

Overlap and refractory effects in a brain–computer interface speller based on the visual P300 event-related potential

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2009 J. Neural Eng. 6 026003

(<http://iopscience.iop.org/1741-2552/6/2/026003>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 199.184.22.55

This content was downloaded on 25/06/2015 at 13:33

Please note that [terms and conditions apply](#).

Overlap and refractory effects in a brain–computer interface speller based on the visual P300 event-related potential

S M M Martens¹, N J Hill¹, J Farquhar² and B Schölkopf¹

¹ Empirical Inference Department, Max Planck Institute for Biological Cybernetics, Tübingen, Germany

² Radboud University Nijmegen Medical Centre, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

E-mail: smm.martens@gmail.com

Received 27 October 2008

Accepted for publication 10 February 2009

Published 2 March 2009

Online at stacks.iop.org/JNE/6/026003

Abstract

We reveal the presence of refractory and overlap effects in the event-related potentials in visual P300 speller datasets, and we show their negative impact on the performance of the system. This finding has important implications for how to encode the letters that can be selected for communication. However, we show that such effects are dependent on stimulus parameters: an alternative stimulus type based on apparent motion suffers less from the refractory effects and leads to an improved letter prediction performance.

1. Introduction

The visual P300 speller [1] is a brain–computer (BCI) interface which enables users to spell words by focusing their attention on letters in a letter grid displayed on a computer screen. While controlled stimulus events take place on the letters, the electroencephalogram (EEG) of the user is recorded. An algorithm provides a prediction of the letters on which the user focuses by classifying segments of the EEG that are time locked to the stimulus events. The standard stimuli are intensifications (flashes) of the letters in a particular row or column of the letter grid. We introduce some terminology here. The *target letter* is the letter that the subject wants to communicate. The intensifications of the row or column containing the letter that the user wants to communicate are *target events* in a stimulus sequence. The intensifications of all other rows and columns are *non-target events*. An *epoch* is a time period during which an evoked response in the EEG is expected to occur in response to one stimulus event. The time interval between the start of one stimulus event and the start of the next event is called the *stimulus onset asynchrony* (SOA). The interval between the start of a target event and the start of a former target event is referred to as the *target-to-target interval* (TTI). Similarly, we introduce the term *target-to-non-target interval* (TNI) to denote the interval between the start of a non-target event and the start of a previous target event. For each

hypothetical target letter, one can represent its stimulus events as a binary string of 1's and 0's denoting the intervals in which a stimulus event did or did not occur to the letter in question. A *codebook* is a collection of strings for all the letters in the letter grid (see figure 1). By producing different brain responses for the target and non-target events of a particular letter, the user implicitly conveys information of its bitstring. A common strategy for the subject is to count the target events and to ignore the non-target events. The resulting epochs following a target event will contain attention-modulated components such as the P300 event-related potential (ERP) at larger amplitudes than the epochs following a non-target event. A classifier, such as a support vector machine (SVM) [2] or a stepwise-linear discriminant analysis (SWLDA) [3], is trained to recognize the presence of attention-modulated components and labels the epochs as targets or non-targets. In this way, an output bitstring is generated that can be matched to one of the bitstrings in the codebook resulting in a prediction of the target letter. Note that there are two classification problems involved: the binary epoch classification problem (demodulation in figure 1) and the letter classification problem (decoding in figure 1).

Relatively few attempts have been made to improve the set-up of the speller system since its introduction in 1988. There are studies which evaluated the performance of the system using letter grids with more letters [4, 5]. It was found that the P300 amplitude was larger in a larger letter grid and this

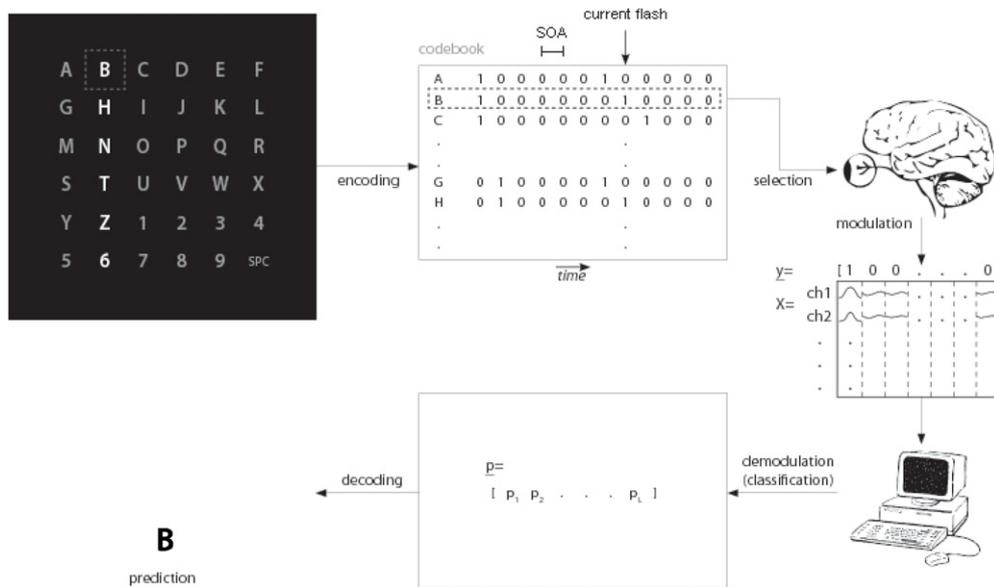


Figure 1. Schematic of the visual speller system. On the left, the letter grid is shown during a stimulus event on the second column. The corresponding column in the codebook is indicated by the arrow. The user has to transmit information of the bitstring corresponding to the letter that the user wants to communicate, e.g. the letter B, by brain signals. A classifier receives an EEG epoch after each stimulus event and decides to which class (0 or 1) it belongs to. The resulting bitstring \underline{p} is then matched to one of the codewords.

was attributed to a different target probability. Here, the target probability refers to the proportion of targets in the sequence of stimulus events. Also, the performance at different settings of the SOA was studied. Farwell and Donchin [1] showed that an SOA of 0.500 s leads to larger P300 amplitudes and a larger letter prediction accuracy than an SOA of 0.125 s for a fixed number of stimulus events. Allison and Pineda [6] found the same relationship between the SOA and P300 amplitude. Sellers *et al* [4], on the other hand, compared SOAs of 0.175 s and 0.350 s and found neither P300 amplitude nor letter prediction accuracy differences. An innovative study by Allison and Pineda [6] involved a comparison of different flashing patterns in the letter grid. They compared the standard stimulus pattern, with rows or columns being flashed, to a multiple stimulus pattern, with half of the letters of the grid being flashed at any one stimulus event. Unexpectedly, they found that for high target probabilities (33% and 50% of the stimulus events being target events) under the multiple stimulus condition, the P300 amplitude did not simply increase as a function of the SOA. An increased task complexity was speculated as a possible factor for the larger P300 amplitude in the multiple stimulus type. Another finding was that subjects make an error of about 4% in counting the target events in the standard row-column flashing speller at an SOA of 0.125 s. A counting error indicates that target events are missed (false negatives), that non-target flashes are mistaken for target flashes (false positives) or that subjects lose count. When using the multiple stimulus pattern, the counting error increased to 19% even though the target probability and SOA were about the same as in the row-column set-up. It was suggested that the increase in counting error was due to subjects losing the count since there was no P300 amplitude effect. In that case, the speller performance should not be affected by the counting error. Unfortunately, this was not investigated in the study.

These studies indicate that a better understanding of the neurophysiological and the psychophysiological effects in the speller system may be advantageous. We will discuss a neurophysiological effect which we believe is related to the observations mentioned above. In visual spellers, it is common to use small SOAs of ~ 0.2 s [3, 7] to achieve high information transfer rates. However, the epoch length is set to about ~ 0.6 – ~ 1.0 s, leading to overlapping epochs. One may wonder to what extent the use of small SOA in the visual speller system introduces overlap and refractory effects of the long-latency ERP components. For example, P300 that appears around 0.3 s after a target flash will also manifest itself in the epoch following the target epoch. Will this confuse the classifier? Evidence of ERP refractory effects comes from auditory studies using targets and non-targets, showing that the P300 amplitude drops and P300 latency increases if the TTI is reduced from 8 to 2 s [8]. The same TTI–P300 amplitude relationship was described in 1980 for even smaller TTIs in a study using only targets [9]. The results of this study indicated a refractory period of about 0.9 s of P300. Although on average the TTI lies around 1 s in the standard speller set-up, in fact the TTI of each individual epoch varies widely since the order of row and column flashes is random (see figure 2). For example, for a 6×6 grid, the TTI may range from $1 \times \text{SOA}$ (~ 0.2 s) to $21 \times \text{SOA}$ (~ 4 s). Consequently, refractory effects may play a role in the targets in the visual P300 speller. The contribution of this paper is threefold, as follows.

- (i) We present an analysis of the characteristics of the brain responses as a function of the stimulus sequence. In particular, we show the presence of overlap and refractory effects of the ERPs in speller data. In addition, we present the impact of the overlap and refractory effects on the classification performance.

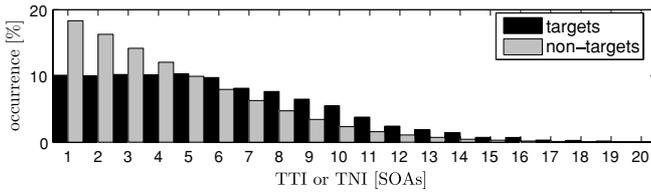


Figure 2. Occurrence of the TTI among the targets (black bars) and TNI among the non-targets (grey bars). Values were derived from competition dataset subject A in which a standard row-column code was used. The total number of non-targets is five times the total number of targets.

- (ii) We present a new stimulus type which will be referred to as the FLIP stimulus. Each letter of the letter grid is placed in a grey box with either a horizontal or vertical orientation; a stimulus event corresponds to a 90° rotation of the box from one video frame to the next (see figure 3). This FLIP stimulus type creates the percept of apparent motion, whereas the standard FLASH stimulus does not. Interestingly, the refractory effects of the ERPs were less prominent or even absent in the proposed stimulus type.
- (iii) We build a simulation model that quantifies the ERP refractory effects.

2. Methods

2.1. Datasets

A standard set-up of the visual speller was implemented using a 6 × 6 grid of alphanumeric characters, presented via an LCD monitor on a desk in a quiet office. Each letter had a low and a high intensity state. The high intensity states lasted 0.033 s. This set-up corresponded to the FLASH stimulus set-up. A second set-up involved the FLIP stimulus as described in point (ii). The monitor refresh rate was 60 Hz. The SOA was set to 0.167 s.

EEG signals were measured using a QuickAmp system (BrainProducts GmbH) in combination with an electrode cap (Electro-Cap International, Inc.). The equipment was set up to measure 58 channels of the EEG, one horizontal EOG at the left eye, one bipolar vertical EOG signal and a synchronization signal from a light sensor attached to the display, all sampled at 250 Hz. All electrode impedances were kept below 10 kΩ.

