2016

Using Motor Imagery to Control Brain-Computer Interfaces for Communication

Jonathan S. Brumberg

Jeremy D. Burnison

Kevin M. Pitt

Using Motor Imagery to Control Brain-Computer Interfaces for Communication

Jonathan S. Brumberg,[1,2] Jeremy D. Burnison,[2] and Kevin M. Pitt[1]

1 Department of Speech-Language-Hearing: Sciences & Disorders, University of Kansas, Lawrence, KS, USA

2 Neuroscience Graduate Program, University of Kansas, Lawrence, KS, USA
    jburnison@ku.edu

Corresponding author — J.S. Brumberg, *email* brumberg@ku.edu

Abstract

Brain-computer interfaces (BCI) as assistive devices are designed to provide access to communication, navigation, locomotion and environmental interaction to individuals with severe motor impairment. In the present paper, we discuss two approaches to communication using a non-invasive BCI via recording of neurological activity related to motor imagery. The first approach uses modulations of the sensorimotor rhythm related to limb movement imagery to continuously modify the output of an artificial speech synthesizer. The second approach detects event-related changes to neurological activity during single trial motor imagery attempts to control a commercial augmentative and alternative communication device. These two approaches represent two extremes for BCI-based communication ranging from simple, "button-click" operation of a speech generating communication device to continuous modulation of an acoustic output speech synthesizer. The goal of developing along a continuum is to facilitate adoption and use of communication BCIs to a heterogeneous target user population.

Keywords: BCI, Brain-computer interface, Motor imagery, AAC

1 Introduction

Since developing the first computers, technologists have been trying to reduce the separation between users and their devices. Moving from punch cards to keyboards was a dramatic advance, and today we now have reliable voice control, touch screens and adaptive keyboards with language prediction. In addition to improving the manner in which people interact with computers, these new methods are quite useful for providing access to communication systems for individuals with mild to moderate speech and movement disorders. Still, these methods require manual or spoken input, which may not be available to individuals with severe motor impairment and paralysis, specifically, those with locked-in syndrome (LIS) [1] due to stroke and neurodegenerative disorders (e.g., amyotrophic lateral sclerosis, ALS). Locked-in syndrome is characterized by near total paralysis, including the limbs and face, with some remaining eye movements and significant amounts of sensation and cognition [2]. Therefore, individuals with LIS require alternative means for accessing language and communication that does not require overt motor control. For those with reliable eye movement control, eye-gaze tracking hardware can be used to control high-tech augmentative and alternative communication devices (AAC), in addition to other user-customized input methods [3,4]. For individuals without reliable oculomotor control, brain-computer interfaces (BCI) may be the only method available to provide access to computers and communication.

The idea behind all BCIs is a straightforward extension of existing assistive technology, albeit to an extreme degree. Consider the case of spelling on a keyboard using an eye-gaze tracker. In this scenario, the user must first determine a desired message, identify all the required elements on the communication display, attend to each element, and make an appropriate oculomotor action to move the eyes and drive the eye-gaze tracking pointer. For an individual with LIS, the final oculomotor stage is impaired to such a degree that eye-gaze tracking is either impossible or unreliable. Individuals who use a button or mouse click to make communication device item selections using linear scanning follow a similar strategy; desired items are identified and attended to, then cortical motor commands are issued to activate

limb muscles necessary to activate the selection device. Here too, the final stage of motor command transmission to the periphery is impaired or absent in individuals with LIS. In both examples, the goal of a BCI is to intercept the last reliable neurological control signal available prior to attempted activation of the disordered periphery. For the visual attention example, it is possible to elicit and record the P300 event-related potential (ERP) [5–9], and in the button-click example neurological markers of motor planning and execution for use in communication interfaces [10–14]. A major research area now is focused on translating BCIs from research settings into practical user settings [15,16]. One barrier to this translation process arises since most BCIs use custom communication interface software that may not necessarily be compatible with existing high-tech AAC devices that are already supported by commercial and clinical professionals. In this paper, we describe two motor imagery based BCIs for communication in which one system directly interfaces with existing AAC devices and the other provides direct speech output without the need for a separate communication device.

2 BCI as an AAC Device

One barrier to clinical translation of BCI devices is the reliance on custom communication interfaces that often are not compatible with existing AAC devices. We therefore designed a BCI that does not rely on its own visual interface, rather, it uses existing AAC devices and software to help elicit neurological activity used for BCI control. Electroencephalography (EEG) signals are obtained as participants make covert (or imagined) motor movements of the limbs as they interact with a Tobii C15 communication device (Tobii-DynaVox). The presentation of communication items in a linear scanning protocol is expected to generate neurological signals related to both movement preparation and movement execution. Specifically, the BCI detects both the contingent negative variation (CNV), a potential related to the preparation of an upcoming, cued movement [17–19], and the event-related (de)synchronization (ERD/S), a change in the sensorimotor rhythm spectral power related to the preparation and execution of overt and covert movement [10–12].

The CNV is an event-related negativity normally elicited in a cue-response paradigm in which a warning stimulus alerts the participant to an upcoming event and a following imperative stimulus instructs the participant to produce a known motor command [17,18]. The linear scanning interface used for this experiment automatically advances sequentially, therefore, all non-target rows and columns can be considered as "warning" stimuli[1] and the target row or column is the "imperative" stimulus. The ERD/S is elicited in response to the preparation and execution of overt and covert motor commands, and are not dependent on a warning stimulus. Typically, an ERD is observed in the μ and β bands (8–14 Hz, 15–25 Hz, respectively) as an attenuation in spectral power around the onset of the intended action. Our BCI uses both movement-related signals to classify neurological activity into intended or unintended actions that are used to send a simulated button-press to the AAC device for item selection, though we only discuss the CNV results below.

2.1 Methods

Six individuals without neurological impairment and one individual with advanced ALS participated in the AAC-based BCI experiment. EEG was recorded continuously from each participant from 62 active electrodes (g.HIAmp, g.tec) at a sampling rate of 512 Hz with notch filters at 58–62 Hz. Participants were seated in a sound-treated booth in front of a simulated AAC device that displayed a preprogrammed communication interface page. The device was configured to automatically highlight each communication item sequentially with a red box (2.5 s interval) and the name of the item was played over the AAC device speakers. Participants were provided with a randomly selected target item and instructed to imagine a movement of their dominant hand every time it was highlighted. EEG signals were bandpass filtered from 0.5–8 Hz to obtain the CNV.

Offline classification of the CNV that preceded overt and covert movement for selecting communication items was accomplished using linear discriminant analysis (LDA) of the average EEG amplitude from −0.23 s to −0.03 s relative to the onset of item highlighting.[2]

---

1 Though likely strongest for the immediately preceding row or column.
2 Specifically the auditory signal onset.

Bipolar surface electromyography was also collected from the limbs to ensure participants adhered to the motor imagery instructions. Data was collected from 80 item highlighting trials per condition (overt and covert) and performance of the LDA classifier was evaluated using a 2-fold cross validation. In this procedure, the first 40 trials per condition were used to train the classifier and the second 40 trials used for validation, then the training and validation sets were switched to obtain a full estimate of the performance.

2.2 Results

Analysis of the EEG data indicated that the CNV was present (statistically significantly less than zero, 1-tailed t-test, fdr corrected $p < 0.05$) and spatially located primarily over bilateral parietal electrodes for all participants (see Figs. 1 and 2). In the overt condition, the CNV was characterized by a slow negativity followed by a peak negativity
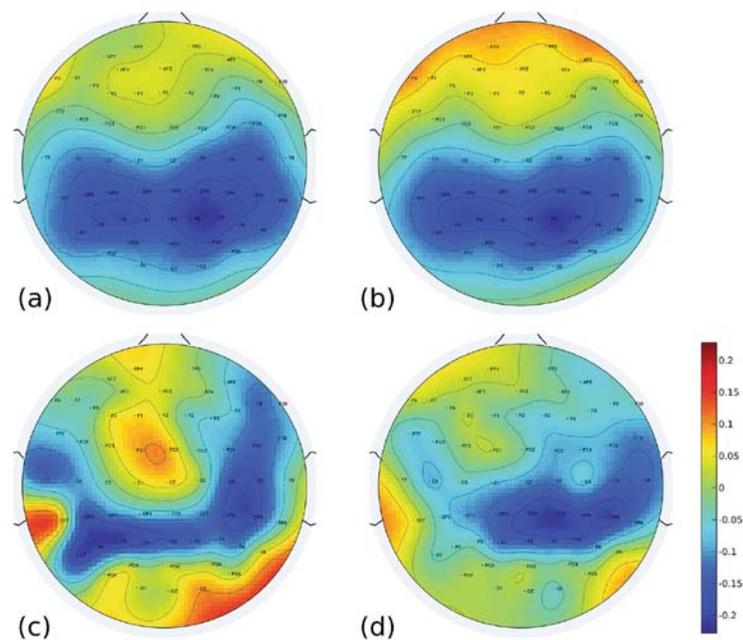

Fig. 1. The topography of average normalized EEG amplitudes in the 200 ms prior to auditory stimulus onset. The average for all participants without impairment in the overt condition is shown in (a) and covert in (b). For both conditions, there is a strong bilateral parietal distribution of negative amplitudes. Similarly, the patterns of negativity for the participant with ALS show bilateral posterior negativity in both the overt (c) and covert (d) conditions, with a slightly right-lateralized response in the covert condition.
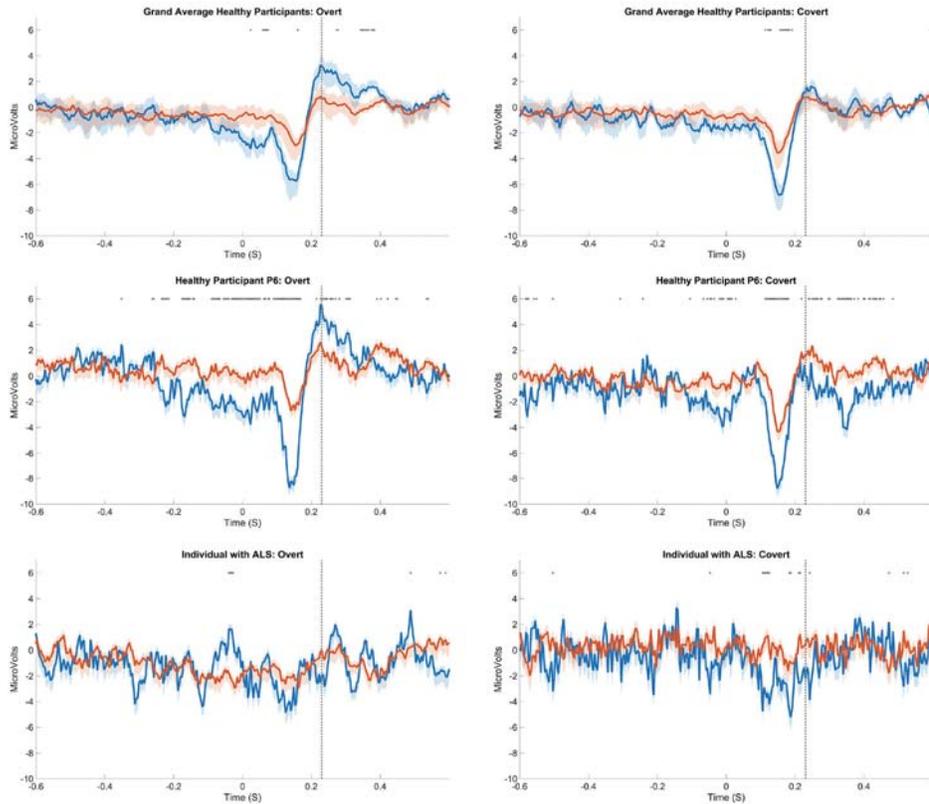
Fig. 2. A graphical summary of average CNV amplitudes for overt (left column) and covert (right column) movements in the AAC selection paradigm. The top row represents average CNVs over all healthy participants (N = 6), the middle row is the average CNV for one healthy participant and the bottom row is the average CNV for one participant with ALS. *: indicates statistically significant differences between target trials (blue) and non-target trials (red), and the shaded ranges indicate the 95% confidence intervals.

immediately prior to auditory stimulus in the overt condition for participants without neurological impairments. In the covert condition, only the peak negativity prior to auditory stimulus onset was observed for participants without neurological impairments. For the participant with ALS, the overt condition did not elicit demonstrable negativity prior to the auditory stimulus onset, however a slight negativity was observed immediately prior to the auditory stimulus onset in the covert condition.

The cross-validation accuracy of the LDA classifier was 64% in the overt condition for participants without neurological impairment and 60% in the covert condition. The cross-validation accuracy of the LDA

Table 1. A summary of offline decoding accuracy for all participants, with (P1–6) and without (ALS1) neuromotor impairment in the overt and covert production tasks.

	P1	P2	P3	P4	P5	P6	Average	ALS1
Overt	75.4%	61.1%	56.8%	65.8%	66.4%	59.3%	64.1%	NA
Covert	60.0%	55.7%	58.9%	60.4%	63.5%	61.3%	60.0%	62.6%

classifier for the individual with ALS was 63% in the covert condition, and was not attempted in the overt condition due to the lack of statistically significant CNV response. Table 1 includes a summary of individual and average classification accuracy. For all participants, the classification was based on a 0.2 s window prior to the auditory stimulus onset, however, the number of electrodes differed based on the location of the peak CNV negativity on the scalp. In general, the electrodes were chosen from the CP, P and PO locations, and more electrodes passed our inclusion criteria in the overt condition (average: 3.5, range: 2–7) than the covert condition (average: 1.8, range 1–3). A similar CNV topography was observed for the individual with ALS, and two electrodes from the CP region were used for decoding.

2.3 Discussion

The AAC-BCI device described in this experiment is designed to simplify BCI control as much as possible for individuals who already use, or may use BCI in the future. Our approach for achieving this aim is to rely on existing AAC technology and techniques that may be most familiar to both users of AAC (who are also potential users of BCI), and AAC professionals. One of the most basic ways for controlling AAC devices is through the use of a physical button and a visual interface with automatically advancing scanning of communication icons. Such a device can be used by individuals with disordered, but present motor control. In this experiment, we extend the existing AAC paradigm to BCI by replacing the AAC item selection mechanism with a "brain switch" controlled using a neurological potential related to motor planning and motor execution in a mental button pressing task.

The contingent negative variation is a very well known neurological potential that precedes movement in a cued paradigm. Classically, this potential is strongest when individuals know they will be required to

make a movement in the near future, but both the action and its timing are uncertain [17,18]. In this experiment, however, only a portion of these classical factors used to elicit the CNV are met: (1) movements are made in a cued paradigm and (2) individuals know they will make movements in the near future. The third factor, uncertainty of both action and timing, is not met because the AAC device automatically scans through all available communication items at a predictable rate. Our preliminary results show that the third requirement is not necessary for eliciting the CNV; it is present in our paradigm for both individuals with and without neurological impairments. Further, our offline classification results show it is possible to accurately predict the occurrence of CNVs in a cued motor control paradigm. On average, the LDA classifier performed better in the overt condition than the covert for individuals without neuromotor impairments. Additionally, the decoder performed marginally better for in the covert condition for the individual with ALS compared to the participants without impairment, though no statistical analysis of these differences was performed. These are promising results that warrant further study of an online decoder for controlling an AAC device in real-time.

3 BCI-Controlled Speech Synthesizer

Our second BCI implementation provides continuous control over a formant frequency based speech synthesizer through detection of modulations to the sensorimotor rhythm [20]. The primary advantage of this system is the distinct lack of a communication interface, rather the user is directly in control of acoustic speech output. This BCI is based on prior work decoding continuous, two-dimensional control signals from the EEG sensorimotor rhythm [14]. In previous studies, participants learned to control a two-dimensional cursor by performing limb motor imagery that modulated the sensorimotor rhythm. In speech, spectral energy of vowel sounds and transitions into, and out of, consonants can be represented by low-dimensional acoustic features known as formant frequencies, or formants. These features are directly related to the dynamically changing configuration and resonant properties of the vocal tract. It is possible to perfectly represent all of the monophthong vowels in American English using just the first two formant frequencies (F1 and F2). In addition, there are a

number of real-time formant frequency speech synthesizers capable of instantaneous auditory feedback. Therefore, our BCI decodes continuous modulation of the sensorimotor rhythm into a two-dimensional formant frequency feature vector that is synthesized and provided back to the user in real-time.

3.1 Methods

Three individuals without any neuromotor impairments were recruited to participate in the BCI-controlled speech synthesizer study. EEG was recorded continuously from 62 active electrodes (g.HIAmp, g.tec) at a sampling rate of 512 Hz with notch filters from 58–62 Hz. The EEG signals were then re-referenced to the common average reference and bandpass filtered from 7–15 Hz to obtain the μ band (i.e., sensorimotor rhythm). Finally, the band power was calculated based on the analytic amplitude from the Hilbert transform. During the experiment, vowel sounds were presented visually as a two-dimensional cursor position on a display with the positions of the three test vowels (/a/, /u/ and /i/). Auditory stimuli (and BCI feedback) were synthesized in real-time using the Snack Sound Toolkit (KTH Royal Institute of Technology) and played through pneumatic insert earphones (ER1, Etymotic, Inc.). Participants were instructed to imagine moving their right hand when presented with an /a/ stimulus, their left hand for the /u/ stimulus and their feet for the /i/ stimulus.

During training, participants were asked to imagine the appropriate movement throughout the entire 3 s stimulus period. A total of 135 trials were presented (45 trials per vowel) with vowels in random order. The sensorimotor bandpower and target vowel formant frequency velocities (bark/s) were used to estimate the state and likelihood models of a Kalman filter decoder. Formant velocities are taken as the change in formant frequency over time. Offline training and performance was evaluated using a two-fold cross-validation of the correlation coefficient of each formant velocity trajectory to the target vector, and the combined 2D formant velocity trajectories.

3.2 Results

The procedure for training the Kalman filter decoder revealed asymmetric linear model weights over the left and right sensorimotor areas
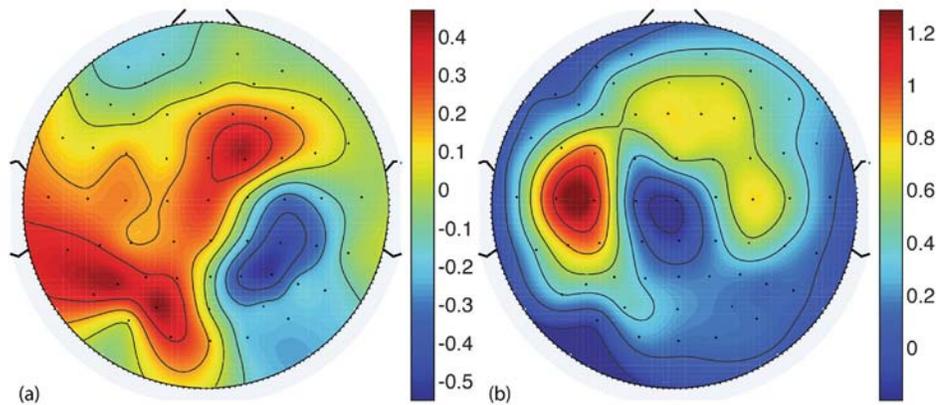
Fig. 3. The Kalman filter decoder linear model weights for each formant velocity dimension reveal contribution of the sensorimotor electrodes to the motor imagery task. (a) The model weights for F1 velocities show asymmetric model weights with sign contralateral to the intended left or right limb movement imagery (/u/ and /a/ vowels). (b) The model weights for F2 velocities show symmetric model weights for the bilateral foot movement imagery task (/i/ sound)

(C, CP and FC electrodes) contralateral to the intended movement imagery. In contrast, the model weights for relating sensorimotor rhythm modulations to the second formant frequency are symmetric and bilateral. The model weights are shown graphically in Fig. 3 and confirm the involvement of sensorimotor areas in the motor imagery task.

A two-fold cross-validation procedure was used to evaluate the offline performance of the trained Kalman filter decoder. The model predicted formant velocities are shown graphically in Fig. 4. In this figure, the average predicted formant frequency velocities are shown on the left for /i/ (top), /a/ (middle) and /u/ (bottom) with 95% confidence intervals (shaded regions) and velocity targets in black. These trajectories are shown on the 2D formant velocity plane in Fig. 4(d) for the /i/ (blue), /a/ (red) and /u/ (yellow) vowels. From this view, it is possible to observe a more faithful prediction of /i/ vowel velocities compared to the /a/ and /u/ vowels; however, there is greater overall congruence when the velocities are integrated in time to obtain final predicted formant frequencies (Fig. 4(d)). These results, quantitatively summarized in Table 2, indicate a moderate ($r = 0.51$) correlation between the predicted and target 2D formant velocity trajectories as well as the correlations of the individual formants to their targets (F1: $r = 0.35$, F2: $r = 0.62$).
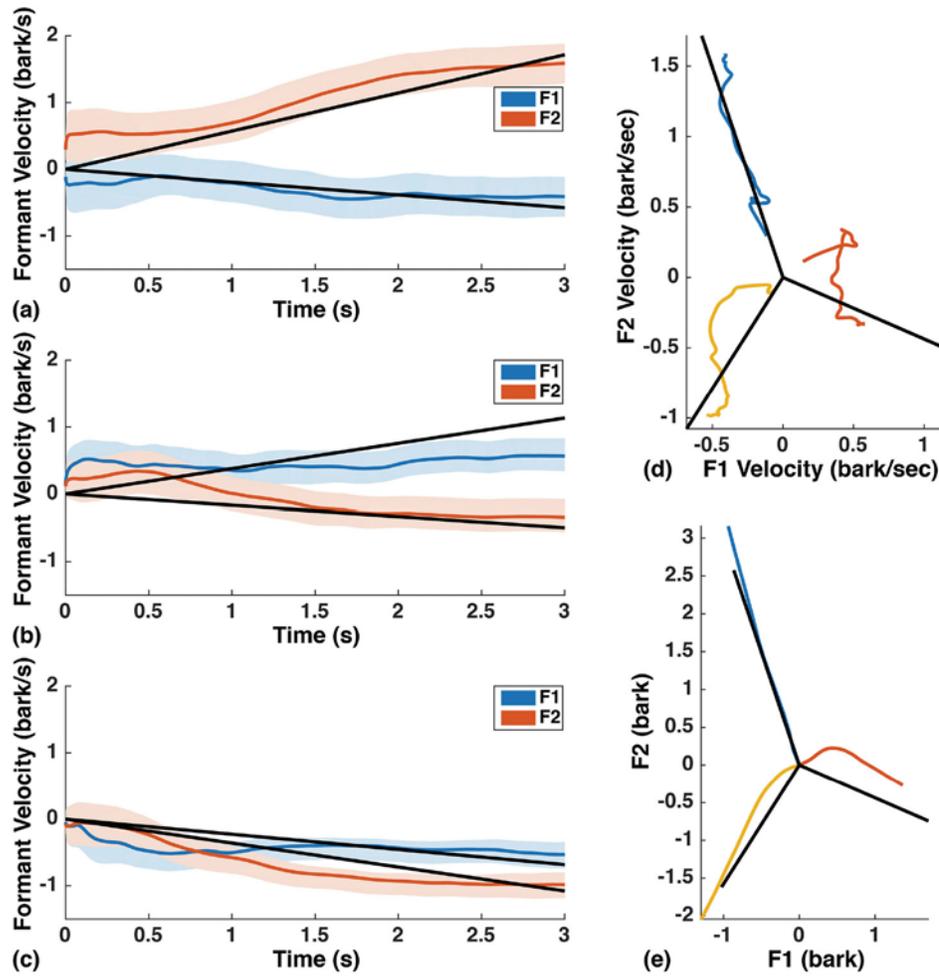
Fig. 4. (Left) A graphical summary of the first and second formant frequency velocities (F1 blue, F2 red) for the /i/, /a/ and /u/ vowels, respectively in subfigures (a), (b) and (c). Shaded regions represent the 95% confidence interval and the black lines represent target formant velocity trajectories. (right) A graphical summary of the predicted formant frequency velocities (d) and integrated formants (e) on the 2D formant plane. Here, the blue line is the average predicted model response for /i/, the red line for /a/ and the yellow line for /u/. The black lines are the target formant responses. Note, the integrated formants are centered based on the average of the three tested vowels.

Table 2. Pearson's correlation coefficient (r) of the Kalman filter decoder for predicting formant frequency velocities in the synthesizer BCI.

	F1	*F2*	*Combined*
Training set	0.50	0.68	0.66
Validation set	0.35	0.62	0.51

3.3 Discussion

The BCI-controlled speech synthesizer is fundamentally a limb motor imagery based BCI for decoding a continuous 2D output vector (similar to [14]), but in both auditory and visual feedback domains related to speech. Therefore, we can examine the results of our single pilot participant to determine whether our protocol, paradigm and BCI algorithms are functioning appropriately. Specifically, we can examine the Kalman filter linear model weights, which represent the relationship between modulations in the 7–15 Hz sensorimotor rhythm and target formant velocity for each vowel production – motor imagery trial. In this way, the model weights themselves are informative for determining the spatial topography of EEG activity participants use to complete the vowel synthesizer task.

The results of our offline decoding analysis reveal a scalp topography of model weights that reflect differential activation for the control of the first formant frequency, and coordinated activation for controlling the second formant frequency. This asymmetric response is expected because we asked participants to use contralateral limb motor imagery for the vowels /a/ and /u/ which differ almost entirely in the first formant. Similarly, a coordinated, bilateral topography for controlling the second format agrees with the instructed task for associating bilateral foot movement imagery with productions of the vowel /i/, which primarily differs from /a/ and /u/ in the second formant. Finally, the moderate correlation of predicted formant frequency velocities to targets is promising for continued investigation in an online control paradigm. The addition of closed-loop feedback of audio and visual information, should help to generate error control signals used to improve the continuous BCI control for the production of the vowels /a/, /i/ and /u/.

4 Conclusions

In the present paper, we examine two BCIs for communication using two separate control techniques. In the first example, we extend existing AAC input signal designs for accessing communication software programs using a "brain switch." This approach decodes and uses the

neurological potentials related to a mental button pressing task to select items on a communication interface. