Predictive spelling with a P300-based brain-computer interface: Increasing the rate of communication

D.B. Ryan[1], **G.E. Frye**[1], **G. Townsend**[2], **D.R. Berry**[1], **S. Mesa-G**[1], **N.A. Gates**[1], and **E.W. Sellers**[1]

[1] East Tennessee State University, Johnson City, TN 37601, USA

[2] Algoma University, Sault Ste. Marie, Ontario, P6A 2G4, Canada

Abstract

This study compared a conventional P300 speller brain-computer interface (BCI) to one used in conjunction with a predictive spelling program. Performance differences in accuracy, bit rate, selections per minute, and output characters per minute (OCM) were examined. An 8×9 matrix of letters, numbers, and other keyboard commands was used. Participants (n = 24) were required to correctly complete the same 58 character sentence (i.e., correcting for errors) using the predictive speller (PS) and the non-predictive speller (NS), counterbalanced. The PS produced significantly higher OCMs than the NS. Time to complete the task in the PS condition was 12min 43sec as compared to 20min 20sec in the NS condition. Despite the marked improvement in overall output, accuracy was significantly higher in the NS paradigm. P300 amplitudes were significantly larger in the NS than in the PS paradigm; which is attributed to increased workload and task demands. These results demonstrate the potential efficacy of predictive spelling in the context of BCI.

1. Introduction

Brain-computer interface (BCI) technology can help people with severe neuromuscular disease communicate (Wolpaw & Birbaumer, 2006). For example, amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease that may eventually cause people to become completely paralyzed, or locked-in to their bodies, and typically causes death within 2–5 years (Kunst, 2004). Until recently, it was assumed that cognitive function remains intact even in advanced stages of ALS; however, current research shows that some people with ALS experience some type of cognitive impairment, although the actual number of people affected is still debated (Murphy et al., 2007). Nonetheless, people with advanced ALS have little or no means of effective communication given existing alternative and augmentative communication (AAC) devices. For these people, a BCI may be the only option for independent communication.

Corresponding author: Eric W. Sellers, Ph.D., Director, Brain-Computer Interface Laboratory, Assistant Professor, Department of Psychology, East Tennessee State University, P.O. Box 70649, Johnson City, TN 37614, Office phone: 423.439.4476, Lab phone: 423.439.5695, Fax: 423.439.5695, sellers@etsu.edu.

The P300 BCI is based on event-related potentials (ERPs). An ERP is a time-locked electrophysiological brain response to a meaningful stimulus. The P300 ERP is a positive going deflection occurring approximately 300 ms post event. P300 BCI has received much attention because it requires little training due to the P300 ERP being elicited by meaningful attended stimuli (Picton, 1992; Ritter & Vaughan, 1969), and as compared to other BCIs it produces high bit rates (e.g., Serby et al., 2005). The first P300 BCI was described by Farwell and Donchin (1988). Since that time, approximately 90 papers addressing the topic have been published. Moreover, people have now begun to use the P300 BCI in their homes on a daily basis (Sellers, Vaughan, & Wolpaw, in press) and Vaughan et al (2006) have described a research program focused on placing BCIs in numerous homes of people with severe communication disorders. The system uses BCI2000 software (Schalk, McFarland, Hinterberger, Birbaumer, & Wolpaw, 2004), and can provide icon selection, alphanumeric character selection, and multiple menus. These components can provide input to other software and even environmental control. It is now clear that a P300 BCI can be an effective method of communication for ALS patients (e.g., Kubler et al., 2005; Nijboer et al., 2008; Sellers et al., in press; Sellers & Donchin, 2006).

The P300 BCI first models a given participants response to attended stimuli and then uses that information to try and determine which of the items being presented is the one that the subject wishes to select. Typically, the P300 BCI can provide between three and eight selections per min; this study examines how a predictive speller can transform these selections into additional output characters and the predictive speller's effects on performance measures. Previous studies have measured performance through accuracy (percentage correct), selections per minute (total selections correct or incorrect in a minute), and bit rate (formulated from accuracy, number of possible choices, and time to complete a task). In this study, we introduce a new performance measure "output characters per minute" or OCM. OCM was calculated by taking the total selections to complete a session (including spaces, and a selection to end the session) and dividing it by the total time to complete the task. This new measure was used to calculate the contribution of the predictive speller program. It is important to include a performance measure such as OCM when examining the effectiveness of a BCI system because it provides more useful information than accuracy and/or bit rate alone. That is, OCM provides information about how "powerful" each selection is in terms of what it can accomplish. In other words, OCM is more or less independent of accuracy and bit rate. In addition, OCM is certainly more important to the BCI user than bit rate because it provides a realistic assessment of the system output, which bit rate cannot.

Predictive spelling applications have previously been examined in the context of AAC devices. Typically these comparisons use interfaces such as manual typing (Venkatagiri, 1994), mouth stick typing (Koester & Levine, 1996; Koester & Levine, 1994), or touch screen typing (Trnka, McCaw, Yarrington, McCoy, & Pennington, 2009). A primary goal of this research is to examine and maximize the benefits of word prediction by reducing user effort and maximizing output; however, these studies have produced conflicting results regarding the efficacy of predictive spelling applications (Garay-Vitoria & Abascal, 2004, 2006). Some researchers have suggested that keystroke savings as high as 50–60% is a realistic limit of the benefits of delayed word prediction with an AAC user (Copestake 1997; Lesher & Rinkus, 2002). Conversely, it has been noted that significant cognitive demands occur with the use of word prediction programs, and that savings in keystrokes do not necessarily lead to an increase in the rate of communication (Koester & Levine, 1994; Venkatagiri, 1994).

