = 2.16, p = 0.12$) and the three-way interaction ($F = 1.35, p = 0.25$) were not significant.

Tests on N200 latency revealed a significant effect of *Speller* ($F = 34.51, p < 0.001$) and *Electrode* ($F = 4.22, p < 0.05$). *Status* ($F = 0.08, p = 0.77$) and all interactions showed no significant effects. Mean N200 latencies averaged across electrodes and status were 180ms (Covert Cake), 164ms (Overt Cake) and 198ms (Motion Center Speller).

**P300** Mean P300 amplitudes (over the subset of electrodes) of the Overt Speller were $1.31\mu V$ for targets and $0.66\mu V$ for nontargets. In the Covert Speller amplitudes were $2.21\mu V$ for targets and $0.93\mu V$ for nontargets. Foveal stimulation was most successful in modulating the P300, the Motion Center Speller showed the biggest difference for stimulus status: $6.21\mu V$ for target and $2.68\mu V$ for nontarget potentials. A repeated-measures ANOVA performed on P300 amplitude indicated a significant effect of *Speller* ($F = 186.63, p < 0.001$) and *Status* ($F = 134.95, p < 0.001$) but not of *Electrode* ($F = 0.17, p = 0.842$). Two-way interaction of *Speller* x *Status* was significant ($F = 31.36, p < 0.001$). The other interactions were not significant.

P300 latency tests revealed a significant effect of *Status* ($F = 29.39, p < 0.001$). No other significant effects could be observed. Mean P300 latencies averaged across electrodes and status were 380ms (Covert Cake), 386ms (Overt Cake) and 394ms (Motion Center Speller).

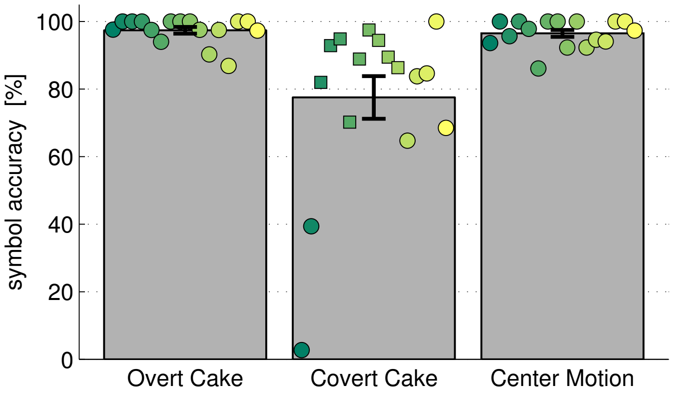

**Figure 3:** Online spelling accuracies for the three spellers, given for each of the participants (color coded). The mean online spelling accuracies are shown spellerwise by the grey bars and the standard error by the whiskers. Accuracy is given as percentage correctly selected symbols (chance $1/30 \approx 3.33\%$).

## 3.2 Classification

Online spelling accuracy, given as percentage correct selections, was $97.4\% \pm 0.8\%$ standard error (SE) for the Overt Cake Speller. For the Covert Cake Speller, selection accuracy dropped to $76.74\% \pm 6.8\%$ SE selection accuracy. Motion Center Speller reached $96.2\% \pm 1.2\%$ SE spelling accuracy (see Figure 3). The two participants (*iac*, *ibu*) with impaired color vision yielded results comparable to the remaining subjects and were therefore included in all analyses. *Effective* spelling speed (i.e., taking into account the time needed to make corrections) with ten sequence repetitions was $1.6 \pm 0.02$ symbols/minute (Overt Cake Speller), $0.97 \pm 0.22$ (Covert Cake Speller), and $1.28 \pm 0.03$ (Motion Center Speller).

