International Journal of Human-Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: https://www.tandfonline.com/loi/hihc20

A Chinese Text Input Brain–Computer Interface Based on the P300 Speller

James W. Minett , Hong-Ying Zheng , Manson C-M. Fong , Lin Zhou , Gang Peng & William S-Y. Wang

To cite this article: James W. Minett , Hong-Ying Zheng , Manson C-M. Fong , Lin Zhou , Gang Peng & William S-Y. Wang (2012) A Chinese Text Input Brain–Computer Interface Based on the P300 Speller, International Journal of Human-Computer Interaction, 28:7, 472-483, DOI: 10.1080/10447318.2011.622970

To link to this article: https://doi.org/10.1080/10447318.2011.622970

Accepted author version posted online: 03 Oct 2011.
Published online: 03 Oct 2011.

Intl. Journal of Human–Computer Interaction, 28: 472–483, 2012

ISSN: 1044-7318 print / 1532-7590 online
DOI: 10.1080/10447318.2011.622970

A Chinese Text Input Brain–Computer Interface Based on the P300 Speller

James W. Minett, Hong-Ying Zheng, Manson C-M. Fong, Lin Zhou, Gang Peng, and William S-Y. Wang
The Chinese University of Hong Kong, Hong Kong, China

A visual speller is a brain–computer interface that empowers users with limited motor functionality to input text into a computer by measuring their electroencephalographic responses to visual stimuli. Most prior research on visual spellers has focused on input of alphabetic text. Adapting a speller for other types of segmental or syllabic script is straightforward because such scripts comprise sufficiently few characters that they may all be displayed to the user simultaneously. Logographic scripts, such as Chinese *hanzi*, however, impose a challenge: How should the thousands of Chinese characters be displayed to the user? Here, we present a visual speller, based on Farwell and Donchin's P300 Speller, for Chinese character input. The speller uses a novel shape-based method called the First–Last, or FLAST, method to encode more than 7,000 Chinese characters. Characters are input by selecting two components, from a set of 56 distinct components, that match the shape of the target character, followed by selection of the character itself. At the input speed of one character per 107 s, 24 able-bodied participants achieved mean online accuracy of 82.8% per component selection and 63.5% per character input. At the faster input speed of one character per 77 s, mean online accuracy was 59.4% per component selection and 33.3% per character input.

1. INTRODUCTION

A substantial minority of people have motor deficits that limit or destroy their ability to communicate by conventional means, that is, by speech, by writing, or by sign language. Various neuromuscular disorders can cause such severe motor impairment, including amyotrophic lateral sclerosis, multiple sclerosis, spinocerebellar ataxia, cerebral palsy, and brainstem stroke. Brain–computer interfaces (BCIs) provide affected individuals an alternative channel that can augment their ability to communicate with other people. Typically, the input signals used to drive BCIs are the fluctuations in electrical potential elicited by cognitive activity in response to sensory input, measured by electrodes placed on the scalp (i.e., electroencephalography [EEG]; e.g., Farwell & Donchin, 1988; Townsend et al., 2010), or by electrode grids implanted directly onto the cortical surface of the brain (i.e., electrocorticography; e.g., Brunner, Ritaccio, van Erp, Aloise, & Cincotti, 2011; Leuthardt, Schalk, Wolpaw, Ojemann, & Moran, 2004). Many BCIs for augmentative communication elicit brain signals using visual stimuli (e.g., Farwell & Donchin, 1988; Guger et al., 2009; Jin et al., 2010; Minett, Peng, Zhou, Zheng, & Wang, 2010; Sellers, Krusienski, McFarland, Vaughan, & Wolpaw, 2006; Serby, Yom-Tov, & Inbar, 2005; Townsend et al., 2010), but effective use has also been made of both auditory stimuli (e.g., Furdea et al., 2009; Höhne, Schreuder, Blankertz, & Tangermann, 2010; Klobassa et al., 2009) and tactile stimuli (e.g., Brouwer & van Erp, 2010; Brouwer, van Erp, Aloise, & Cincotti, 2010).

In the P300 Speller, Farwell and Donchin's (1988) original visual speller, a 6×6 matrix of characters or symbols is displayed onscreen to the user. The 12 rows and columns of this stimulus matrix are briefly intensified, one by one, in pseudorandom order, forming an oddball sequence (Fabiani, Gratton, Karis, & Donchin, 1987), while the user attends to the particular target character that is to be input. Intensification of a row or column that contains the target tends to elicit a P300 event-related potential (ERP; Sutton, Braren, Zubin, & John, 1965) with greater amplitude than that elicited by intensification of other rows or columns, allowing the user's intended target character to be inferred after it has been intensified one or more times. The original P300 Speller allowed able-bodied users to select 2.3 characters per minute on average at 95% accuracy, corresponding to a mean input rate of about 12 bits/minute (Farwell & Donchin, 1988).

The P300 Speller continues to be the most commonly used BCI system (Mak et al., 2011). Numerous refinements have been proposed that adjust various features of the speller, including matrix size (Allison & Pineda, 2003)

This work was supported by research grants awarded to William S-Y. Wang by the Innovation and Technology Commission, The Government of the Hong Kong SAR (Grant Nos.: ITS/068/09, InP/241/09, and InP/229/09), the Patent Committee of The Chinese University of Hong Kong (Grant No.: TBF/11/ENG/01), the Shun Hing Institute of Advanced Engineering at The Chinese University of Hong Kong (Grant No.: BME 8115020), and the Research Grants Council of Hong Kong (CUHK Grant No.: 2050461).

Address correspondence to James W. Minett, Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong. E-mail: jminett@ee.cuhk.edu.hk

and interstimulus interval (Sellers et al., 2006), background noise and interference color contrast (Nam, Li, & Johnson, 2010), interface type and screen size (Li, Nam, Shadden, & Johnson, 2010), stimulus presentation (Frye, Hauser, Townsend, & Sellers, 2011; Townsend et al., 2010), stimulus motion (Hong, Guo, Liu, Gao, & Gao, 2009), predictive spelling (Ryan et al., 2011), and method of data processing (Krusienski et al., 2006; Serby, Yom-Tov, & Inbar, 2005). The use of the checkerboard paradigm (Townsend et al., 2010), for example, which intensifies sets of six characters that are located in distinct rows and columns of an 8×9 matrix—rather than entire rows or columns—allows able-bodied users to input 4.36 characters per minute at 91.52% accuracy, maintaining a mean practical input rate of 22.59 bits/minute, almost double that of Farwell and Donchin's original speller.

1.1. Nonalphabetic Script

Previous research on the use of visual spellers to augment communication has focused mostly on languages that are written with *alphabetic script*, including English and German. Alphabets consist of characters that map to the phonemes of the corresponding spoken language (Daniels & Bright, 1996), although the mapping is not necessarily one to one.[1] Implementation of a visual speller for other alphabetic scripts (e.g., Greek and Cyrillic) is straightforward because, by appropriately adjusting the size of the stimulus matrix, all the characters in the script can be displayed onscreen to the user simultaneously. Like alphabetic scripts, other types of *segmental script*, particularly *abjads* (e.g., Arabic) and *abugidas* (e.g., Devanagari), as well as *syllabaries* (e.g., Japanese *kana*) also represent the sounds of the corresponding spoken language (Daniels & Bright, 1996) and can also be input directly with a visual speller because they too comprise a small set of characters that may be composed sequentially to form written words. For example, the two *kana* scripts of Japanese, *katakana* and *hiragana*, each comprise 48 characters. In contrast, *logographic scripts*—notably Chinese *hanzi*—consist of large numbers of characters that represent units of meaning, rather than units of sound (Wang, 1981). The most commonly used dictionary of Chinese, Xinhua Zidian (2004), for example, has more than 11,000 entries. It is, therefore, impossible to display simultaneously in the stimulus matrix of a visual speller the many thousands of Chinese characters that the user might wish to input.