Since a predictive speller may enable the user to produce more information with fewer selections, it has the ability to enhance communication for those who depend on a P300 BCI.

While predictive spellers have been used in-home with ALS patients (Sellers et al., in press), a formal comparison between the use of a predictive P300-speller and a conventional P300 speller has never been conducted. Therefore, we integrated a predictive speller software package into a P300 BCI and compared its performance to a non-predictive (i.e., conventional) system.

1.1 The present study

To approximate in-home use, participants were required to accurately copy a sentence and stop the session once complete. This is the first study to hold the participant to the same simple, yet tedious, demands of an in-home user. To make P300 BCIs more viable for every day home use for individuals that rely on communication devices, the program must be able to quickly output words without sacrificing accuracy. Conventional performance measures (i.e., accuracy, bit rate) were not designed for an additional output from a second program such as a predictive speller. These performance measures are only based on single selections made by the user; they do not encompass the potential output of a selection. Thus, OCM was used to accurately measure the advantage or disadvantage of the predictive speller.

We predicted that the predictive spelling (PS) paradigm will improve performance, in terms of OCM, as compared to the non-predictive spelling (NS) paradigm because the same number of selections per minute (or bit rate) should allow participants to select several items at a time (i.e., words).

We also predict that in the PS paradigm, P300 amplitude may be reduced and P300 latency may be lengthened due to increases in workload or dual task interference in the PS paradigm (e.g., Isreal, Chesney, Wickens, & Donchin, 1980; Isreal, Wickens, Chesney, & Donchin, 1980; Kramer, Wickens, & Donchin, 1985; Wickens, Kramer, Vanasse, & Donchin, 1983). It is reasonalbe to assume that using a PS in addition to a BCI is more cognitively demanding than using a conventional BCI. In the conventional method, other than attending to the desired item, the only task of the participant is to evaluate the feedback between selections and determine what to select next, either backspace or the next character. Using a PS requires more attentional resources than the conventional method. An invidiviual using a PS must a) evaluate whether or not an item is correct, b) decide if an incorrect item must be corrected, c) evaluate the list of suggested words from the predictive speller, and d) determine whether the next selection will be a backspace, an undo, a word from the list, or the next character of a word. Indeed, predictive spellers used in non-BCI context have shown an increase in cognitive demand (Koester & Levine, 1994; Venkatagiri, 1994). These cognitive effects will become evident in the performance measures, but any negative effects will be overshadowed by the increase in communication rate.

2. Methods

2.1 Participants

Twenty nine able-bodied adults were recruited from the East Tennessee State University undergraduate subject pool. Twenty four (10 men, 14 women; age range 18–47) completed the experiment. All were naïve to BCI use and none had uncorrected visual impairments or any known cognitive deficit. The study was approved by the East Tennessee State University Institutional Review Board and each subject gave informed consent.

2.2 Experimental paradigm

Each participant completed two experimental sessions on separate days within a one-week period. Participants completed one PS and one NS session; sessions were counter-balanced to control for order effects. Each session consisted of a calibration phase and an online test

phase using an identical 8×9 matrix. Classification coefficients (described below) were generated with data collected during the calibration phase and subsequently applied during the online test phase. In each phase, participants were provided target items to select. In the calibration phase, items were displayed at the top of the monitor with the next item-to-spell (the target item) indicated in parentheses at the end of the word. As shown in Figure 1A, if the assigned word was "DRIVING," it would appear at the beginning of the run as: DRIVING (D). The participant's task was to attend to (or count) the number of times the item in parentheses flashed. After the first item, there was a 3.5-second pause before the next target appeared in parentheses (e.g. DRIVING (R)). This process repeated until the word was complete (one run). Data were collected from five such runs (4 words and 1 numeric string). For both the PS and NS, each set of items flashed for 62.5ms. This was followed by a 62.5ms inter-stimulus interval. Thus, a flash occurred every 125ms (i.e., 8 flashes/second). For each of the 36 calibration items, five complete sequences (i.e., including 10 flashes of the target item) occurred. The flashes were presented using the checkerboard paradigm (CBP), which presented items in a quasi-random format. The CBP does not allow adjacent items to flash in the same group nor does it allow any item to flash without a minimum of six intervening flashes (for more details see Townsend et al., 2010).

During the online test phase of the NS paradigm, participants copied a sentence from a Notepad "target window" to a blank Notepad "output window" (Figure 1B, top and middle left). The target sentence consisted of 58 selections, including spaces between words, a period, and a Sleep command to end the session. At the beginning of the test phase the output window was blank and the participant's task was to copy the entire sentence correctly, lowercase letters were used for the output window to reduce possible confusion between the target and output windows. After each item selection feedback was presented to the participant (as a translucent character that filled approximately 30% of the screen), and the keystroke was entered into the Notepad output window. In the event of an incorrect selection, the participant was required to use the Backspace (Bs) command to erase the error and then correct the selection. After each selection a 6 second pause was provided before the next set of sequences began to flash. This pause was provided to ensure that the participant had sufficient time evaluate the feedback presented by the BCI, decide what the next item selection should be, and to find the correct item in the 8×9 matrix.