Chance level for selecting a single letter out of 30 is 3.33%. A one-way repeated measures ANOVA with factor *Speller* yielded a significant effect of speller ($F = 8.63, p < 0.01$). Tukey-Kramer post-hoc tests showed that the performance of the Covert Cake Speller is significantly lower

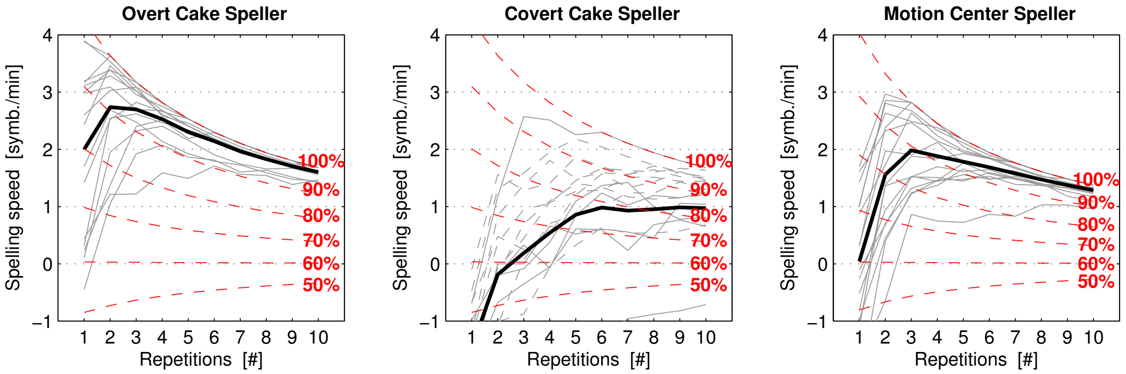

**Figure 4:** Spelling speed in symbols/minute for each of the three spellers plotted against number of repetitions. Thin grey lines depict results for single participants and the solid black line depicts the mean. Red dashed lines represent the spelling speed for fixed levels of symbol selection accuracy. Spelling accuracy for the empirical data (solid black line) can be deduced by comparing the black solid line to the red dashed lines. Note that the accuracy and spelling for 10 sequences corresponds to the data from the online experiment. For the Covert Cake Speller, participants using the eyetracker are depicted as solid lines, uncontrolled participants as dashed lines.

than the performance of the other two spellers, with no significant difference between Overt Cake Speller and Motion Center Speller. For the Covert Cake Speller, a t-test for the two groups (eyetracker/ no eyetracker) revealed that the participants using an eyetracker performed significantly worse than the other participants ($t = 2.29; p < 0.05$). Mean accuracy for eyetracker-controlled participants was $61.15\% \pm 8.71\%$ versus $88.87\% \pm 2.55\%$ for non-controlled participants. Possibly, this discrepancy is due to participants making unintentional saccades in the direction of the target, thereby easing the deployment of spatial attention.

We also performed offline simulations wherein we deteremined the spelling speed (symbols/minute) as a function of the number of sequences. The results are shown in Figure 4 for each speller separately. A two-way repeated measures ANOVA with factors *Speller* and *Sequence* was conducted to investigate increases in accuracy with number of intensification sequences. There was a significant effect of the factors *Speller* ($F = 247.77, p < 0.001$) and *Sequence* ($F = 19.73, p < 0.001$) and also the interaction was significant ($F = 9.22, p < 0.001$). Tukey-Kramer post-hoc tests revealed that, in terms of spelling speed the following relation holds: Overt Cake Speller $>$ Motion Center Speller $>$ Covert Cake Speller.

To investigate spatial distributions of obtained classification data and advise possible future electrode configurations, offline analyses were re-run for each electrode separately. Each sample in the 100–800ms post-stimulus interval served as an individual feature. The results are shown in Figure 5. For the Overt Cake Speller, peak accuracies of 50% are obtained over occipital electrode sites, suggesting that classification success is based mainly on visual and visual-attentional components. The selection accuracy for the Covert Cake Speller is comparably low (25% - 30%) for almost all electrode sites with left-parietal sites performing slightly better. For the Center Speller, there are performance peaks over fronto-central and left-occipital electrode sites, suggesting that both the P300 component and visual-attentional components are discriminative. To investigate whether the left-hemispheric dominance in terms of classification performance is significant, we conducted a two-way RM-ANOVA with factors *Speller* and *Hemisphere*. There were significant effects for *Speller* ($F = 39.65, p < 0.001$) and *Hemisphere* ($F = 12.56, p < 0.001$), but there was no significant interaction ($p = .5$).