To resolve this obstacle to implementation of a visual speller for Chinese text input, we first consider how Chinese text may be input to a computer via a keyboard. There are two broad classes of methods for this purpose: sound-based methods, in which sequences of keystrokes are input to represent the sounds of the spoken language, and shape-based methods, in which sequences of keystrokes are input to represent the visual forms of the written language (Fong & Minett, 2012). Foremost among the sound-based input methods are those using *Hanyu Pinyin* (Yin & Felley, 1990), or just *Pinyin*, the official Romanization of standard Chinese (i.e., Putonghua, which is based on the Beijing dialect of Mandarin Chinese; Gu, 2009) that has been adopted throughout the mainland of China and elsewhere, both to teach Putonghua pronunciation and to transcribe Chinese words into alphabetic script. *Pinyin* uses just the 26 letters of the Roman alphabet to represent the entire inventory of Putonghua consonants and vowels.[2] A Chinese character may be typed using *Pinyin* by entering the sequence of letters that corresponds to its pronunciation. Multiple characters may have the same pronunciation, necessitating an additional selection to type the particular character required. For example, the *Pinyin* sequence 'tian' represents the pronunciation of the Chinese characters 天 ("sky") and 添 ("to append"), as well as many others. Sound-based input methods using the *Pinyin* Romanization may not be suitable for all Chinese users. According to Ethnologue (Lewis, 2009), the web-based resource maintained by the Summer Institute of Linguistics, about 70% of the Chinese population speaks a dialect of Mandarin, about 80% of whom speak Putonghua, the dialect upon which *Pinyin* is based. Of the remaining population, the majority speak related—but mutually unintelligible—dialects of Han Chinese (i.e., Sinitic; Van Driem, 2001).

In contrast to sound-based methods, shape-based input systems specify a set of forms to represent the visual appearance of the target character. The Chinese character, or *sinogram* (Wang & Tsai, 2011), has a hierarchical structure, consisting of one or more strokes with particular spatial arrangement (Wang, 1981). For example, the sinogram 天 ("sky") is made up of four strokes: two horizontal strokes, one leftward going stroke, and one rightward going stroke. The same four strokes may be recombined with different spatial arrangement to form the distinct character 夫 ("husband"). Certain spatial arrangements of strokes, which we refer to here as *components*, reoccur with great frequency in the Chinese lexicon—for example, each of the two sinograms just stated contains the two components 人 and 二 (which are themselves both sinograms). Popular shape-based input methods include *wubizixing* (or just *wubi*), used mainly in the Chinese mainland, and *simplified Cangjie*, more commonly used in Hong Kong. The simplified Cangjie method, for example, uses a set of 24 basic shapes, each shape designating a set of stroke patterns with similar spatial arrangements, each mapped to a key on the keyboard. This method requires two basic shapes to be selected, followed by selection of a particular

[1]The lack of a necessary one-to-one correspondence between alphabetic characters and phonemes is exemplified by the letters *-ough*, which represent various combinations of phonemes, such as /ʌf/, /ɒf/, /uː/, /ɔː/, /aʊ/, or /əʊ/, depending on the English word in which they appear.

[2]In addition to consonants and vowels, the Chinese syllable comprises lexical tone: the use of distinct pitch patterns to provide lexical contrast among syllables that are otherwise pronounced similarly. Pinyin uses a set of optional diacritics to mark the lexical tone of each syllable.

sinogram from among a list of sinograms that match the chosen basic shapes. For example, the sinogram 天 is represented by the basic shapes 一 and 大, which are selected by the keys "m" and "k," respectively. This same encoding also represents a number of other sinograms, including 更 ("change") and 政 ("government").

1.2. BCIs for Chinese Text Input

Recently, attempts have been made to develop BCIs for Chinese text input (Jin et al., 2010; Minett et al., 2010; Wu et al., 2009), all using shape-based methods rather than sound-based methods. Although the standard P300 Speller could be adapted relatively simply to input characters by selecting appropriate *Pinyin* sequences, developing a universal system for application throughout China would be challenging—if not impossible—because, as noted previously, *Pinyin* reflects the pronunciation of Putonghua, based on the Beijing dialect of Mandarin Chinese, which differs considerably from the pronunciations of other Han Chinese dialects.

In Jin et al. (2010), the BCI makes use of an existing input system—the T9 stroke input system (Tegic Communications, Seattle, WA)—built into a mobile phone. A four-by-four stimulus matrix is used to display a mixture of single strokes, digits, and controls. The system distinguishes five distinct types of stroke: horizontal, vertical, leftward going, rightward going or dot, and hook, as does the system reported in (Wu et al., 2009). The rows and columns of the stimulus matrix are intensified in pseudo-random sequence while the user attends to the item to be selected, as in Farwell and Donchin's P300 Speller. A sinogram is input by selecting the sequence of strokes by which that sinogram may be written. After selection of each stroke, a shortlist of seven sinograms is displayed that matches the current stroke sequence. The user may then (a) input one of these seven sinograms by selecting the corresponding digit (1–7), (b) request that more matching sinograms be displayed by selecting the digit 8, (c) select an additional stroke, or (d) delete the previous selection. Using Bayesian linear discriminant analysis and particle swarm optimization, this system delivered offline stroke input accuracy ranging from 38.67% up to 100% by averaging a block of fifteen 2-s trials from each of 11 able-bodied participants. One participant maintained 100% accuracy by averaging four trials, achieving an information transfer rate (Lenhardt, Kaper, & Ritter, 2008) of 40.34 bits/minute, equivalent to about eight stroke selections per minute. Based on the authors' estimate that six to seven stroke selections are required to input a sinogram, the online system may allow slightly better than one sinogram to be input per minute by the most proficient users. However, the online performance of this system has yet to be validated (Jin et al., 2010).

1.3. The Present Study—The FLAST Speller

As just noted, users of the system in (Jin et al., 2010) were required to make six to seven selections on average in order to input one sinogram. In online operation, an incorrect selection in any one of those six to seven selections would necessitate deletion of the erroneous selection, followed by an additional attempt to make the intended selection. For users who can maintain high selection accuracy (e.g., one of their 11 participants maintained mean classification accuracy above 90%), there is no great negative consequence. However, for individuals unable to maintain high selection accuracy (e.g., two of their 11 participants achieved mean classification accuracy below 50%), it is likely that repeated correction of errors in stroke selection would greatly increase the time required to input a sinogram correctly. We have therefore sought to develop an alternative method for encoding the Chinese lexicon that allows sinograms to be input in fewer selections.

In this article, we describe and present a performance analysis of a novel visual speller for Chinese text input that requires at most three selections per sinogram. This online system uses a set of 56 shape-based components, shown in Figure 1, extending the 35-component offline system proposed in Minett et al. (2010), to represent a lexicon of 7,072 distinct traditional Chinese sinograms.[3] Each sinogram is represented by two components: the *first component*, representing the pattern of strokes that would be written first when writing the sinogram, and the *last component*, representing the pattern of strokes that would be written last—we therefore call this speller the *First–Last Speller*, or *FLAST*. For example, the sinogram 天 is represented by the components 二 and 人. Of the 56 × 56 possible pairs of components, only 1,601 pairs encode sinograms. Of these, 491 pairs each encode a single sinogram. The 56 components defined in FLAST were chosen for the express purpose

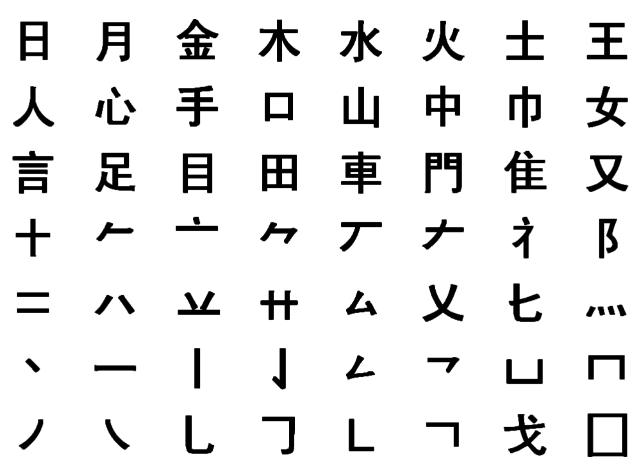

FIG. 1. The 56 shape-based components of the FLAST speller.

[3]There are two sets of sinograms with which Chinese may be written: traditional characters, used mainly in Hong Kong, Macau, and Taiwan, and simplified characters, used in the mainland of China and elsewhere. Simplified characters were introduced in China in 1956 with the aim of enhancing literacy (Wang, 1973).

that no pair would encode more than 56 sinograms—the most numerous pair, 丿 and 口, encodes 41 sinograms, leaving ample room for future expansion of the lexicon.

The FLAST Speller has two modes of operation: a sinogram input mode and a component input mode. Sinogram input proceeds in three stages: (a) selection of the first component, (b) selection of the last component, and (c) selection of the sinogram itself. During each stage, an 8 × 8 stimulus matrix is displayed onscreen, the bottom row of which displays eight system controls: sleep, delete component, delete sinogram, delete sentence, copy, paste, toggle sinogram/alphanumeric input (allowing both sinograms and alphanumeric text to be input), and enter (allowing sinograms that also appear as components in the stimulus matrix, e.g., 二, meaning "two," to be input). During Stages 1 and 2, the first seven rows of the matrix comprise the 56 components. During Stage 3, these rows comprise a set of 56 sinograms, a subset of which is encoded by the pair of components selected during the first two stages; the remaining cells of the stimulus matrix are filled by high-frequency sinograms. The 7,072 sinograms that are encoded in FLAST are distributed across 140 stimulus matrices, with some high-frequency sinograms occurring in multiple matrices. To assist the user to locate target sinograms more easily, the position of each sinogram in the particular stimulus matrix that contains it is held constant.

At the beginning of each stage, the entire stimulus matrix is intensified for a period of time to allow the user time to locate the target component or sinogram, after which the rows and columns of the stimulus matrix are intensified in pseudo-random order. Figure 2 illustrates the FLAST procedure, showing the stimulus matrix at the beginning of Stage 2 after the first component, 𠂇, has been selected in Stage 1 (see Figure 2A), and at the beginning of Stage 3 after the last component, ㇏, has been selected in Stage 2 (see Figure 2B). During sinogram input mode, component selections are displayed between underline-characters (e.g., _𠂇㇏_), to distinguish them from sinogram selections, which are shown without such underlines (e.g., 中文). Once a sinogram has been selected during Stage 3, the component selections are deleted from the screen and the selected sinogram displayed. During component input mode, only components may be input, and they are displayed without underlines (see Figure 2C).

2. METHODS

2.1. Participants

Twenty-four students (12 female, 12 male, M_{age} = 21.7 years, age range = 19–30 years) from The Chinese University of Hong Kong, all native Chinese readers, were paid to participate in this study. All participants were able-bodied, were naive to BCI use, had normal or corrected-to-normal vision, and reported no history of neurological illness. Each participant gave informed consent in compliance with a protocol approved by the Survey and Behavioural Research Ethics Committee of The Chinese University of Hong Kong.

2.2. Data Acquisition

Each participant was seated in a quiet, dimly lit room about 90 cm before a LCD monitor that displayed the experimental stimuli. EEG data were acquired using a 32-channel ActiveTwo EEG system (BioSemi B.V., Amsterdam, The Netherlands) with Ag/AgCl active electrodes at positions Fp1, Fp2, AF3, AF4, Fz, F3, F4, F7, F8, FC1, FC2, FC5, FC6, Cz, C3, C4, T7, T8, CP1, CP2, CP5, CP6, Pz, P3, P4, P7, P8, PO3, PO4, Oz, O1, and O2 (Sharbrough, Lesser, Lüders, Nuwer, & Picton, 1991). Two additional electrodes, *common mode sense* and *driven right leg*, positioned on either side of vertex at positions C1 and C2, respectively, were used to provide a feedback loop to drive the average electrical potential as close as possible to the amplifier reference voltage (BioSemi, 2007). The EEG data were sampled at a rate of 256 Hz, and a 0.1 Hz to 40 Hz digital band-pass filter applied. No artifact correction was applied to deal with eye blinks or eye movements. Stimulus presentation, data collection, and data processing were all managed by the general-purpose BCI software BCI2000 (Schalk, McFarland, Hinterberger, Birbaumer, & Wolpaw, 2004) as follows.

2.3. Experimental Procedure

Each participant completed the experiment in a single session that consisted of two parts: *calibration* followed by *online testing*. During calibration, EEG data were elicited from participants while the FLAST Speller ran in component input mode in order to determine the coefficients of a classifier (described in section 2.5). In each run, participants were required to input a string of five pseudo-randomly selected components that was displayed at the top of the screen (see Figure 2C). The current target component was shown within parentheses to the right of the component string. After a 10-s pause, during which time the entire stimulus matrix was intensified—allowing the participant time to locate the target component and to blink, if necessary—the rows and columns of the matrix were intensified in pseudo-random order. Participants were instructed to attend to the target throughout this period, keeping a mental count of the number of times it was intensified (Farwell & Donchin, 1988). Each intensification was 62.5 ms in duration, followed by a 62.5 ms interstimulus interval during which time no stimulus was intensified. Thus there were eight intensifications per second. For each intensification, an 800-ms segment of EEG data (comprising 205 samples) was extracted from each channel, forming a vector of 6,560 features (32 channels × 205 samples) for use in classification.

In each 2-s trial of 16 intensifications, each row and column was intensified once. After 10 such trials, requiring a total duration of 20 s, the component that was identified using the current classifier was presented onscreen below the component string (see Figure 2C). This procedure was repeated