Six subjects, indicated by {I, II, III, IV, V, VI}, five males and one female, were included in the study with ages between 20 and 40 years. The subjects had little to no experience with the standard visual speller (0 to 3 previous sessions). All subjects participated in both the FLASH and the FLIP set-ups. They all gave informed consent prior to the EEG session. The session was divided into 16 blocks of copy-spelling, each

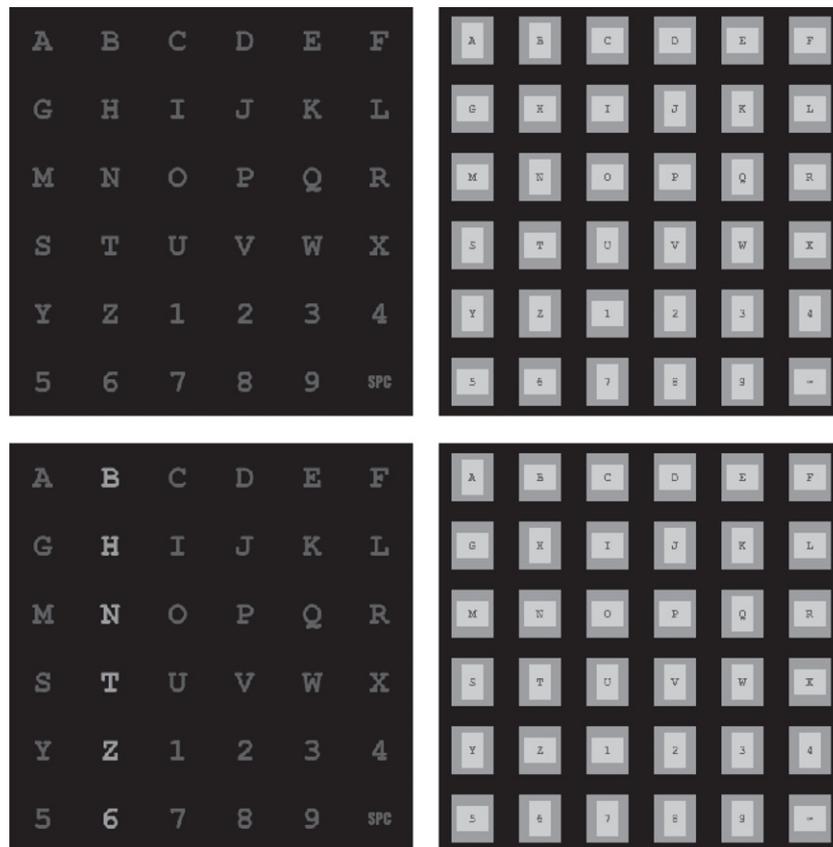


Figure 3. Standard FLASH (left plots) and proposed FLIP (right plots) letter grids. The top plots show the letter grids at the start of the trial (no stimulus event) and the bottom plots show a stimulus event on the second column.

block containing 20 trials (letters to spell) of either the FLASH or FLIP stimulus. In the copy-spelling mode, the subjects had to attend to prescribed letters such that the target letters were always known by the experimenter. During this copy-spelling part, the subjects did not receive any feedback about the letter prediction performance. We applied different letter encoding methods (described in [10]). Here, we restrict our analysis to the subset of trials in which a traditional row–column stimulus code was used. The distributions of TTI and TNI values were the same for the FLASH and the FLIP data. In total, each subject spelled 64 letters per stimulus type. One subtrial, a part of the trial during which each row and column of the letter grid would flash or flip once, lasted $12 \times \text{SOA} = 2.0$ s. Each trial contained six subtrials, such that it took 12 s to complete a letter. The stimulus type was alternated between blocks to prevent a possible bias due to fatigue in either the FLASH or FLIP stimulus. Each trial began with a red box surrounding one of the letters in the letter grid which indicated to the subject the letter (randomly chosen on each trial) they should attend to—this cue came on for a second and was removed 1 s before the start of the stimulus sequence. Subjects were instructed to count the stimulus events at the target location and to minimize the blinking, moving or swallowing during the sequence.

We also included two publicly available datasets in our analysis, i.e. the training sets of subjects A and B of the Wadsworth competition dataset III [11] in which the standard FLASH stimulus was used at an SOA of 0.175 s. The flash duration was 0.100 s. These data will be referred to as FLASH COMP data.

2.2. Signal analysis

Preprocessing. The signal analysis was performed offline in Matlab (The MathWorks, Inc.). The EEG system sampled the data at a sampling frequency of 250 Hz. The EEG channels had a common average reference. We applied an FIR Bartlett–Hanning low-pass filter (cut-off frequency 0.5 Hz, order 125) and high-pass filter (cut-off frequency 10 Hz, order 1000), both with a linear phase response. We compensated for the delay introduced by the filters. The EEG was cut up in 0.600 s epochs synchronized by the stimulus cues. These epochs were downsampled to 25 Hz. The stimulus cues were derived from the unfiltered synchronization signal by setting a threshold for the amplitude. Each cue corresponded to the beginning of each FLASH or FLIP stimulus. Per stimulus type and per subject, we collected $64 \text{ trials} \times 6 \text{ subtrials per trial} \times 2 \text{ target epochs per subtrial} = 768$ target epochs, and similarly $64 \times 6 \times 10 = 3840$ non-target epochs. For each epoch, the time interval between the start of the current epoch and the start of the previous target epoch was derived from the stimulus sequence. Thus, each epoch was tagged with a TTI or a TNI value.

Average ERP. Per TTI and TNI values, the average target and non-target ERPs were calculated. Epochs with absolute amplitude values larger than $100 \mu\text{V}$ were excluded from

the average ERPs. Each average ERP was constructed by averaging at least 50 epochs.