The online test phase of the PS paradigm was identical to that of the NS except for the addition of the Quillsoft WordQ2 (version 2.5) predictive spelling program (Figure 1B, left bottom). BCI2000 (Schalk et al., 2004) includes a UDP that can send output to peripheral programs. The interface between WordQ2 and BCI2000 was achieved using the BCIKeyboard, a program written and supported by the BCI2000 software project. Once an item had been selected and appeared in the output window, the WordQ2 window would populate with seven words, each preceded by a number. In the event that the participant desired to select a word from the list, they could "select" the corresponding number in the 8×9 matrix on the next selection by attending to the flashes of the desired number. In Figure 1B, once the "y" had been selected the WordQ2 window generates the word "your" as choice 1. Thus, to select the word "your" the participant would select the number 1 from the matrix. Upon selecting the 1 from the matrix, WordQ2 would type the remaining characters "our" and a space, thus completing the word in the output window. At this time, WordQ2 would populate with the seven most probable words. If the participant's target word did not appear in the WordQ2 list, it was necessary to provide additional characters until the word appeared in the predictive window, or it was completed. As every participant was spelling the same sentence, the learning vocabulary feature of WordQ2 was disabled to prevent the program from listing each target word after a single selection. In the event that a word was incorrectly selected (e.g., 2 was selected instead of 1), the participant could select Escape (Esc) from the matrix and WordQ2 would undo the selection. Thus, returning the participant

to the previous location in the sentence. However, if a participant was attending to Esc and the resulting selection was incorrect, the participant was required to backspace all of the incorrect characters individually (a limitation of WordQ2 for the current application). In this way, a predictive speller can provide powerful correct selections with time savings and powerful errors with time losses.

Not all errors required a correction. Under certain conditions, the predictive speller also corrected misspelled words. For example, if the output window read "plos" the predictive speller would still list "please" as one of the options and would correct the errors if "please" was selected. If End or RtArw was selected the cursor in the output window would not move; it only cost the participant a single selection. Participants were not required to correct an error if F5 was selected. In this case, a date/time stamp would appear in the Notepad window. The participant was asked ignore the mistake and attend to the next selection. Once this error was observed it was addressed by changing F5 to F6 in the matrix, which has no output in Notepad thus keeping the error and correct selection count consistent across all participants.

2.3 Sentence selection

The length of the sentence is typical of a moderately easy sentence in English, the selected words are representative of the mean length of words in English, and five of the 10 words are in the 200 most common English words (Brysbaert & New, 2009). Thus, the sentence used in the online test phase was made up of 50% of the 200 most commonly used words in the English language.

2.4 Data acquisition, processing

Participants were seated in a chair approximately 1meter from a computer monitor that displayed an 8×9 matrix of letters, numbers, and other keyboard commands. A 72-item speller matrix was used because it is similar to the one designed for home use (Sellers et al., 2010). Moreover, larger matrices have been shown to increase P300 amplitude as the probability of the desired item is reduced (Allison & Pineda, 2003; Sellers, Krusienski, McFarland, Vaughan, & Wolpaw, 2006).

Electroencephalograph (EEG) was recorded with a 32-channel electrode cap embedded with tin electrodes (Electro-Cap International, Inc.). All channels were referenced to the right mastoid and grounded to the left mastoid. Impedance on each channel was reduced below 10.0 k before testing began. Two g.tec (Guger Technologies) 16-channel biosignal amplifiers (version 2) were used. The amplifiers have a +/−250 mV input sensitivity and are amplified to +/−2 V before the ADC converts the signals to digital format. Signals were sampled at a rate of 256 Hz, high-pass filtered at 0.5 Hz, and low-pass filtered at 30 Hz. Before analyses EEG data were moving average filtered and downsampled to 20Hz. Thirty-two channels were collected for the possibility of future analysis, but only electrodes Fz, Cz, P3, Pz, P4, PO7, PO8, and Oz (Sharbrough, Lesser, Lüders, Nuwer, Picton, 1991) were used for BCI operation (Krusienski, Sellers, McFarland, Vaughan, & Wolpaw, 2008).

Due to the P300s low signal-to-noise ratio, each item must be flashed multiple times and the results averaged (Cohen & Polich, 1997). During calibration, the number of target item flashes was constant across participants and presentation methods. Item sets of six were flashed in quasi-random groups, with two flashes of each of the 72 items of the matrix flashing twice per sequence, and 10 times in the 5 sequences of each selection. In the calibration phase for the PS and NS conditions, 36 target items were presented; each of the 36 item selections contained 120 flashes (360 targets and 3960 non-targets).

2.5 Classification

Classification coefficients were determined with a stepwise linear discriminate analysis (SWLDA) algorithm (Draper & Smith, 1981) implemented in MATLAB (version 7.6 R2008a, stepwisefit function). The SWLDA algorithm performs forward and backward partial regression procedures to select the spatiotemporal features (i.e., features determined by the combination of electrode location and specific time points during the recording epoch) that account for the most unique variance. Initially, the single feature that accounts for the most unique variance is added to the model (forward regression), then the feature accounting for the most unique remaining variance is added (forward regression). The model is then tested to determine if each feature of the two-feature model still accounts for a significant amount of unique variance (backward regression), if so both features remain in the model and a third is selected. This forward and backward process continues until the model includes the maximum number of features (set to 60) or until no additional features reach the criteria for entry or removal from the model ($p<.10$ for entry and $p>.15$ for removal). SWLDA outputs a set of spatiotemporal classification coefficients that are subsequently applied to the averaged ERP responses during the online phase.

Before the online phase, the number of sequences was optimized for each participant using the maximum written symbol rate (WSR, or symbols/minute; (Furdea et al., 2009; Townsend et al., 2010)). This metric determines the number of item selections a participant can correctly make in one minute, taking into account error correction. Using the WSR, nearly all participants were presented with fewer than five sequences during the online test phase. In theory, the calibration phase should yield equal numbers of sequences for each participant in each paradigm because the calibration tasks are identical for each session. Given our goal of comparing the PS and NS in an unbiased means, we sought to match the number of sequences in the PS and NS conditions. Thus, five of the participants were removed from the study due to having a difference in optimal sequences equal to or greater than two after calibration. Each sequence of flashes requires three seconds; thus, a difference of two or more sequences yields a minimum of six additional seconds per selection. Such a large difference would have confounded the primary goal of the study. By eliminating these five participants the two paradigms were better matched for time and accuracy.

After the matrix flashed the predetermined number of times during online testing, ERPs were averaged for each channel and each cell of the 72-matrix item locations, and then the spatiotemporal coefficients were multiplied by the amplitude value of each model feature. The matrix item with the highest summed score was selected by the classifier and presented to the participant as feedback. The method used was analogous to that used by Krusienski et al, (2008), with the exception that eight channels were used.

The present experimental paradigm derived a classifier for each session independently because within participant differences between sessions could influence performance. For example, if a participant has had a variable amount of sleep or caffeine it is possible that such variables would affect attentional processes and waveform morphology. In addition, removing and replacing the cap may result in electrodes being located at slightly different locations, contributing to deleterious effects on classification performance in the subsequent session. Thus, performing two calibration sessions should have provided classifiers best suited for a given session.

2.6 Dependent measures

Accuracy was measured by taking the number of correct selections (i.e. feedback matched the character to which the participant was attending) and dividing this value by the total number of selections per session. The formula for calculating bit rate described by Pierce

(1980) incorporates the number of possible targets (*N*) and the probability that the target is accurately classified (*P*):

$$\text{Bit Rate} = \log_2 N + P + (1 - P)\log_2(1 - P/N - 1)$$

(1)

the result divided by number of minutes in a session yields bits per minute. The calculation "selections per minute" was performed by taking the total number of selections and dividing by the total time of the session. "Output characters per minute" (OCM) was calculated by taking the 58 total selections in each session (including sleep) and dividing it by the total time of the PS session. OCM was used to calculate the contribution of the predictive speller program. This calculation includes the time it took for the participant to correct errors while the number of correct target selections (58) remained static. Therefore, the more errors a participant made the more time it took to finish the session, resulting in lower output characters per minute. However, PS and NS selections per minute were a direct result of sets per sequence and time, thus not affected by error correction.

3. Results

A 2×2 mixed model analysis of variance (Order: (NS first vs. PS first) X Condition: (NS vs. PS)) was used to examine if an order effect was present in the data. The results provided insufficient evidence to reject the null hypothesis (F (1, 22) = 0.185, p = 0.671). Thus, we collapsed across the conditions and analyzed the data using paired *t*-tests to examine the differences between the PS and NS conditions on the measure of mean accuracy, selections per minute, bit rate, theoretical bit rate, output characters per minute, and waveform latency and amplitude.

3.1 Online Accuracy, bit rate, and theoretical bit rate

Table 1 shows raw scores and means for accuracy, bit rate and theoretical bit rate. Online accuracy was significantly higher for NS, (M = 89.80%, SD = 7.78) than for the PS, (M = 84.88%, SD = 10.59), t (23) = 2.15, p = 0.04, d = 0.40. We suspect lower accuracy in PS is attributed to the higher workload and/or dual task processing requirements of the PS paradigm. In addition, we found marginal differences between PS bit rate and NS bit rate (M = 17.71, SD = 5.38, M = 19.39, SD = 5.37, respectively), t (23) = 2.04, p = 0.053, d = 0.39. Theoretical bit rate (i.e., bit rate with the time between selections removed) is presented for comparison to studies that report bit rate with the time between selections removed, in this study six seconds were provided between each item selection.