To investigate which time intervals contribute most to classification success, we considered a time window of 30ms wherein voltages were averaged. Classification was repeated for different positions of the time window, in order to track the distribution of discriminative information across time. The results, depicted in Figure 6, are consistent with the picture of the ERP analysis, which advocates that it was mainly classified on the intended signals.

Accuracy of the Overt Speller rises sharply around 165ms, which corresponds to mean peak

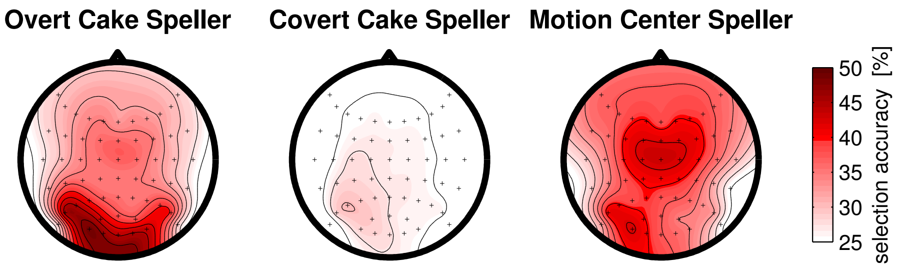

**Figure 5:** Spatial distribution of classification performance. To this end classification was done for each electrode at a time, using sample times as temporal features. Grand averages for each of the spellers are shown and percentage loss is indicated by a color gradient. Black lines delineate regions of high accuracies from surrounding lower values.

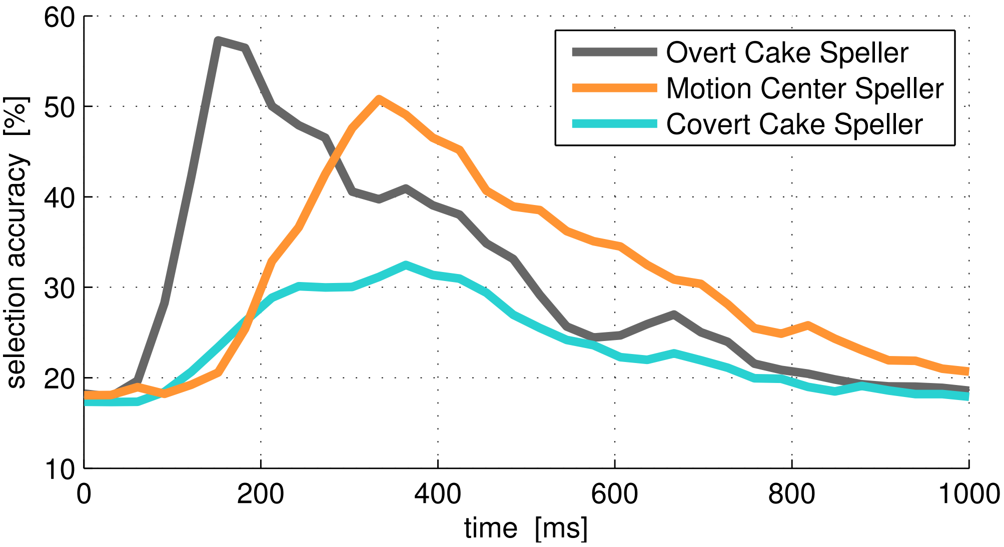

**Figure 6:** Temporal classification on average voltage for a running window width of 30ms. Accuracy is given as percentage correct.