Classification. We used a ν -SVM classifier [12] using all 58 EEG channels after a whitening transformation which decorrelated the signals in a sensor space and made the variance of all the decorrelated components equal to 1. We found that the whitening step increased the letter prediction accuracy significantly. The data were split into a training and test set (see paragraph *Letter prediction*). The SVM parameter ν was found by a ten-fold cross-validation on each training set. In order to alleviate class imbalance problems (see, for example, [13]), the number of target and non-target epochs in the training sets was balanced by throwing away excessive non-target epochs. We calculated the per-class classification accuracy on the test set per TTI and TNI as follows. The target classification accuracy was derived as the percentage of correctly classified targets among the target epochs at a particular TTI; the non-target classification accuracy was derived as the percentage of correctly classified non-targets among the non-target epochs at a particular TNI.

Letter prediction. Letter prediction was performed on a leave-one-letter-out basis; i.e. for the prediction of letter with index i , all epochs corresponding to trials $j \neq i$ were used to train the classifier and all epochs corresponding to trial i constituted the test set. For each letter, the classifier outputs for the nonzero entries in the codeword corresponding to that letter were summed up. The letter with the largest total output value was then selected. After running this for all the trials per dataset, we calculated the letter prediction accuracy as the fraction of correctly predicted letters.

3. Results

We found overlap and refractory effects in the average ERPs in our speller datasets. Figure 4 shows the average target and non-target ERPs per TTI and TNI value from one subject who used both the FLASH and FLIP stimuli. For $\text{TTI} = \text{TNI} = 6 \times \text{SOA}$ (first two columns, top plots), there is a clear difference between the target and non-target ERPs. The P300 attention-modulated component appears as a positive deflection between 0.15 and 0.55 s in the target ERPs. Now observe the ERPs for increasingly smaller TTI and TNI values and note that the shape of the target and the non-target ERPs changes. For $\text{TTI} = \text{TNI} = 1 \times \text{SOA}$ (first two columns, bottom plots) the P300 component seems to be present, but at an earlier latency and in both the target and non-target ERPs. In fact, this is the P300 component of a preceding target event which overlaps with the current target and non-target ERPs. The P300 component which one would expect in the target ERP between 0.15 and 0.55 s is highly attenuated for $\text{TTI} = 1 \times \text{SOA}$ in the FLASH stimulus (first column, bottom plot) and, consequently, the target and non-target ERPs are almost indiscriminable. This attenuation of the attention-modulated components at small TTI values is referred to as the ERP refractory effect. Note that for the FLIP data, the P300 component is preserved at

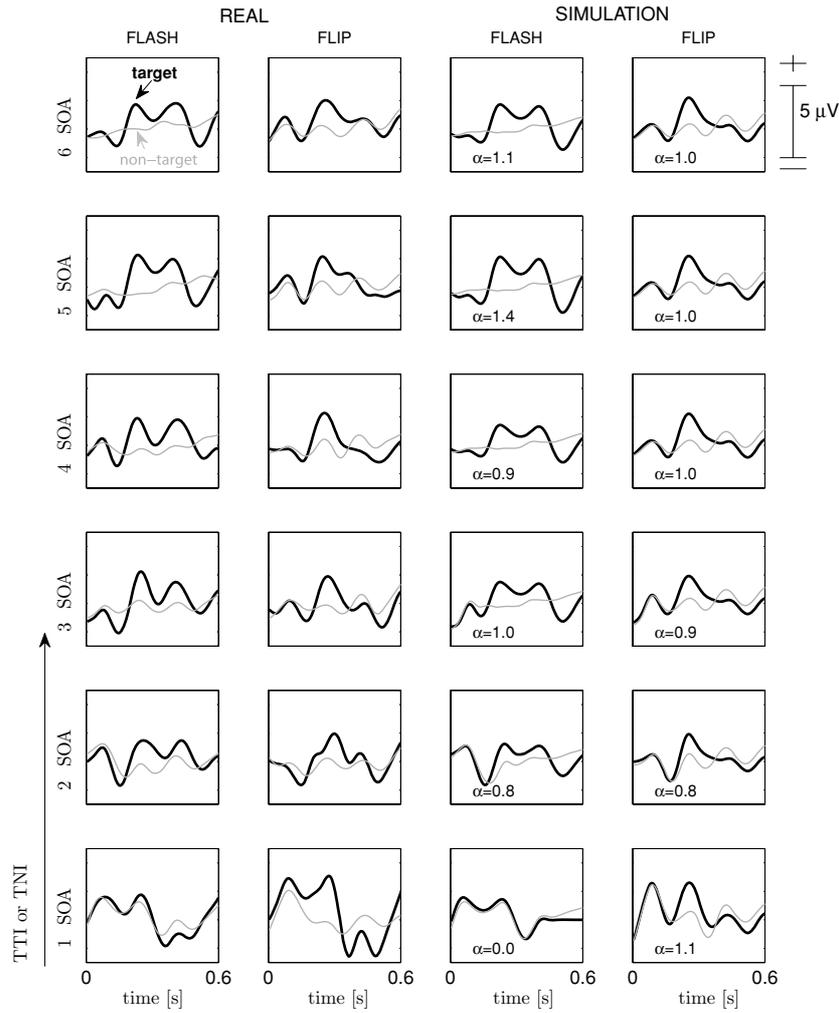


Figure 4. The plots in the first two columns show average target (thick black line) and non-target (thin grey line) ERPs on Cz from subject V. The ERPs have been plotted as a function of the TTI and TNI values for the FLASH and FLIP stimulus types. From top to bottom, the TTI and TNI decrease from $6 \times \text{SOA}$ to $1 \times \text{SOA}$ ($\text{SOA} = 0.167 \text{ s}$). The plots in the last two columns show the simulated ERPs for subject V for which the scaling factor α (TTI) for refractory effects is also depicted.

$\text{TTI} = 1 \times \text{SOA}$ (second column, bottom plot), indicating that there the refractory effect is less prominent or even absent.

The target classification accuracy, averaged over all subjects per stimulus type, decreased with smaller TTI values for $\text{TTI} < 5 \times \text{SOA}$ for the FLASH COMP and FLASH data (figure 5). At $\text{TTI} = 1 \times \text{SOA}$, the target classification accuracy even approached the chance level. This TTI dependence was not found in the FLIP data, although the target classification accuracy was reduced for $\text{TTI} = 2 \times \text{SOA}$. The non-target classification accuracy did not show a TNI dependence in any of the stimulus types. The per-subject target classification accuracy was lower at $\text{TTI} = 1 \times \text{SOA}$ than at $\text{TTI} = 10 \times \text{SOA}$ for both subjects from the FLASH COMP, all six subjects from the FLASH and two out of six subjects from the FLIP stimulus (Fisher’s exact test, $p < 0.05$). The per-subject target classification accuracy at $\text{TTI} = 2 \times \text{SOA}$ was significantly lower in both FLASH COMP subjects, five out of six FLASH subjects and three out of six FLIP subjects.

The FLIP stimulus type resulted in better letter prediction performance than the FLASH stimulus in five out of six subjects

(figure 6). However, the difference was only significant for two subjects (Fisher’s exact test, $p < 0.05$).

4. Simulation model

We constructed a simulation model to verify and quantify the refractory effects in speller data.

4.1. Description of the model

We denote the discrete time by k and the sampling frequency in Hz by f_s . We assume that a target response $\hat{i}(k, \text{TTI})$ and a non-target response $\hat{n}(k, \text{TNI})$ at a given TTI and TNI value can be constructed as follows:

$$\hat{i}(k, \text{TTI}) = \bar{i}(k) \cdot \alpha(\text{TTI}) + \bar{i}(k + \tau_t), \quad (1)$$

$$\hat{n}(k, \text{TNI}) = \bar{n}(k) + \bar{i}(k + \tau_n), \quad (2)$$

with $k = (1, 2, \dots, K)$. Here, the target and non-target responses are constructed by summing up a number of components. The terms $\bar{i}(k)$ and $\bar{n}(k)$ denote a target and

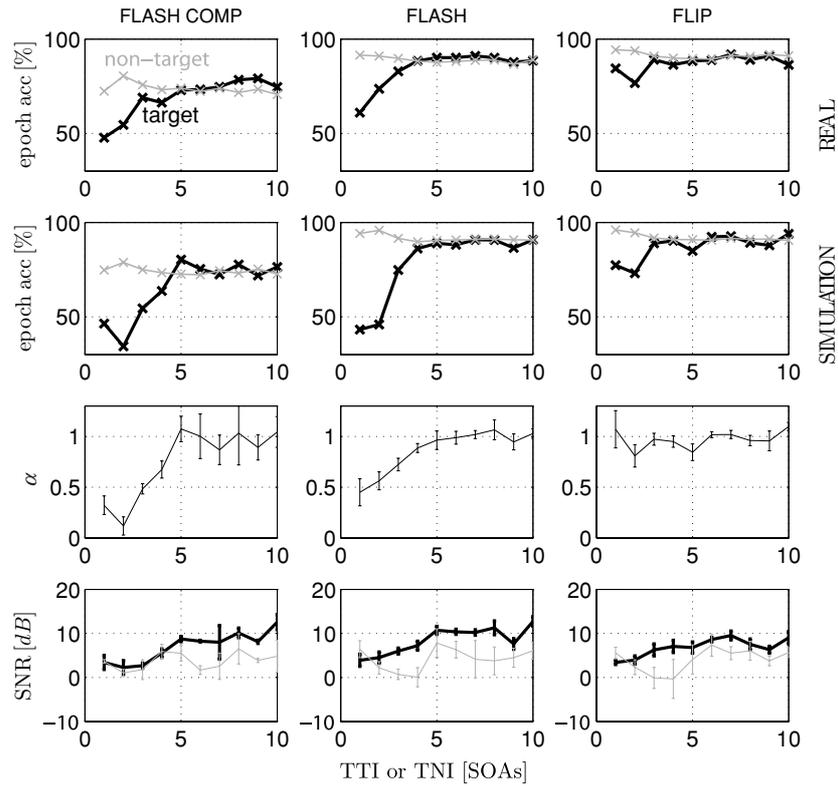


Figure 5. Target (black line) and non-target (grey line) epoch classification accuracies as a function of TTI and TNI (first row of plots). Each column shows the result for a different stimulus type; values are averaged over the subjects. Since the classes were balanced, the chance performance for this binary epoch classification problem is 50%. Also, the classification accuracy for the simulated data is shown, averaged over the subjects (second row). The third and fourth rows show the scaling factor α (TTI) and SNR_P (black line) and SNR_N (grey line), averaged over subjects with standard error bars of the mean.