3.2 Selections per minute

Table 2 shows raw scores and means for PS and NS sets per sequence, time to complete the sentence, selections per minute and OCM. We compared means of PS selections per minute against NS selections per minute (M = 3.71, SD = 0.75, M = 3.76, SD = 0.75, respectively) and found no difference between groups, t (23) = 0.49, p = .62, d = 0.10. Although this comparison provided null findings, when compared to OCM significant differences were revealed. OCM was significantly higher than PS selections per minute (M = 5.28, SD = 1.67), t (23) = 6.05, p < .001, d = 0.78. Similarly, OCM was significantly higher than NS selections per minute, t (23) = 5.61, p < .001, d = 0.76. Moreover in total time to complete the sentence (in minutes), the PS was significantly faster than the NS paradigm (M = 12.43, SD = 4.96, M = 20.20, SD = 5.98, respectively), t (23) = 7.52, p < .001, d = 0.84.

3.3 Waveform Morphologies

The PS and NS produced virtually identical waveforms. Our analyses focused on the electrodes Cz, Pz, Po7, and Po8 because most of the P300 amplitude change in BCI applications is captured in these four electrodes (Kaper, Meinicke, Grossekathoefer, Lingner, & Ritter, 2004; Krusienski et al., 2008). Figure 2A shows average target waveforms for each of the 24 participants. Figure 2B shows the grand mean waveforms for the target waveforms (top row) and the non-target waveforms (bottom row). The difference in the positive peak at electrode location Cz around 200 ms was marginally higher in the NS than in the PS paradigm (M = 3.45, SD = 1.47, M = 2.82, SD = 1.71, respectively), t (23) = 2.06, p = .051, d = 0.39. Additionally, the NS peak at electrode location Pz around 200 ms was significantly larger than the PS peak (M = 3.82, SD = 1.49, M = 3.24, SD = 1.81, respectively), t (23) = 2.34, p = .028, d = 0.43.

4. Discussion

The primary goal of this study was to test the efficiency of a predictive speller program in conjunction with a P300 BCI. The main hypotheses were that the predictive speller should improve overall character output and possibly affect waveform morphology. The first hypothesis was supported, even though accuracy was significantly lower in the PS paradigm, and bit rate and selections per minute were statistically equivalent in both paradigms. Despite the NS advantage in accuracy, the PS showed an average time advantage of 7minutes and 46 seconds over the NS, and OCM were significantly higher for the PS than the NS by 1.51 characters/minute. Given the current maximum character selection rate of approximately four selections per minute in P300 BCIs (this work also see (Lenhardt, Kaper, & Ritter, 2008; Townsend et al., 2010), these results impressively convert to an additional 91.2 output characters per hour, or nearly one and a half per minute. These results suggest that a predictive speller can provide a substantial advantage to an individual communicating via a P300 Speller in an online environment.

The significant difference in accuracy between the two paradigms may be a result of increased workload and/or task difficulty associated with the PS. This hypothesis is indirectly supported by the finding of lower amplitude responses in the PS condition at the Cz and Pz electrode locations. Previous P300 research has shown that workload (i.e., the measure of the interaction between task difficulty and an individual's ability to perform a given task (Gopher & Donchin, 1986), and dual task interference can significantly reduce P300 amplitude and increase P300 latency (Gopher & Donchin, 1986; Isreal, Chesney, et al., 1980; Isreal, Wickens, et al., 1980; Kramer, Wickens, & Donchin, 1983; Kramer et al., 1985; Wickens et al., 1983). The relatively small amplitude differences in the current study may be due to the fact that the increase in workload was discontinuous (i.e., increased during the time in which target stimuli were not flashing). This is in contrast to studies investigating workload which typically use continuous increases in task demands (e.g., tracking a stimulus). In addition, the AAC literature also suggests that cognitive demand is increased when a predictive speller is used (Koester & Levine, 1994; Venkatagiri, 1994).

As this study used naïve participants, we believe that with training PS accuracy will increase, thus increasing OCM. Gopher and Donchin (1986) suggest that the effects of workload decrease with practice. In addition, the predictive speller can learn to adapt to the individual over time, which we did not allow in the current study.

Further support of the inefficiency of the naïve participants to use a predictive speller is shown by the number of selections required for an ideal user to complete the sentence; only 31 selections were necessary using the untrained predictive speller. However, many

participants failed to select a word from the predictive speller at the first opportunity, leading to additional unnecessary selections.

5. Conclusions

These results demonstrate the potential efficacy of predictive spelling in the context of BCI. Future research should be conducted in an ALS population to determine if similar improvements in output character selections are obtained.

Acknowledgments

References

Allison BZ, Pineda JA. ERPs evoked by different matrix sizes: implications for a brain computer interface (BCI) system. IEEE Trans Neural Syst Rehabil Eng 2003;11(2):110–113. [PubMed: 12899248]

Brysbaert M, New B. Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. Behav Res Methods 2009;41(4):977–90. [PubMed: 19897807]

Cohen JD, Polich J. On the number of trials needed for P300. International Journal of Psychophysiology 1997;25(3):6.

Copestake, A. Augmented and alternative NLP techniques for augmentative and alternative communication. Paper presented at the Natural Language Processing for Communication Aids; 1997.

Draper, NR.; Smith, H. Applied regression analysis. 2. New York: Wiley; 1981.

Farwell LA, Donchin E. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroencephalogr Clin Neurophysiol 1988;70(6):510–523. [PubMed: 2461285]

Furdea A, Halder S, Krusienski DJ, Bross D, Nijboer F, Birbaumer N, et al. An auditory oddball (P300) spelling system for brain-computer interfaces. Psychophysiology 2009;46(3):617–625. [PubMed: 19170946]

Garay-Vitoria N, Abascal J. A comparison of prediction techniques to enhance the communication rate. User-Centered interaction paradigms for universal access in the information society 2004:18.

Garay-Vitoria N, Abascal J. Text prediction systems: a survey. Universal Access in the Information Society 2006;4:20.

Gopher, D.; Donchin, E. Workload: An examination of the concept. In: Lloyd, KRK Boff; Thomas, James P., editors. Handbook of Perception and Human Performance, Vol. 2: Cognitive Processes and Performance. Oxford: John Wiley & Sons; 1986. p. 1-49.

Guger C, Daban S, Sellers E, Holzner C, Krausz G, Carabalona R, et al. How many people are able to control a P300-based brain-computer interface (BCI)? Neurosci Lett. 2009

Isreal JB, Chesney GL, Wickens CD, Donchin E. P300 and tracking difficulty: evidence for multiple resources in dual-task performance. Psychophysiology 1980;17(3):259–273. [PubMed: 7384376]

Isreal JB, Wickens CD, Chesney GL, Donchin E. The event-related brain potential as an index of display-monitoring workload. Hum Factors 1980;22(2):211–224. [PubMed: 7390505]

Kaper M, Meinicke P, Grossekathoefer U, Lingner T, Ritter H. BCI Competition 2003--Data set IIb: support vector machines for the P300 speller paradigm. IEEE Trans Biomed Eng 2004;51(6): 1073–1076. [PubMed: 15188881]

Koester HH, Levine SP. Learning and performance of able-bodied individuals using scanning systems with and without word prediction. Assist Technol 1994;6(1):42–53. [PubMed: 10147209]

Koester HH, Levine SP. Modeling the speed of text entry with a word prediction interface. IEEE Transactions on Rehabilitation Engineering 1994;2(3):10.

Koester HH, Levine SP. Effect of a word prediction feature on user performance. Augmentative and Alternative Communication 1996;12(3):23.

Kramer AF, Wickens CD, Donchin E. An analysis of the processing requirements of a complex perceptual-motor task. Hum Factors 1983;25(6):597–621. [PubMed: 6671646]

Kramer AF, Wickens CD, Donchin E. Processing of stimulus properties: evidence for dual-task integrality. J Exp Psychol Hum Percept Perform 1985;11(4):393–408. [PubMed: 3161983]

Krusienski DJ, Sellers EW, McFarland DJ, Vaughan TM, Wolpaw JR. Toward enhanced P300 speller performance. J Neurosci Methods 2008;167(1):15–21. [PubMed: 17822777]

Kubler A, Nijboer F, Mellinger J, Vaughan TM, Pawelzik H, Schalk G, et al. Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. Neurology 2005;64(10):1775–1777. [PubMed: 15911809]

Kunst CB. Complex genetics of amyotrophic lateral sclerosis. Am J Hum Genet 2004;75(6):933–947. [PubMed: 15478096]

Lenhardt A, Kaper M, Ritter HJ. An adaptive P300-based online brain-computer interface. IEEE Trans Neural Syst Rehabil Eng 2008;16(2):121–130. [PubMed: 18403280]

Lesher, G.; Rinkus, G. Domain-specific word prediction for augmentative communication. Paper presented at the RESNA Annual Conference; 2002.

Murphy JM, Henry RG, Langmore S, Kramer JH, Miller BL, Lomen-Hoerth C. Continuum of frontal lobe impairment in amyotrophic lateral sclerosis. Arch Neurol 2007;64(4):530–534. [PubMed: 17420314]

Nijboer F, Sellers EW, Mellinger J, Jordan MA, Matuz T, Furdea A, et al. A P300-based brain-computer interface for people with amyotrophic lateral sclerosis. Clin Neurophysiol 2008;119(8): 1909–1916. [PubMed: 18571984]

Picton TW. The P300 wave of the human event-related potential. J Clin Neurophysiol 1992;9(4):456–479. [PubMed: 1464675]

Pierce, JR. An Introduction to Information Theory. Dover; New York: 1980. p. 145-165.

Ritter W, Vaughan HG Jr. Averaged evoked responses in vigilance and discrimination: a reassessment. Science 1969;164(3877):326–328. [PubMed: 5776645]

Schalk G, McFarland DJ, Hinterberger T, Birbaumer N, Wolpaw JR. BCI2000: a general-purpose brain-computer interface (BCI) system. IEEE Trans Biomed Eng 2004;51(6):1034–1043. [PubMed: 15188875]

Sellers EW, Donchin E. A P300-based brain-computer interface: initial tests by ALS patients. Clin Neurophysiol 2006;117(3):538–548. [PubMed: 16461003]

Sellers EW, Krusienski DJ, McFarland DJ, Vaughan TM, Wolpaw JR. A P300 event-related potential brain-computer interface (BCI): the effects of matrix size and inter stimulus interval on performance. Biol Psychol 2006;73(3):242–252. [PubMed: 16860920]

Sellers EW, Vaughan TM, Wolpaw JR, et al. A brain-computer interface for long-term independent home use. Amyotrophic Lateral Sclerosis. (in press).

Sharbrough FCG, Lesser RP, Lüders H, Nuwer M, Picton W. AEEGS guidelines for standard electrode position nomenclature. Clinical Neurophysiology 1991;8:202–204.

Townsend GT, LaPallo BK, Boulay C, Krusienski DJ, Frye GE, Hauser CK, Schwartz NE, Vaughan TM, Wolpaw JR, Sellers EW. A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns. Clinical Neurophysiology. 201010.1016/j.clinph. 2010.01.030

Trnka K, McCaw J, Yarrington D, McCoy KF, Pennington C. User interaction with word prediction: the effects of prediction quality. ACM Transactions on Accessible Computing 2009;1(3):34.

Vaughan TM, McFarland DJ, Schalk G, Sarnacki WA, Krusienski DJ, Sellers EW, et al. The Wadsworth BCI Research and Development Program: at home with BCI. IEEE Trans Neural Syst Rehabil Eng 2006;14(2):229–233. [PubMed: 16792301]

Venkatagiri HS. Effect of window size on rate of communication in a lexical prediction AAC system. AAC Augmentative and Alternative Communication 1994 June;10:8.

Wickens C, Kramer A, Vanasse L, Donchin E. Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. Science 1983;221(4615):1080–1082. [PubMed: 6879207]

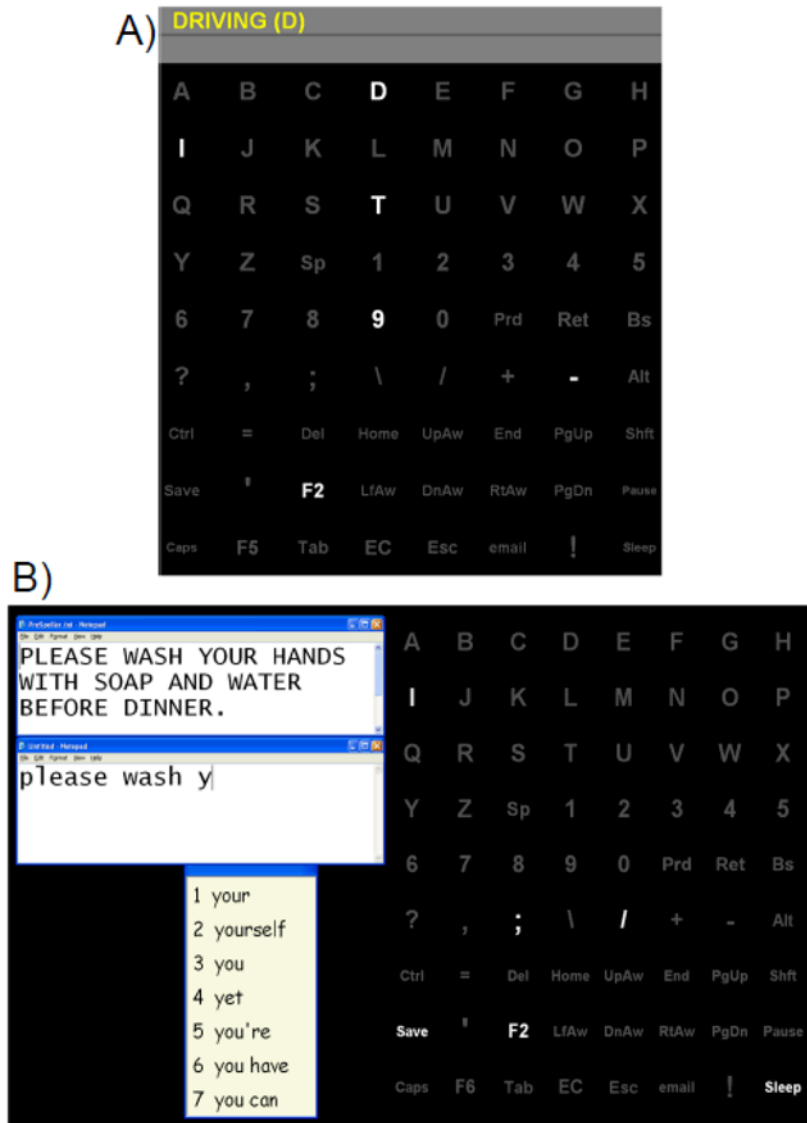Wolpaw, JR.; Birbaumer, N. Brain Communication Interfaces for Communication and Control. 2006. p. 602-614.

Figure 1.
A) The 8×9 matrix used during the calibration phase and online spelling phase of the experiment. In this example, the target item "D" is noted by the letter in parentheses at the end of the word. Participants are instructed to count the number of time the target item flashes. After all items have flashed a predetermined number of times, there is a 3.5 second pause in which the item in parentheses changes to the next letter of the word to indicate the next target item. B) The 8×9 matrix and additional windows used during the online spelling phase of the experiment. Right: the flashing matrix used to make item selections. Left top: the sentence target window. Left middle: the sentence output window. Left bottom: the predictive spelling window used in the PS condition (see text for details).
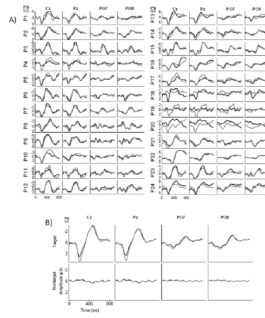
Figure 2.
A) Target waveforms for electrode locations Cz, Pz, Po7, and Po8 for each of the 24 participants; PS paradigm data are presented in black and NS paradigm data are presented in grey. (Amplitude units are µV.) B) Grand mean waveforms for all 24 participants at electrode locations Cz, Pz, Po7, and Po8. The top row consists of target responses for both paradigms, and the bottom row consists of non-target responses for both paradigms. PS data is presented in black and NS data is presented in grey.

Table 1

Online test phase accuracy (Acc), bit rate (BR), and theoretical bit rate (Theo BR) for the predictive speller (PS) and non-predictive speller (NS).

Subject	PS Acc	NS Acc	PS BR	NS BR	PS Theo BR	NS Theo BR
1	96.88	95.31	23.70	28.26	39.33	56.09
2	88.89	87.50	19.93	19.54	32.62	32.38
3	70.00	88.16	11.48	16.46	17.11	24.58
4	79.59	89.86	18.78	20.41	33.52	33.82
5	91.89	92.65	17.71	15.39	26.33	21.50
6	87.18	95.31	21.73	22.58	38.70	37.39
7	91.67	100.00	21.21	24.85	34.96	41.13
8	81.13	87.50	15.79	21.72	24.66	38.86
9	80.95	70.83	17.35	17.60	28.64	35.05
10	82.35	98.33	22.28	29.98	44.12	59.45
11	80.00	91.18	12.11	14.91	16.87	20.79
12	77.59	82.50	11.55	12.69	16.10	17.70
13	82.22	93.94	17.61	22.00	28.91	36.45
14	94.29	77.17	22.01	14.57	36.00	22.81
15	91.18	95.31	19.10	20.51	29.70	32.05
16	94.29	85.25	18.52	15.62	27.51	23.29
17	72.50	77.23	8.18	11.45	10.48	15.98
18	91.89	100.00	26.69	31.12	52.65	61.70
19	100.00	100.00	25.00	24.85	41.13	41.13
20	96.88	91.18	21.25	19.00	33.01	29.70
21	86.67	91.43	16.06	14.95	23.92	20.82
22	57.58	83.67	5.02	15.13	6.19	22.62
23	67.07	96.77	11.80	16.55	18.46	23.06
24	94.44	84.15	20.27	15.28	31.54	22.82
Mean	84.88	89.80	17.71	19.39	28.85	32.13
StDev	10.59	7.78	5.38	5.39	10.95	12.83
SE	2.16	1.59	1.10	1.10	2.24	2.62

Table 2

Online test phase sets per sequence (Sets/Seq), time to complete the sentence (Comp(min)), and selections per minute (Sel/min) in the predictive speller (PS) and non-predictive speller (NS) paradigms, and the predictive output characters per minute (OCM).

Subject	PS Sets/Seq	NS Sets/Seq	PS Comp(min)	NS Comp(min)	PS Sel/min	NS Sel/min	PS OCM
1	3.00	2.00	7.80	12.70	4.10	5.04	7.44
2	3.00	3.00	9.00	17.90	4.00	4.02	6.44
3	4.00	4.00	24.00	22.70	3.33	3.35	2.42
4	2.50	3.00	10.92	17.15	4.49	4.02	5.31
5	4.00	5.00	11.00	23.58	3.36	2.88	5.27
6	2.50	3.00	8.67	15.90	4.50	4.03	6.69
7	3.00	3.00	8.90	14.40	4.04	4.03	6.52
8	3.50	2.50	14.47	16.10	3.66	4.47	4.01
9	3.00	2.00	10.40	23.90	4.04	5.02	5.58
10	2.00	2.00	10.10	11.90	5.05	5.04	5.74
11	5.00	5.00	19.15	23.70	2.87	2.87	3.03
12	5.00	5.00	20.20	27.90	2.87	2.87	2.87
13	3.00	3.00	11.25	16.40	4.00	4.02	5.16
14	3.00	3.50	8.75	25.20	4.00	3.65	6.63
15	3.50	3.50	9.25	17.50	3.68	3.66	6.27
16	4.00	4.00	10.40	18.20	3.37	3.35	5.58
17	5.00	5.00	17.75	35.25	2.25	2.87	3.27
18	2.00	2.00	7.30	11.50	5.07	5.04	7.95
19	3.00	3.00	7.65	14.40	4.05	4.03	7.58
20	3.50	3.50	8.70	18.60	3.68	3.66	6.67
21	4.00	5.00	13.40	24.45	3.36	2.86	4.33
22	3.50	4.00	16.95	29.30	1.95	3.34	3.42
23	3.50	5.00	22.45	21.60	3.65	2.87	2.58
24	3.50	4.00	9.80	24.50	3.67	3.35	5.92
Mean	3.42	3.54	12.43	20.20	3.71	3.76	5.28
StDev	0.830	1.062	4.963	5.978	0.745	0.749	1.666
SE	0.169	0.217	1.013	1.220	0.152	0.153	0.340