latencies detected for N200 components as described before. It peaks around 220ms and decays until 800ms. This suggests early incorporation of the N200. P300 components were not shown to be distinctive but other VEPs were likely to be accumulated later for highly successful classification. N200 induction rendered peak classification accuracies. Temporal performance distribution of the Covert Cake Speller does not show a clear peak, it is rather a small plateau that rises around 220ms and pertains until 600ms. This is in accordance with the ERP analysis, N200 responses are only informative to some extent and classification was mainly due to modulations of the P300. Compared to the Overt Speller, accuracy of the Motion Center Speller develops later: it starts around 200ms and peaks about 400ms. Which is conform with the observed higher latency of N200 (190ms) in this setting. Accuracy peak is in accordance with peak time of the P300, nevertheless it is not clear-cut and unimodal but more broadly distributed. Possibly a smearing of the effects of N200 and P300 responses leads to this relatively slow accumulation that leaks out late (800ms).

### 3.3 Exogenous effects

To investigate stimulus-specific effects, target ERPs in each individual stimulus-class were compared with all nontargets. Stimulus-specific effects for both P300 and N200 could be observed for the Covert Cake Speller but not for the other two spellers. Results for N200 components are depicted in Figure 7. Exogenous effects on ERPs were also reflected in selection accuracy. A one-way repeated measures ANOVA with factor *Symbol* was run for each of the spellers. It proved that there were significant differences in selection accuracy regarding the different stimuli in the Covert Cake Speller ($F = 5.82, p < 0.001$). No differences could be observed in the Overt Cake Speller ($F = 0.82, p = 0.54$), and Motion Center Speller ($F = 2.31, p = 0.052$).

The fact that the target/nontarget patterns differ depending on the spatial location for the Covert Cake Speller seems to suggest that a single binary classifier may not be optimal. To test this, we trained a separate classifier for each spatial location. However, performance was not better than for a single classifier (results not shown).

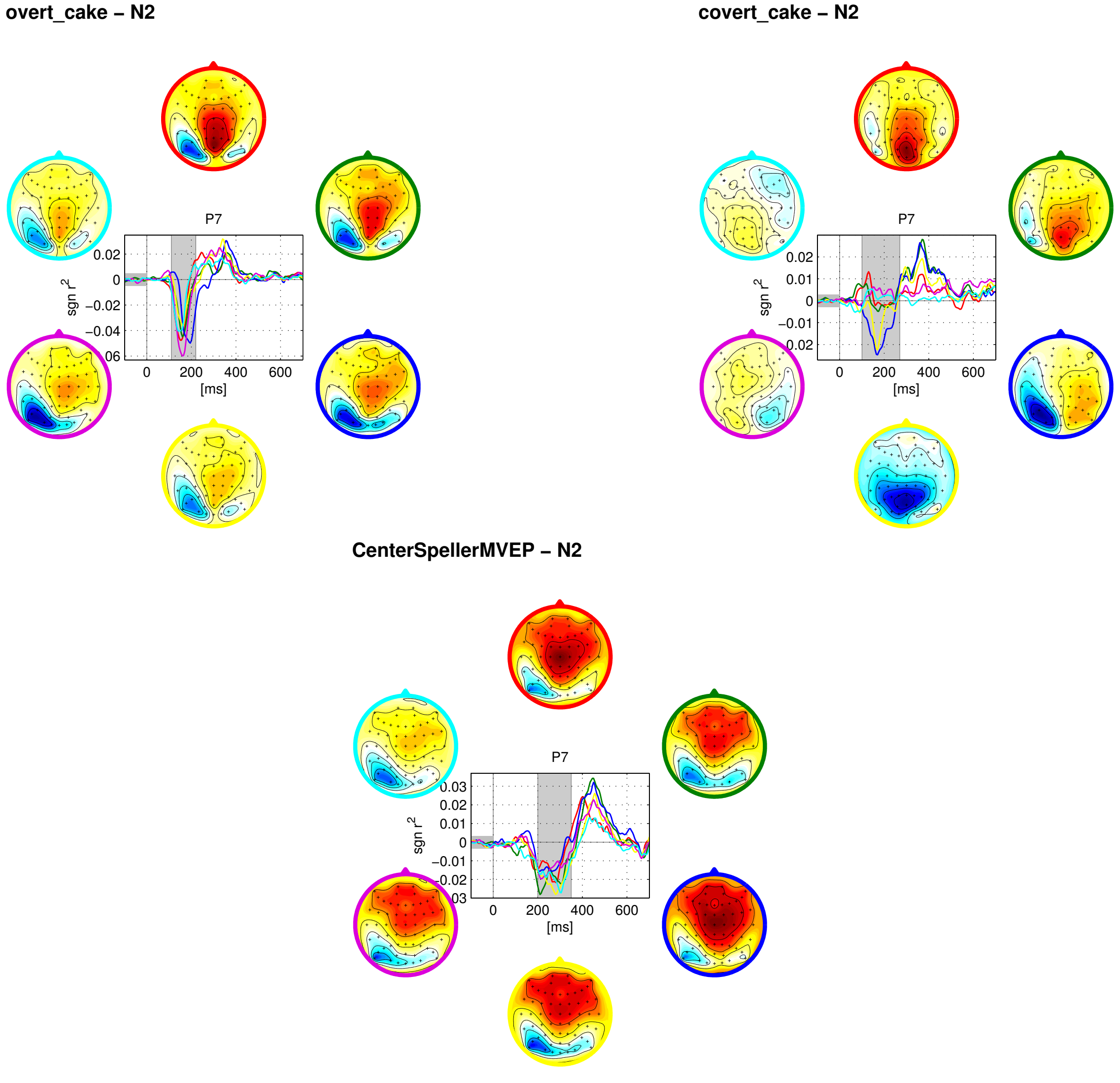

**Figure 7:** Average N200 waves for each of the six stimulus classes for Overt Cake, Covert Cake and Motion Center Speller. The graphs illustrate average time courses (x-axis) of *sign* $r^2$ values (y-axis) for each stimulus class at electrode P7. Colored lines indicate the six stimulus classes. Scalp topographies the depict the spatial distribution of *sign* $r^2$ values with the intervals corresponding to the shaded areas in the graph.

## 4 Discussion

The purpose of this study was to evaluate whether mVEPs can serve as basis for gaze-independent communication. Three different designs have been under investigation in order to compare overt versus covert attention and central versus peripheral stimulation: Overt Cake Speller, Covert Cake Speller and Motion Center Speller. All participants were able to successfully online control the BCI speller and spelling speed was equivalent to about 1.5 characters per minute. Online accuracy results were high for all of the spellers: 97.4% for the Overt Cake Speller, 76.7% in the Covert Cake Speller setting and the Motion Center Speller reached 96.2%. The accuracy rate of the Motion Center Speller, though entirely gaze-independent, is up to the standard of the Overt Cake Speller. However, information throughput is slightly higher in the Overt Cake Speller due to its shorter SOA (200ms compared to 266ms). Stimulus specific effects were only observed in the Covert Cake Speller and are presumably due to peripheral vision and attention constraints.

Overt attention results confirm previous findings [25]. Peripheral covert attention remains a challenge, though allocation of attention was supported in many ways. Control of the Covert Cake speller was remarkably harder than control of the other two spellers. Partly, this may have been caused by the dual-task requirements of the task (i.e., fixating the center, attending to the periphery). However, classification loss was mainly due to the attenuation of the N200 component that has been shown to be largely dependent on eye gaze [13]. Overall performance of the Motion Center Speller was high, allocation of attention was easy and (almost) entirely feature-bound, eye movements were not at all necessary since stimulation occurred in foveal regions. Concluding it can be said that overt attention designs are more effective than covert ones and foveal stimulation is superior to peripheral stimulation if covert attention is a prerequisite. As a consequence in mVEP-based online spellers visual stimulation has to be foveal if eye movements are not permitted.