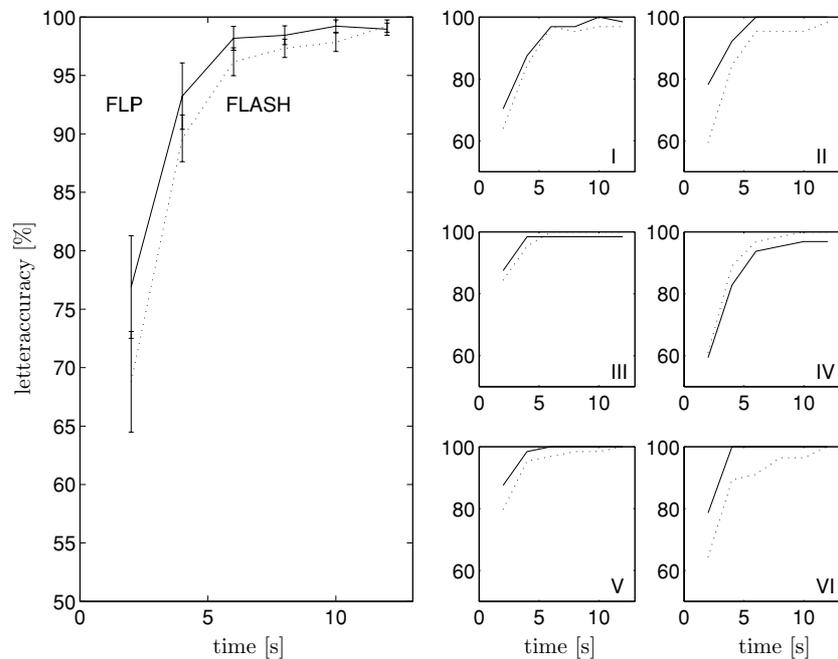


Figure 6. Letter prediction accuracy as a function of trial duration for the FLASH (dashed line) and FLIP (solid line) stimulus type. The left plot shows the average over all the subjects in our study, as well as the standard error bars of the mean. The smaller plots show the letter prediction accuracy per subject. One subtrial corresponds to 2 s. The chance performance for this 36-class letter prediction problem is 3%.

a non-target template, respectively. The terms $\bar{i}(k + \tau_t)$ and $\bar{i}(k + \tau_n)$ are shifted versions of the target template and represent

the ERP overlap due to a target response from a preceding target event, with $\tau_t = \lfloor TTI \cdot f_s \rfloor$ and $\tau_n = \lfloor TNI \cdot f_s \rfloor$ samples,

$[\cdot]$ being the nearest integer function. Both TTI and TNI are expressed in s. We set these overlap components to 0 for $k > K - \tau_t$ and $k > K - \tau_n$. Note that the constructed target and non-target responses with sufficiently large TTI and TNI values are free from overlap. For example, given an SOA of 0.167 s or 0.175 s and $K = \lfloor 0.6 \cdot f_s \rfloor$, the overlap components are zero for target and non-target responses with $\{\text{TTI}, \text{TNI}\} \geq 4 \times \text{SOA}$ s. The target template $\bar{t}(k)$ in (1) is scaled by $\alpha(\text{TTI})$ to model the ERP refractory effects as a function of the TTI value, where $\alpha < 1$ results in an attenuated target response. The special case $\alpha = 1$ corresponds to the absence of refractory effects.

4.2. Dataset generation

For each subject of the FLASH COMP, FLASH and FLIP data, we established the TTI or TNI value of each epoch, where TTI or TNI values larger than $10 \times \text{SOA}$ were set to $10 \times \text{SOA}$. Epochs had length $K = \lfloor 0.6 \cdot f_s \rfloor$. We constructed an average target ERP $t(k, \text{TTI})$ and non-target ERP $n(k, \text{TNI})$ per TTI and TNI value for channel Cz. We assumed that the overlap and refractory effects in epochs with TTI and TNI values equal to or larger than the 0.9 s P300 refractory period from [9] were negligible. Consequently, we derived a target $\bar{t}(k)$ and non-target template $\bar{n}(k)$ by averaging the target and non-target epochs with $\{\text{TTI}, \text{TNI}\} \geq 5 \times \text{SOA}$. We upsampled $t(k, \text{TTI})$, $n(k, \text{TNI})$, $\bar{t}(k)$ and $\bar{n}(k)$ to 500 Hz to avoid mismatch in the following steps due to rounding errors. For each TTI and TNI value, a target response $\hat{t}(k, \text{TTI})$ and non-target response $\hat{n}(k, \text{TNI})$ with $K = \lfloor 0.6 \cdot f_s \rfloor$ was derived using (1) and (2). The scaling factor $\alpha(\text{TTI})$ was estimated by finding the least-squares solution of $t(k, \text{TTI}) - \hat{t}(k, \text{TTI})$ according to

$$\alpha(\text{TTI}) = \frac{1}{R_{\bar{t}}} (R_{t, \bar{t}} - R_{\bar{t}, \text{shift}, \bar{t}}), \quad (3)$$

with $R_{\bar{t}} = E[\bar{t}(k) \cdot \bar{t}(k)]$, $R_{t, \bar{t}} = E[t(k, \text{TTI}) \cdot \bar{t}(k)]$ and $R_{\bar{t}, \text{shift}, \bar{t}} = E[\bar{t}(k + \tau_t) \cdot \bar{t}(k)]$. All $\alpha < 0$ were set to 0. As measures of goodness of fit, we define SNR_P and SNR_N :

$$\text{SNR}_P(\text{TTI}) = 10 \log_{10} \frac{E[t(k, \text{TTI})^2]}{E[(t(k, \text{TTI}) - \hat{t}(k, \text{TTI}))^2]}, \quad (4)$$

$$\text{SNR}_N(\text{TNI}) = 10 \log_{10} \frac{E[n(k, \text{TNI})^2]}{E[(n(k, \text{TNI}) - \hat{n}(k, \text{TNI}))^2]}. \quad (5)$$

For each TTI and TNI for each subject, a number of synthetic epochs were created by adding Gaussian white noise $r(k) \sim \mathcal{N}(0, \sigma^2)$ with zero mean and standard deviation σ to the simulated target and non-target responses. Each simulated epoch saw a different realization of the noise. The noise represented the non-ERP-related EEG signals, such as background EEG and artefacts, which the classifier has to deal with when labelling the epochs. Although in reality these non-ERP-related EEG signals do not have a white power spectrum, we used this simple noise model assuming that the spectral characteristics of the noise would not affect the simulation outcome.

4.3. Classification

We trained a ν -SVM classifier on 1000 targets and 1000 non-targets in each synthetic dataset. The distribution of the TTI values over the targets and of the TNI values over the non-targets in the training set was the same as in the real dataset of the subject. We tested on 500 targets and 500 non-targets per TTI and TNI value. Then, we collected the classifier outputs per TTI and per TNI and calculated the per-class epoch classification accuracy. Note that a larger value for σ decreases the signal-to-noise ratio and therefore leads to a poorer epoch classification accuracy. The value for σ was set such as to match the resulting balanced classification accuracy with the real classification accuracy at $\{\text{TTI}, \text{TNI}\} \geq 5 \times \text{SOA}$.

4.4. Simulation results

Figure 4 shows the simulated target and non-target responses of one subject. Since the responses with $\{\text{TTI}, \text{TNI}\} > 3 \times \text{SOA}$ are free from overlap components, they depict the target and non-target templates, apart from a scaling factor $\alpha(\text{TTI})$ in the target response. At small TTI and TNI values, the overlap effects are visible. Note that for the small TTI values, $\alpha < 1$ for the FLASH stimulus whereas $\alpha \approx 1$ for the FLIP stimulus type. Figure 5 shows the per-class classification accuracies for the real dataset and the simulation model. It also shows the least-squares-based $\alpha(\text{TTI})$ for the different stimulus types and goodness-of-fit measures SNR_P and SNR_N . We derived Spearman's rank correlation coefficient between α and TTI value per stimulus type, taking into account TTI values of up to 0.9 s. Rankings were derived per subject and per stimulus type, after which all subjects per stimulus type were used per correlation test. A significant correlation at the 5% level was found for the FLASH COMP (Spearman's rank correlation coefficient $\rho = 0.9$, $p = 0.0002$) and the FLASH DATA ($\rho = 0.8$, $p = 0.0000$), but not for the FLIP data ($\rho = 0$, $p = 0.5000$).

5. Discussion

There is no obvious explanation for the reduction in refractory effects and the gain in performance for the proposed FLIP stimulus. We want to stress that the psychophysical and neurophysiological findings in the literature on visual stimuli are not easily interpolated to our data since the paradigms often use stimuli with completely different physical properties. Nevertheless, we will discuss some ideas that we thought relevant in the light of our findings.

The larger number of edges and corners in the FLIP stimulus may lead to larger responses in area V1 of the primary visual cortex [14]. Furthermore, the detection of FLASH events is a detection of a luminance change (first-order stimuli), whereas the FLIP events do not entail overall changes in luminance but rather in object orientation (second-order stimuli). There are indications that the detection mechanisms for first-order stimuli versus second-order stimuli are distinct [15–18]. We, however, could not find any evidence of this while comparing the scalp topologies of the classification information between the FLASH and FLIP stimulus types.

With respect to the reported errors in counting the target flashes in the visual speller [6], a study by Boynton [19] reported that the temporal resolution for the detection of a double flash depends on the temporal length of the stimulus. Thus, the critical SOA required to perceive two flashes as two separate visual events is related to the flash duration. When the SOA is below this critical value, the subject perceives the two flashes as one flash. For instance, for flash durations around 0.100 s the critical SOA would be around 0.120 s [19, 20]. According to these results, the SOAs used in our study are safely above the critical SOA value. However, the speller system might require a larger SOA due to fatigue [6, 21]. In some subjects, we could see a negative trend in the percentage of correctly classified target epochs at $TTI = 1 \times SOA$ from the beginning to the end of the subject's session, which supports this idea. Thus, we speculate that a second target flash is frequently missed when two target flashes happen right after each other, whereas the strong motion signal of a double flip makes this less likely in the FLIP stimulus. A small-TTI target detection failure would explain the decrease in performance for the targets at small TTI. An experiment in which subjects count the target events at different SOAs and different flash durations could resolve this question.

Another issue worth mentioning is the interfering effect of the non-target events in the periphery. Studies have shown that a luminance flicker, presented peripheral to a foveal test target, increases thresholds for target detection [22, 23]. It was shown that the motion sensitivity of some second-order stimuli in peripheral vision is reduced compared to that of first-order stimuli [24]. We speculate that the luminance changes of non-attended peripheral letters in the FLASH stimulus may have a larger distracting impact than the iso-luminance flips of non-attended peripheral letters in the FLIP stimulus. In correspondence with this, the horizontal and vertical EOG, which were recorded simultaneously with the EEG, showed a larger variance in the FLASH data than in the FLIP data, but only the horizontal eye activity difference was significant (Wilcoxon signed-rank test, $p < 0.05$). Also, some subjects reported that the non-target events in the FLASH were more distracting than in the FLIP stimulus. Eye movements may lead to EEG artefacts and to missed target events. Moreover, attention directed to distractors leads to a reduced attention to target events. Therefore, the smaller distracting effect of non-target events in the FLIP stimulus could be another explanation for the superior performance of the speller using this stimulus type.

The simulation model reproduced the small-TTI decrease in target classification accuracy for the FLASH COMP and FLASH data and the relatively constant target classification accuracy for the FLIP data. However, the reduction in goodness of fit for the small TTI values in the simulation model indicates that there is room for improvement. The current model does not allow for latency shifts of the attention modulation components, whereas TTI-dependent latency shifts of P300 have been reported [8, 9]. If latency shifts are present, the least-squares estimation of the scaling factor for the refractory effects will be prone to errors. Also, the model only uses channel Cz to construct the templates and to find the scaling

factor. Other channels may contain additional information about refractory and overlap effects. Therefore, extending the model to multiple channels could give a more complete picture of the ERP effects. Furthermore, the model does not correct for the fact that target events following a target or a non-target event may also cause ERP overlap components.

6. Conclusion

We have shown that both overlap and refractory effects of the ERPs are present in the standard visual speller. Epochs characterized by a small TTI or TNI display long-latency components of a preceding target ERP. Moreover, our simulation model showed that the amplitude of the attention-modulated components decreases for smaller TTI values. As a result, the target epochs with a TTI of about ~ 0.2 s are characterized by a severely reduced classification accuracy approaching the chance level. Our findings provide insight into how to optimize the codebook from a neurophysiological viewpoint. For instance, one may design the stream of stimulus events such that the number of small TTIs is minimized.

We proposed a new stimulus type for the speller system based on two-frame apparent motion of blocks with letters instead of flashing letters. The subjects in this study had an offline letter prediction accuracy close to 100%, independent of the stimulus type. The scaling factor α showed no significant correlation with the TTI value. This implies that the refractory effects of the target epochs at small TTI are reduced or even absent for the proposed FLIP stimulus. Correspondingly, the classification performance of the targets at small TTI is less affected in the FLIP stimulus than in the standard FLASH stimulus. The absence or reduced amount of refractory effects in the FLIP stimulus is very promising for the use of letter encodings that involve error correcting techniques in the visual speller; see also [10].

An evaluation of the proposed FLIP stimulus type on patients with severe motor disabilities, such as ALS and tetraplegic patients, would be of great interest. Future work will be on designing a method that exploits knowledge about overlap and refractory effects in the speller. Possibly, this will bring the visual speller performance to the next level.

Acknowledgments

The authors wish to thank Felix Bießmann for exploratory work and Professor Lynne Kiorpes for her input on the perceptual characteristics of the different stimulus types.

References

- [1] Farwell L A and Donchin E 1988 Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials *Electroencephalogr. Clin. Neurophysiol.* **70** 510–23
- [2] Kaper M, Meinicke P, Grossekhoefer U, Lingner T and Ritter H 2004 BCI Competition 2003–data set IIb: support vector machines for the P300 speller paradigm *IEEE Trans. Biomed. Eng.* **51** 1073–6

- [3] Krusienski D J, Sellers E W, Cabestaing F, Bayouhd S, McFarland D J, Vaughan T M and Wolpaw J R 2006 A comparison of classification techniques for the P300 speller *J. Neural Eng.* **3** 299–305
- [4] Sellers E W, Krusienski D J, McFarland D J, Vaughan T M and Wolpaw J R 2006 P300 event-related potential brain–computer interface (BCI): the effects of matrix size and inter stimulus interval on performance *Biol. Psychol.* **73** 242–52
- [5] Allison B Z and Pineda J A 2003 ERPs evoked by different matrix sizes: implications for a brain computer interface (BCI) system *IEEE Trans. Neural Syst. Rehabil. Eng.* **11** 110–3
- [6] Allison B Z and Pineda J A 2006 Effects of SOA and flash pattern manipulations on ERPs, performance and preference: implications for a BCI system *Int. J. Psychophysiol.* **59** 127–40
- [7] Serby H, Yom-Tov E and Inbar G F 2005 An improved P300-based brain–computer interface *IEEE Trans. Neural Syst. Rehabil. Eng.* **13** 89–98
- [8] Gonsalvez C L and Polich J 2002 P300 amplitude is determined by target-to-target interval *Psychophysiology* **39** 388–96
- [9] Woods D L, Hillyard S A, Courchesne E and Galambos R 1980 Electrophysiological signs of split-second decision-making *Science* **207** 655–7
- [10] Hill N J, Farquhar J, Martens S M M, Bießmann F and Schölkopf B 2009 Effects of stimulus type and of error-correcting code design on BCI speller performance *Advances in Neural Information Processing Systems* vol 21 ed D Koller, Y Bengio, D Schuurmans, L Bottou and A Culotta <http://books.nips.cc/nips21.html>
- [11] Wadsworth BCI Dataset. BCI Competition III Challenge 2004 http://ida.first.fraunhofer.de/projects/bci/competition_iii
- [12] Schölkopf B, Smola A J, Williamson R and Bartlett P L 2000 New support vector algorithms *Neural Comput.* **12** 1207–45
- [13] Japkowicz N and Stephen S 2002 The class imbalanced problem: a systematic study *Intell. Data Anal.* **6** 429–49
- [14] Hubel D H and Wiesel T N 1959 Receptive fields of single neurones in the cat's striate cortex *J. Physiol.* **148** 574–91
- [15] De Valois R L, Albrecht D G and Thorell L G 1982 Spatial frequency selectivity of cells in macaque visual cortex *Vis. Res.* **22** 545–59
- [16] Hubel D H and Wiesel T N 1968 Receptive fields and functional architecture of monkey striate cortex *J. Physiol.* **195** 215–43
- [17] Albright T D, Desimone R and Gross C G 1984 Columnar organization of directionally selective cells in visual area MT of the macaque *J. Neurophysiol.* **51** 16–31
- [18] Maunsell J H and Van Essen D C 1983 Functional properties of neurons in middle temporal visual area of the macaque monkey: I. Selectivity for stimulus direction, speed, and orientation *J. Neurophysiol.* **49** 1127–47
- [19] Boynton R M 1972 Discrimination of homogeneous double pulses of light *Handbook of Sensory Physiology: Visual Psychophysics* vol VII/4 (New York: Springer) pp 202–32
- [20] Poggel D A, Treutwein B, Calmanti C and Strasburger H 2006 Increasing the temporal gain: double-pulse resolution is affected by the size of the attention focus *Vis. Res.* **46** 2998–3008
- [21] Carandini M 2000 Visual cortex: fatigue and adaptation *Curr. Biol.* **10** R605–7
- [22] Breitmeyer B, Valberg A, Kurtenbach W and Neumeier C 1980 The lateral effect of oscillation of peripheral luminance gratings on the foveal increment threshold *Vis. Res.* **20** 799–805
- [23] Breitmeyer B G and Valberg A 1979 Local foveal inhibitory effects of global peripheral excitation *Science* **203** 463–4
- [24] Pantle A 1992 Immobility of some second-order stimuli in human peripheral vision *J. Opt. Soc. Am. A* **9** 863–7