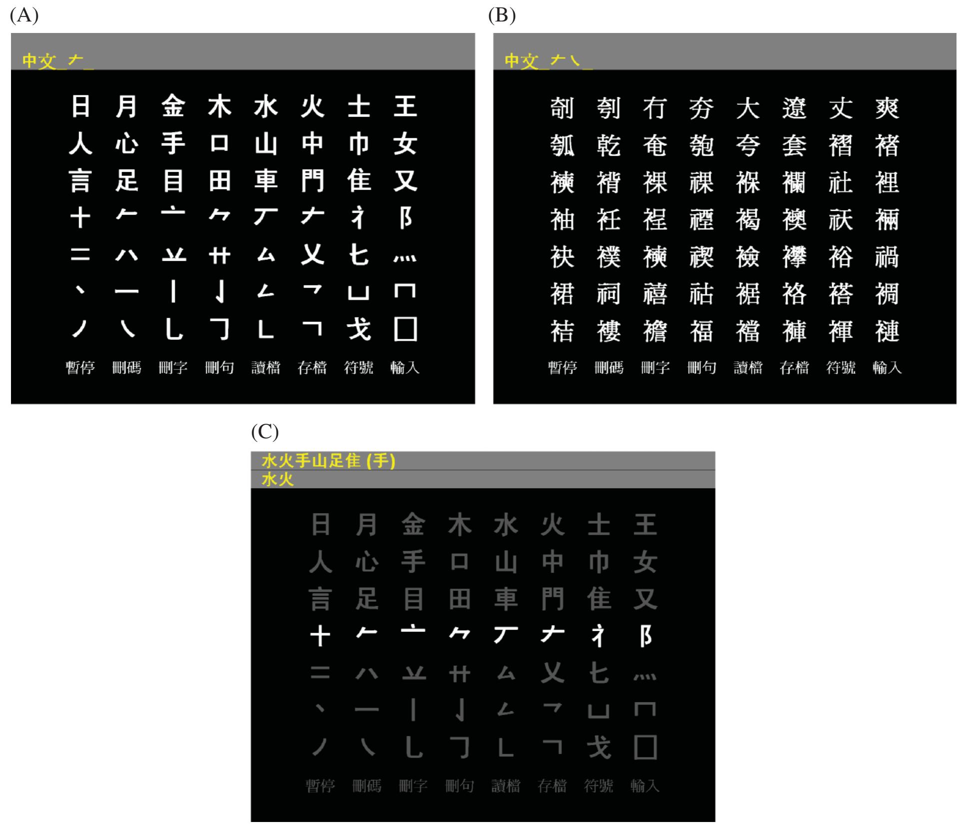

FIG. 2. Illustration of the FLAST procedure. (A) The component matrix at the beginning of Stage 2, showing the intensification of the entire stimulus matrix while the participant locates the target component. (B) An example of the sinogram matrix at the beginning of Stage 3. (C) The component matrix during calibration, showing the intensification of one row of the matrix and the current progress in selecting the target string of five components.

until all components in the string had been presented. After an initial practice run of five components, six calibration runs were carried out, each run lasting 2½ min. After calibration was complete, each participant performed three additional runs in component input mode in order to evaluate offline component selection accuracy, each run requiring five components to be selected, as in the calibration runs.

The participant was then instructed to use the FLAST Speller online in sinogram input mode to input four phrases, each consisting of a four-sinogram string, for example, 燃眉之急 ("of utmost urgency," literally *burning-eyebrow urgency*). Before each run, the participant was shown the target phrase, together with the component encodings of each of the four sinograms, on a handout to which they could refer throughout the run. As summarized in section 1.3 (see also Figures 2A and B), sinogram input proceeded in three stages. Stages 1 and 2 followed the same component selection procedure as in the calibration runs (but without displaying the target components to be input), permitting the first and last component of the target sinogram to be selected. Upon selection of the last component, Stage 3 commenced by displaying for 15 s the intensified sinogram matrix containing the sinograms that were encoded by the component pair identified by the classifier in the previous two stages. This pause allowed the participant ample time to locate the target sinogram and to blink, if necessary. If either of the two component selections was incorrect, the sinogram matrix displayed during Stage 3 would not contain the target sinogram. The participant was not permitted to correct such errors, but rather was instructed to select the sinogram at the top left corner of the sinogram matrix. After the 15-s pause, the sinogram matrix was intensified following the same procedure as for components.

During both component selection (Stages 1 and 2) and sinogram selection (Stage 3), the rows and columns of the matrix were intensified in pseudo-random order for 62.5 ms, followed by a 62.5-ms interstimulus interval. Each trial of 16 intensifications took 2 s. During input of the first two runs, 10 trials were presented per selection (i.e., 20 target intensifications and 140 nontarget intensifications). During input of the second two runs, five trials were presented per selection (i.e., 10 target intensifications and 70 nontarget intensifications). The time required to input each sinogram was 107 s during Runs 1 and 2—comprising 60 s of stimulus matrix intensifications and 47 s of overhead time for interstimulus pauses and target component/sinogram search—and 77 s during Runs 3 and 4—comprising 30 s of intensifications and 47 s of overhead. Four sinograms were input in each run. Runs 1 and 2 were each completed in 7 min 8 s, and Runs 3 and 4 in 5 min 8 s.

2.4. Performance Measures

Accuracy was calculated as the percentage of either components or sinograms that were correctly input. To input a sinogram correctly, the participant was required to make three consecutive correct selections, that is, first component, last component, and then the sinogram itself; an error in any one of these three selections resulted in an incorrect sinogram input. The number of bits transmitted per selection was calculated according to the formula (Wolpaw et al., 2000):

$$B = \log_2 N + P \log_2 P + (1 - P) \log_2 \left[\frac{1 - P}{N - 1}\right]$$

where B denotes the number of bits per selection, N is the total number of possible component or sinogram selections, and P is the proportion of correct selections. The bit rate in bits/minute was then obtained by multiplying B by the number of selections carried out per minute. During component input, 56 distinct components could be selected. During sinogram input, 7,072 distinct sinograms could be input. In addition to bit rate, the theoretical bit rate in bits/minute was calculated by dividing B by the duration in minutes of each selection excluding the 47-s overhead time during online operation for interstimulus pauses and target component/sinogram search.

2.5. Classification

The EEG signals elicited by the speller were classified using stepwise linear discriminant analysis (SWLDA; Draper & Smith, 1981) following the procedure in (Krusienski, Sellers, McFarland, Vaughan, & Wolpaw, 2008). SWLDA is an efficient and commonly used method for binary classification (e.g., Farwell & Donchin, 1988; Ryan et al., 2011; Townsend et al., 2010). The binary classification problem consists of determining a separating hyperplane,

$$w \cdot x - b = 0, \tag{1}$$

where x denotes the feature vector (described in section 2.3), w denotes the feature weights, and b denotes the bias term, with w and b to be estimated by SWLDA. The row and column of the stimulus matrix predicted to contain the target were calculated as those for which the distance of the response from the hyperplane was greatest (Krusienski et al., 2008). SWLDA estimates w and b iteratively by alternating between forward and backward stepwise regression to construct a multiple regression model. In particular, at each forward regression, the statistically most significant predictor (with $p < .1$) was added to the model. At each backward regression, the least significant predictors (with $p > .15$) were removed from the model. This procedure was repeated for 60 iterations, or until no predictors were either added or removed. For more details of the SWLDA method, refer to Krusienski and colleagues (Krusienski et al., 2006; Krusienski et al., 2008).

Calibration of the classifier coefficients was performed offline in two phases, each requiring the participants to perform 15 component selections. This two-phase calibration process was used so that the improved accuracy arising from using the classifier obtained during the first calibration phase would encourage the participant to focus their attention and seek to optimize their performance during the second calibration phase. During both phases, each component selection consisted of 10 trials, comprising data from 20 target intensifications and 140 nontarget intensifications. Classifier coefficients were derived to discriminate the features of the EEG data elicited across all channels by target intensifications from those elicited by nontarget intensifications. Upon completion of each phase, five sets of classifier coefficients were calculated using the *stepwisefit* function in Matlab (version R2010A, MathWorks, Natick, MA), each set obtained by sampling 50% of the EEG data elicited during that phase. In total, 10 min of EEG data elicited from 4,800 intensifications per participant were used for the calibration, of which 600 were target intensifications and 4,200 were nontarget intensifications. Of the five classifiers obtained from the first phase of calibration, the one that provided peak offline accuracy for the smallest number of trials per selection when applied to EEG data obtained during the first phase was used as the classifier during the second phase of calibration. Of the 10 classifiers obtained from the two calibration phases, the one that provided peak offline accuracy for the fewest number of trials per selection when applied to the EEG data obtained during both phases was used in the subsequent online testing. The classification accuracies for each of the ten classifiers obtained during calibration are shown in Figure 3 for all 24 participants—the figures illustrates that the 10 accuracy curves were tightly packed for most participants.

The classifier's spatiotemporal filter obtained during calibration was applied to the EEG data, averaged over a specified number of trials per selection (either 10 trials during Runs 1 and 2 of online testing or five trials during Runs 3 and 4; see section 2.3), to determine the classification score for each row and column of the stimulus matrix. The classification score for each stimulus was then calculated by summing the scores for the row and column containing it. The stimulus with the greatest score was then displayed onscreen (see Figure 2) to the participant to indicate which component or sinogram had been input.

3. RESULTS

Figure 4 shows the map of spatial features obtained during calibration summed across all 24 participants—the figure indicates the proportion of participants for whom each of the 32 channels was included in the respective participant's classifier. In total, 10 channels (F7, P7, Pz, PO3, O1, Oz,

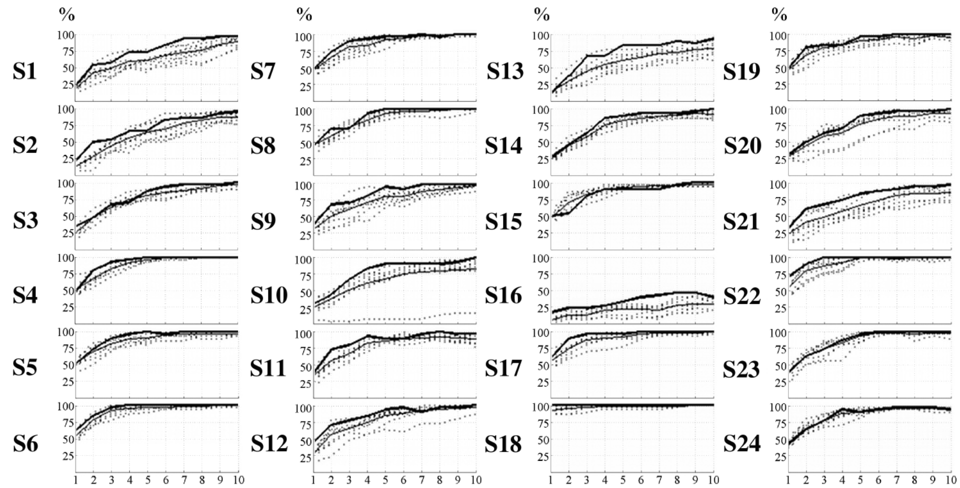

FIG. 3. Offline classification accuracy for each of the 10 classifiers obtained during calibration for each participant. *Note.* The accuracy during calibration of the classifier selected for use in online testing is indicated by the thick solid line; the mean accuracy across all 10 classifiers is indicated by the thin solid line; the accuracy of each of the other nine classifiers is indicated by a dashed line.

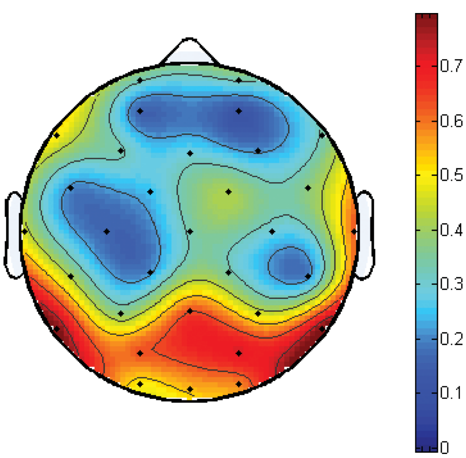

FIG. 4. Map of spatial features summed across all 24 participants. *Note.* The map indicates the proportion of participants for whom each electrode was included in their respective classifiers during online testing (color figure available online).

O2, PO4, P8, T8) contained spatiotemporal features that were selected in the classifiers of at least half of the participants.

The offline accuracy of the FLAST Speller for component input obtained by each of the 24 participants is summarized in Figure 5.

3.1. Online Accuracy and Bit Rate

Table 1 shows the accuracy and bit rate of the FLAST Speller for sinogram input obtained during online testing by each of the 24 participants at 10 trials per selection, corresponding to an input speed of one sinogram per 107 s (i.e., 0.56 sinograms per minute; Runs 1 and 2). The mean component accuracy at this selection rate was 82.8%, delivering 12.93 bits/minute. Five participants (S4, S5, S7, S18, S23) achieved 100% component accuracy. Participants input sinograms with mean accuracy at 63.5%. The mean bit rate was 4.23 bits/minute, with a theoretical bit rate of 7.55 bits per minute (excluding from the calculation the accumulated overhead of 47 s per sinogram for interstimulus pauses and visual search time). Five of the 24 participants achieved 100% sinogram input accuracy at this input speed.

Table 2 shows the online accuracy and bit rate obtained by participants at five trials per selection, corresponding to an input speed of one sinogram per 77 s (i.e., 0.78 sinograms per minute; Runs 3 and 4). The mean component accuracy was 59.4% and the corresponding bit rate was 15.97 bits/minute, slightly faster than for the 10-intensification selection rate. However, sinogram accuracy was 33.3% and the corresponding bit rate was 2.87, lower than for the 10-intensification selection rate. Only one participant (S17) achieved 100% sinogram accuracy at this input speed.

3.2. ERP Waveforms

The ERP waveforms elicited by the FLAST Speller are shown in Figure 6, which displays each participant's averaged waveforms (see Figure 6A) as well as the grand-average waveform (see Figure 6B) for both targets and nontargets at four representative channels: Cz, Pz, P7, and P8. For the majority of participants, there is a clear negative deflection in the parietal channels P7 and P8, peaking about 200 ms after onset of target intensification. For most participants, there is no clear positive deflection after 300 ms (i.e., P300 component), although the grand-averaged waveform does exhibit a positive deflection in parietal channels, peaking at about 300 ms, albeit with negative voltage with respect to baseline, and in the Cz channel, peaking at about 250 ms.

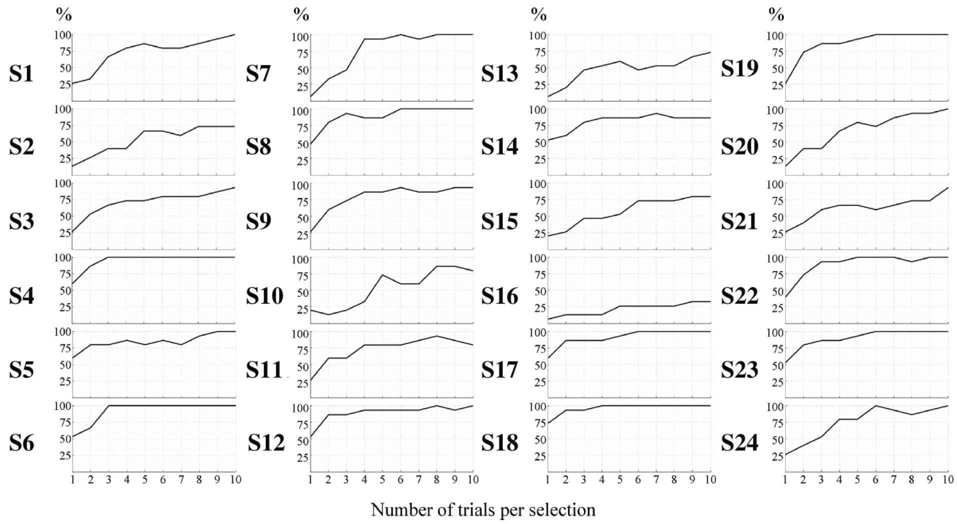

FIG. 5. Offline classification accuracy obtained by each participant during component selection after calibration.

TABLE 1
Online Classification Accuracy and Bit Rate Obtained by Each Participant During Sinogram Input at 10 Trials per Selection (i.e., One Character per 107 s)

	Component Selection		Sinogram Selection		
Participant	Accuracy	Bit Rate	Accuracy	Bit Rate	Theoretical Bit Rate
S1	87.5%	13.62	50.0%	3.02	5.39
S2	68.8%	9.31	25.0%	1.34	2.39
S3	87.5%	13.62	75.0%	4.92	8.78
S4	100.0%	17.42	100.0%	7.17	12.79
S5	100.0%	17.42	100.0%	7.17	12.79
S6	93.8%	15.33	75.0%	4.92	8.78
S7	100.0%	17.42	100.0%	7.17	12.79
S8	81.3%	12.08	50.0%	3.02	5.39
S9	87.5%	13.62	75.0%	4.92	8.78
S10	68.8%	9.31	12.5%	0.59	1.06
S11	68.8%	9.31	12.5%	0.59	1.06
S12	93.8%	15.33	87.5%	5.97	10.65
S13	50.0%	5.75	12.5%	0.59	1.06
S14	81.3%	12.08	62.5%	3.95	7.04
S15	93.8%	15.33	75.0%	4.92	8.78
S16	25.0%	1.98	0.0%	0.00	0.00
S17	93.8%	15.33	87.5%	5.97	10.65
S18	100.0%	17.42	100.0%	7.17	12.79
S19	87.5%	13.62	75.0%	4.92	8.78
S20	75.0%	10.65	37.5%	2.15	3.84
S21	62.5%	8.05	50.0%	3.02	5.39
S22	93.8%	15.33	87.5%	5.97	10.65
S23	100.0%	17.42	100.0%	7.17	12.79
S24	87.5%	13.62	75.0%	4.92	8.78
M	82.8%	12.93	63.5%	4.23	7.55
SD	18.27%	4.03	32.12%	2.38	4.24
SE	3.73%	0.82	6.56%	0.49	0.87

TABLE 2
Online Classification Accuracy and Bit Rate Obtained by Each Participant During Sinogram Input at Five Trials per Selection (i.e., One Character per 77 s)

	Component Selection		Sinogram Selection		
Participant	Accuracy	Bit Rate	Accuracy	Bit Rate	Theoretical Bit Rate
S1	37.5%	7.44	12.5%	0.82	2.11
S2	50.0%	11.50	0.0%	0.00	0.00
S3	37.5%	7.44	12.5%	0.82	2.11
S4	75.0%	21.30	37.5%	2.99	7.68
S5	87.5%	27.25	75.0%	6.84	17.56
S6	43.8%	9.40	0.0%	0.00	0.00
S7	62.5%	16.11	25.0%	1.86	4.77
S8	31.3%	5.62	12.5%	0.82	2.11
S9	62.5%	16.11	37.5%	2.99	7.68
S10	43.8%	9.40	0.0%	0.00	0.00
S11	37.5%	7.44	12.5%	0.82	2.11
S12	75.0%	21.30	50.0%	4.20	10.79
S13	25.0%	3.96	0.0%	0.00	0.00
S14	75.0%	21.30	50.0%	4.20	10.79
S15	81.3%	24.16	50.0%	4.20	10.79
S16	12.5%	1.23	0.0%	0.00	0.00
S17	100.0%	34.84	100.0%	9.96	25.58
S18	93.8%	30.65	87.5%	8.30	21.29
S19	68.8%	18.63	37.5%	2.99	7.68
S20	50.0%	11.50	37.5%	2.99	7.68
S21	43.8%	9.40	0.0%	0.00	0.00
S22	68.8%	18.63	37.5%	2.99	7.68
S23	87.5%	27.25	62.5%	5.48	14.08
S24	75.0%	21.30	62.5%	5.48	14.08
M	59.4%	15.97	33.3%	2.87	7.36
SD	23.39%	9.03	29.64%	2.81	7.22
SE	4.77%	1.84	6.05%	0.57	1.47

4. DISCUSSION

Accurate sinogram input using the FLAST Speller requires the user to make three consecutive selections correctly: first component, last component, and then the sinogram itself. During sinogram input using 10 trials per selection—that is, 0.56 sinograms per minute—the mean component selection accuracy that participants achieved was 82.8%. Based on this figure, the predicted accuracy for sinogram input was $(82.8\%)^3$ that is, 56.8%. The mean sinogram accuracy that participants actually achieved was 63.5%, similar to the predicted accuracy. We infer that participants were able to locate and attend to sinogram targets with similar ease as component targets. Using five trials per selection—that is, 0.78 sinograms per minute—sinogram input accuracy was only 33.3%, too low on average for effective online use, although five participants maintained sinogram input accuracy greater than 50% at this input speed.

To consider ways to enhance performance, it is informative to examine the component selection and sinogram input bit rates. Participants achieved 12.93 bits/per minute while selecting component targets. However, when calculated in terms of correct sinogram inputs, the mean bit rate was only 4.23 bits/minute. This reduction in bit rate had two main causes. First, the time required to input one sinogram included an overhead period of 47 s—for interstimulus pauses and visual search time—during which time no EEG data were processed. Neglecting this overhead period from the calculation of bit rate generated 7.55 bits/minute for the 10-intensification selection rate and 7.36 bits/minute for the five-intensification selection rate. Second, the three-stage FLAST procedure potentially allows $56 \times 56 \times 56$ (i.e., 175,616) distinct sinograms—that is, 12.07 bits of information—to be input. In practice, however, the system allows only 7,072 distinct sinograms—that is, 8.86 bits

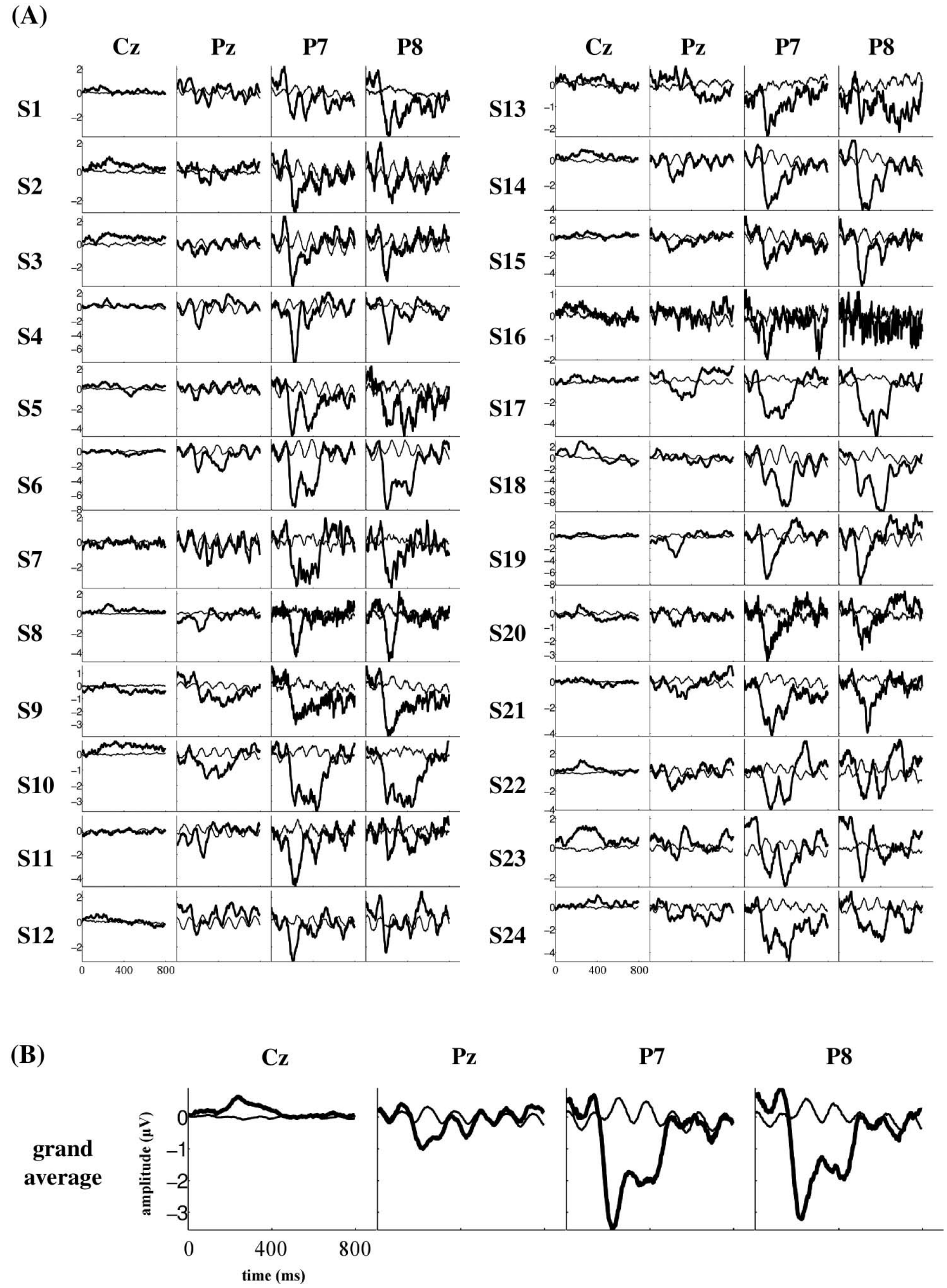

FIG. 6. (A) ERP waveforms elicited from each participant by both target (thick line) and nontarget (thin line) intensifications at four electrode locations: Cz, Pz, P7, and P8. (Amplitude units are μV; time units are ms.) (B) Grand-averaged ERP waveform elicited by both target (thick line) and nontarget (thin line) intensifications for all 24 participants at four electrode locations: Cz, Pz, P7, and P8.

of information—to be input, resulting in a "loss" of 3.21 bits per sinogram input.

This bottleneck on performance may be alleviated by revising the stimulus presentation in three ways[4]: First, by reducing the number of components, the stimulus matrix size may be reduced, resulting in faster trials and, potentially, more accurate selections, although more selections must consequently be made before a sinogram can be input. If the components are defined such that only component combinations that encode sinograms are available to the user, a higher resultant bit rate may be maintained. The five-stroke methods (Jin et al., 2010;

[4]Performance improvement could also be sought by revising the classification algorithm, by additional training of the participants, and so on. The focus in this article is to present the FLAST speller for Chinese text input and to assess its performance based on the use of a standard classification algorithm, SWLDA.

Wu et al., 2009) take this approach. Although suitable for users who achieve high accuracy, these systems may be unusable in practice for users with low selection accuracy who may be required to frequently correct the six to seven selections necessary to input one sinogram. In the future, a comparison of online sinogram input accuracy and speed may be performed for the FLAST method and the five-stroke methods to determine under what circumstances each approach achieves the better performance. Second, by reducing the overhead duration—both the time allocated to locate target components and sinograms, and the interstimulus intervals—sinograms may be input more rapidly. We anticipate that the overhead may be significantly reduced below the present 47 s with minimal reduction in performance. Third, the mean number of selections required to input sinograms can be reduced by incorporating character prediction into the speller (Ryan et al., 2011). For the sinogram, this can be achieved on three levels: at the word level, by assessing which words commonly co-occur to form phrases; at the sinogram level, by predicting the second sinogram of a two-sinogram word after input of the first sinogram; and at the component level, by predicting which sinograms are most likely to match a single selected component. Implementation of both overhead reduction and character prediction is in progress, and is expected to enhance both the online accuracy and input speed of the FLAST Speller.

The majority of visual spellers use the P300 response as the primary source of spatiotemporal features by which to conduct classification. However, as introduced in the results presented earlier, no clear P300 component was elicited from participants in response to target intensification in the FLAST procedure. Instead, an apparent N200 component, peaking about 200 ms after target onset at both left- and right-hemisphere parietal sites, was observed. The N-200 Speller (Hong et al., 2009) has made use of the N200 visual motion response (Kuba & Kubova, 1992), elicited by the use of motion stimuli, for accurate alphabetic text input. In the present work, however, the source of the apparent N200 cannot be attributed to a visual motion response because no motion stimuli were used. Additional experiments are necessary to investigate the causes and implications of the weak P300 and strong apparent N200 response in FLAST.

Validation of the performance of the system by users with neuromuscular disability also remains a necessary step to be carried out. Nevertheless, the results reported here demonstrate that online input of Chinese text from a large lexicon using the FLAST Speller is feasible.

REFERENCES

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

BioSemi B.V. (2007, July 3). *Active two user manual* (Version 3.2). Amsterdam, the Netherlands: BioSemi B. V.

Brouwer, A-M., & van Erp, J. B. F. (2010). A tactile P300 brain-computer interface. *Frontiers in Neuroscience*, *4*(19). doi:10.3389/fnins.2010.00019

Brouwer, A-M., van Erp, J. B. F., Aloise, F., & Cincotti, F. (2010). Tactile, visual, and bimodal P300s: Could bimodal P300s boost BCI performance? *SRX Neuroscience*. doi:10.3814/2010/967027

Brunner, P., Ritaccio, A. L., Emrich, J. F., Bischof, H., & Schalk, G. (2011). Rapid communication with a 'P300' matrix speller using electrocorticographic signals (ECoG). *Frontiers in Neuroscience*, *1*(5). doi:10.3389/fnins.2011.00005

Daniels, P. T., & Bright, W. (1996). *The world's writing systems*. New York, NY: Oxford University Press.

Draper, N. R., & Smith, H. (1981). *Applied regression analysis* (2nd ed.). New York, NY: Wiley.

Fabiani, M., Gratton, G., Karis, D., & Donchin, E. (1987). Definition, identification, and reliability of measurement of the P300 component of the event-related brain potential. In P. K. Ackles, J. R. Jennings, & M. G. H. Coles (Eds.), *Advances in psychophysiology, Vol. 2* (pp. 1–78). Amsterdam, the Netherlands: JAI Press.

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

Fong, M. C-M., & Minett, J. W. (2012). Chinese input methods: Overview and Comparisons. *Journal of Chinese Linguistics*, 40, 102–138.

Frye, G. E., Hauser, C. K., Townsend, G., & Sellers, E. W. (2011). Suppressing flashes of items surrounding targets during calibration of a P300-based brain–computer interface improves performance. *Journal of Neural Engineering*, *8*. doi:10.1088/1741-2560/8/2/025024

Furdea, A., Halder, S., Krusienski, D. J., Bross, D., Nijboer, F., Birbaumer, N., & Kübler, A. (2009). An auditory oddball (P300) spelling system for brain–computer interfaces. *Psychophysiology*, *46*, 617–625.

Gu, Y. (2009). Chinese. In K. Brown & S Ogilvie (Eds.), *Concise encyclopedia of languages of the world* (pp. 213–220). Oxford, UK: Elsevier.

Guger, C., Daban, S., Sellers, E., Holzner, C., Krausz, G., Carabalona, R., . . . Edlinger, G. (2009). How many people are able to control a P300-based brain–computer interface (BCI)? *Neuroscience Letters*, *462*, 94–98.

Höhne, J., Schreuder, M., Blankertz, B., & Tangermann, M. (2010). Two-dimensional auditory p300 speller with predictive text system. *Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 4185–4188.

Hong, B., Guo, F., Liu, T., Gao, X., & Gao, S. (2009). N200-speller using motion-onset visual response. *Clinical Neurophysiology*, *120*, 1658–1666.

Jin, J., Allison, B. Z., Brunner, C., Wang, B., Wang, X., Zhang, J., . . . Pfurtscheller, G. (2010). P300 Chinese input system based on Bayesian LDA. *Biomedizinische Technik: Biomedical Engineering*, *55*(1), 5–18.

Klobassa, D. S., Vaughan, T. M., Brunner, P., Schwartz, N. E., Wolpaw, J. R., Neuper, C., & Sellers, E. W. (2009). Toward a high-throughput auditory P300-based brain-computer interface. *Clinical Neurophysiology*, *120*, 1252–1261.

Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., & Wolpaw, J. R. (2006). A comparison of classification techniques for the P300 speller. *Journal of Neural Engineering*, *3*, 299–305.

Krusienski, D. J., Sellers, E. W., McFarland, D. J., Vaughan, T. M., & Wolpaw, J. R. (2008). Toward enhanced P300 speller performance. *Journal of Neuroscience Methods*, *167*(1), 15–21.

Kuba, M., & Kubova, Z. (1992). Visual evoked potentials specific for motion onset. *Documenta Ophthalmologica*, *80*, 83–89.

Lenhardt, A., Kaper, M., & Ritter, H. J. (2008). An adaptive P300-based online brain-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *16*, 121–130.

Leuthardt, E. C., Schalk, G., Wolpaw, J. R., Ojemann, J. G., & Moran, D. W. (2004). A brain–computer interface using electrocorticographic signals in humans. *Journal of Neural Engineering*, *1*(2), 63–71.

Lewis, M. P. (Ed.). (2009). *Ethnologue: Languages of the World* (16th ed.). Dallas, TX: SIL International. Available from http://www.ethnologue.com/

Li, Y., Nam, C. S., Shadden, B. B., & Johnson, S. L. (2010). A P300-based brain–computer interface: Effects of interface type and screen size. *International Journal of Human–Computer Interaction*, *27*, 52–68.

Mak, J. N., Arbel, Y., Minett, J. W., McCane, L. M., Yuksel, B., Ryan, D., . . . Erdogmus, D. (2011). Optimizing the P300-based brain–computer interface: current status, limitations and future directions. *Journal of Neural Engineering*, *8*. doi:10.1088/1741-2560/8/2/025003

Minett, J. W., Peng, G., Zhou, L., Zheng, H-Y., & Wang, W. S-Y. (2010, June). An assistive communication brain–computer interface for Chinese text input (ID 41183). *Proceedings of the 4th International Conference on Bioinformatics and Biomedical Engineering*.

Nam, C. S., Li, Y., & Johnson, S. (2010). Evaluation of P300-based brain–computer interface in real-world contexts. *International Journal of Human–Computer Interaction*, *26*, 621–637.

Ryan, D. B., Frye, G. E., Townsend, G., Berry, D. R., Mesa-G., S., Gates, N. A., & Sellers, E. W. (2011). Predictive spelling with a P300-based brain-computer interface: Increasing the rate of communication. *International Journal of Human-Computer Interaction*, *27*, 69–84.

Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., & Wolpaw, J. R. (2004). BCI2000: A general-purpose brain–computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering*, *51*, 1034–1043.

Sellers, E. W., Krusienski, D. J., McFarland, D. J., Vaughan, T. M., & Wolpaw, J. R. (2006). A P300 event-related potential brain-computer interface (BCI): The effects of matrix size and inter stimulus interval on performance. *Biological Psychology*, *73*, 242–252.

Serby, H., Yom-Tov, E., & Inbar, G. F. (2005). An improved P300-based brain–computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *13*, 89–98.

Sharbrough, F. C. G., Lesser, R. P., Lüders, H., Nuwer, M., & Picton, W. (1991). AEEGS guidelines for standard electrode position nomenclature. *Clinical Neurophysiology*, *8*, 202–204.

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Information delivery and the sensory evoked potential. *Science*, *155*, 1436–1439.

Townsend, G., LaPallo, B. K., Boulay, C. B., Krusienski, D. J., Frye, G. E., Hauser, C. K., . . . Sellers, E. W. (2010). A novel P300-based brain–computer interface stimulus presentation paradigm: Moving beyond rows and columns. *Clinical Neurophysiology*, *121*, 1109–1120.

Van Driem, G. (2001). *Languages of the Himalayas: An ethnolinguistic handbook of the greater Himalayan region*. Leiden, the Netherlands: Brill.

Wang, W. S-Y. (1973). The Chinese language. *Scientific American*, *228*, 50–60.

Wang, W. S-Y. (1981). Language structure and optimal orthography. In O. J. L. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experimental psychology* (pp. 223–236). Hillsdale, NJ: Erlbaum.

Wang, W. S-Y., & Tsai, Y. (2011). The Alphabet and the sinogram: Setting the stage for a look across orthographies. In P. McCardle, B. Miller, J. R. Lee, & O. Tzeng (Eds.), *Dyslexia across languages*. (pp. 1–16) Baltimore, MD: Brookes.

Wolpaw, J. R., Birbaumer, N., Heetderks, W. J., McFarland, D. J., Peckham, P. H., Schalk, G., . . . Vaughan, T. M. (2000). Brain-computer interface technology: A review of the first international meeting. *IEEE Transactions on Rehabilitation Engineering*, *8*, 164–173.

Wu, B., Su, Y., Zhang, J-H., Li, X., Zhang, J-C., Cheng, W-D., & Zheng, X-X. (2009). A virtual Chinese keyboard BCI system based on P300 potentials [Chinese]. *Acta Electronica Sinica*, *37*, 1733–1745.

Xinhua Zidian (10th ed.). [Chinese]. (2004). Beijing, China: Shang wu yin shu guan.

Yin, B., & Felley, M. (1990). *Chinese Romanization. Pronunciation and orthography*. Beijing, China: Sinolingua.

ABOUT THE AUTHORS

James W. Minett received his PhD in Electronic Engineering from City University of Hong Kong. He is currently Research Associate in the Language Engineering Laboratory at the Chinese University of Hong Kong, where he works on brain-computer interfaces, psycholinguistics, and evolutionary linguistics.

Hong-Ying Zheng received her PhD in Electronic Engineering at the Language Engineering Laboratory, Chinese University of Hong Kong. She continued to work at the laboratory as a Visiting Scholar, conducting research on psycholinguistics and brain-computer interfacing.

Manson C-M. Fong graduated from the Hong Kong University of Science and Technology with an MPhil in Physics in 2009. He is now pursuing a PhD in Electronic Engineering at the Language Engineering Laboratory of the Chinese University of Hong Kong. His research interests include brain-computer interfacing and signal processing.

Lin Zhou graduated from Hunan University with a Bachelor degree in Software Engineering in 2007. She is now a Research Assistant in Electronic Engineering at the Language Engineering Laboratory of the Chinese University of Hong Kong, doing research on psycholinguistics and brain-computer interfaces.

Gang Peng received his PhD in Language Engineering from City University of Hong Kong. He is Research Associate Professor of the Department of Linguistics and Modern Languages at the Chinese University of Hong Kong. He is also Professor of Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences.

William S-Y. Wang received his PhD from the University of Michigan. He was Professor of Linguistics at the University of California at Berkeley for thirty years. Currently he directs the Language Engineering Laboratory at the Chinese University of Hong Kong. He is also Academician of Academia Sinica in Taiwan.