The Overt Cake Speller showed mainly modulation of N200 responses and classification was accordingly for the most part based on this early component. The topography of the component was asymmetric with a clear left-hemisphere dominance. This is in line with previous BCI studies that showed a left-hemisphere dominance in N200 amplitude [37, 30]. The endogenous P300 responses to targets and nontargets were not separable. The opposite is true for the Covert Speller. P300 was larger for targets and classification was almost entirely based on it, whereas N200 target modulations were small. It is likely that this is attributable to the difficulties related to allocating attention to the visual periphery [5] and visual eccentricity of the target reducing response amplitudes of early visual responses, such as mVEPs. Increasing neural receptive field sizes and decreasing cortical magnification make peripheral vision less accurate. It has been shown that spatial-frequency selectivity changes according to retinal eccentricity, smaller stimuli are finer resolved in foveal regions [15]. In the Motion Center Speller both the attention-based P300 and the N200 were modulated for target responses and classification was accordingly based on both components. Evidently, attention allocation was feasible and task difficulty moderate. The effectiveness of the Motion Center Speller setting might be attributable to several factors. First of all, it was centered in the fovea, the area with highest resolution and cortical magnification. Second, the motion stimulus was larger in size, subtending more visual space and therefore activating larger neuronal clusters. And third, attention was necessary and effective. It remains subject to discussion, however, whether the Motion Center Speller can be regarded as a covert attention paradigm. Orienting in this setting does not require eye movements and for an external observer the locus of attention is not obvious, hence covert. This represents the line of reasoning followed in this paper. But nevertheless the attentional spot and stimulus are both centrally oriented, covering the same space. If covert attention is considered to be spatial by nature, the Motion Center Speller has to be understood as an overt paradigm.

To sum up, when comparing covert and overt attention spellers, it is evident that both P300 and N200 components are susceptible to attentional enhancement. Nevertheless, there are differences: if stimuli are peripheral, hard to distinguish and eye movements are not desired, the P300 suits better for classification. On the other hand, if stimuli are centrally presented and good to perceive N200 suits best. It provides an excellent option for successful BCI control. The question remains to what extent covert attention paradigms with peripheral motion stimulation can be improved e.g. by increasing stimulus size. A comparison of foveal and peripheral stimulation identifies central stimulation as favorable in all respects. This is also in accordance with a different speller design that uses foveal stimulation but usual, non motion-onset VEPs [19]. The performance reported in [19] is comparable to the results presented in this paper. When exploring the fully peripheral stimulation in the Covert Speller versus central stimulation, it seems that the N200 is particularly prone to eccentricity effects, which fits its exogenous characteristic. Further investigation is desirable especially since the moving grid pattern was larger in size. So, to fully disentangle the effects of peripheral versus foveal stimulation, an otherwise identical speller design should be compared.

The advantages of the presented mVEP Spellers are clear. Owing to the qualities of mVEPs visual presentation is convenient. A large vocabulary of 30 letters can be employed and participants do not need any prior training. The problem of limited clinical applicability due to the dependence on eye movements can be solved in the Motion Center Speller. Spelling speed of the current spellers can be significantly enhanced by reducing the number of sequence repetitions. Recent methods allow for a dynamic number of repetitions, by having stimulus presentation stop as soon as the statistical confidence in the classifier's decision is sufficiently high [38]. In the overt attention as well as the central stimulation condition, trade-off between accuracy and spelling speed allows accuracies between 85% and 95% with 5 sequence repetitions.

# Acknowledgments

We acknowledge financial support by the BMBF grant nos 01GQ0850, 01IB001A, DFG grant no MU 987/3-2 and the FP7-ICT Programme of the European Community, under the PASCAL2 Network of Excellence, ICT-216886. This publication only reflects the authors' views. Funding agencies are not liable for any use that may be made of the information contained herein

# References

[1] Jonathan R. Wolpaw, Niels Birbaumer, Dennis J. McFarland, Gert Pfurtscheller, and Theresa M. Vaughan. Brain-computer interfaces for communication and control. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, 113(6):767–791, 2002.

[2] J.R. Wolpaw. Brain-computer interfaces as new brain output pathways. *The Journal of Physiology*, 579:613–619, Mar 2007.

[3] N. Birbaumer and L.G. Cohen. Brain-computer interfaces: communication and restoration of movement in paralysis. *The Journal of Physiology*, 579:621–636, 2007.

[4] L.A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical neurophysiology*, 70:510–523, 1988.

[5] Matthias Sebastian Treder and Benjamin Blankertz. (C)overt attention and visual speller design in an ERP-based brain-computer interface. *Behavioral and Brain Functions*, 6:28, May 2010.

[6] S Luck, editor. *An introduction to the event-related potential technique*. MIT Press, Cambridge, MA, 2007.

[7] M Bach and D Ullrich. Motion adaptation governs the shape of motion-evoked cortical potentials. *Vision Research*, 34:1541–1547, 1994.

[8] J. Polich. Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, 118:2128–2148, Oct 2007.

[9] Craig Gonsalvez and John Polich. P300 amplitude is determined by target-to-target interval. *Psychophysiology*, 39(03):388–396, 2002.

[10] E. Donchin, K. M. Spencer, and R. Wijesinghe. The mental prosthesis: Assessing the speed of a P300-based brain-computer interface. *IEEE transactions on rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, 8(2):174–179, June 2000.

[11] Eric W. Sellers, Theresa M. Vaughan, and Jonathan R. Wolpaw. A brain-computer interface for long-term independent home use. *Amyotrophic Lateral Sclerosis*, 11(5):449–455, 2010.

[12] P. Brunner, S. Joshi, S. Briskin, J. R. Wolpaw, H. Bischof, and G. Schalk. Does the "P300" speller depend on eye gaze? *Journal of neural engineering*, 7:056013, 2010.

[13] S Frenzel, E Neubert, and C Bandt. Two communication lines in a 3 × 3 matrix speller. *Journal of neural engineering*, 8(3):036021, 2011.

[14] R Gattass, C Gross, and J Sandell. Visual topography of V2 in the macaque. *The Journal of Comparative Neurology*, 201:519–539, 1981.

[15] R Tootell, N Hadjikhani, J Mendola, S Marrett, and A Dale. From retinotopy to recognition: fMRI in human visual cortex. *Trends in cognitive sciences*, 2(5):174–183, 1998.

[16] A. B. Usakli, S. Gurkan, F. Aloise, G. Vecchiato, and F. Babiloni. On the use of electrooculogram for efficient human computer interfaces. *Computational intelligence and neuroscience*, 2010:135629, 2010.

[17] G. Bin, X. Gao, Y. Wang, Y. Li, B. Hong, and S. Gao. A high-speed BCI based on code modulation VEP. *Journal of neural engineering*, 8:025015, Apr 2011.

[18] Laura Acqualagna and Benjamin Blankertz. A gaze independent speller based on rapid serial visual presentation. In *Conf Proc IEEE Eng Med Biol Soc*, volume 2011, pages 4560–4563, 2011.

[19] Matthias Sebastian Treder, Nico Maurice Schmidt, and Benjamin Blankertz. Gaze-independent brain-computer interfaces based on covert attention and feature attention. *Journal of neural engineering*, 8(6):066003, 2011. Open Access.

[20] Martijn Schreuder, Thomas Rost, and Michael Tangermann. Listen, you are writing! Speeding up online spelling with a dynamic auditory BCI. *Frontiers in Neuroscience*, 5:112, 2011.

[21] Johannes Höhne, Martijn Schreuder, Benjamin Blankertz, and Michael Tangermann. A novel 9-class auditory ERP paradigm driving a predictive text entry system. *Frontiers in Neuroscience*, 5:99, 2011.

[22] Anne-Marie Brouwer and J. B. F. van Erp. A tactile P300 brain-computer interface. *Frontiers in Neuroscience*, 4(19):036003, 2010.

[23] Marieke Thurlings, Jan Van Erp, Anne-Marie Brouwer, Benjamin Blankertz, and Peter Werkhoven. Control-display mapping in brain-computer interfaces. *Ergonomics*, 2012. in press.

[24] Marieke Thurlings, Anne-Marie Brouwer, Jan Van Erp, Benjamin Blankertz, and Peter Werkhoven. Bimodal versus unimodal: Does bimodal stimulus presentation increase usable ERP components for BCI performance? *Journal of neural engineering*, 2012. in press.

[25] Tao Liu, Leslie Goldberg, Shangkai Gao, and Bo Hong. An online brain-computer interface using non-flashing visual evoked potentials. *Journal of neural engineering*, 7(3):036003, 2010.

[26] S Heinrich. A primer on motion onset visual evoked potentials. *Documenta Ophthalmologica*, 114:83–105, 2007.

[27] M Kuba, Z Kubová, J Kremlácek, and J Langrová. Motion-onset veps: characteristics, methods, and diagnostic use. *Vision Research*, 47:189–202, 2007.

[28] T Probst, H Plendl, W Paulus, E Wist, and M Scherg. Identification of the visual motion area (area V5) in the human brain by dipole source analysis. *Experimental brain research. Experimentelle Hirnforschung. Experimentation cerebrale*, 93:345–351, 1993.

[29] S Zeki, J Watson, C Lueck, K Friston, C Kennard, and R Frackowiak. A direct demonstration of functional specialization in human visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 17:641–649, 1991.

[30] B. Hong, F. Guo, T. Liu, X. Gao, and S. Gao. N200-speller using motion-onset visual response. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, 120:1658–1666, Sep 2009.

[31] E Cameron, J Tai, and M Carrasco. Covert attention affects the psychometric function of contrast sensitivity. *Vision Research*, 42:949–967, 2002.

[32] B Giesbrecht, D Weissman, M Woldorff, and G Mangun. Pre-target activity in visual cortex predicts behavioral performance on spatial and feature attention tasks. *Brain Research*, 1080:63–72, 2006.

[33] L Anllo-Vento and S Hillyard. Selective attention to the color and direction of moving stimuli: electrophysiological correlates of hierarchical feature selection. *Perception & Psychophysics*, 58:191–206, 1996.

[34] Bastian Venthur, Simon Scholler, John Williamson, Sven Dähne, Matthias S Treder, Maria T Kramarek, Klaus-Robert Müller, and Benjamin Blankertz. Pyff – a pythonic framework for feedback applications and stimulus presentation in neuroscience. *Frontiers in Neuroscience*, 4:179, 2010.

[35] A. D. Straw. Vision Egg: an open-source library for realtime visual stimulus generation. *Frontiers in Neuroinformatics*, 2:4, 2008.

[36] Benjamin Blankertz, Steven Lemm, Matthias Sebastian Treder, Stefan Haufe, and Klaus-Robert Müller. Single-trial analysis and classification of ERP components – a tutorial. *NeuroImage*, 56:814–825, 2011.

[37] Fei Guo, Bo Hong, Xiaorong Gao, and Shangkai Gao. A brain computer interface based on motion-onset VEPs. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2008:4478–4481, 2008.

[38] Martijn Schreuder, Johannes Höhne, Matthias Sebastian Treder, Benjamin Blankertz, and Michael Tangermann. Performance optimization of ERP-based BCIs using dynamic stopping. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4580–4583, 2011.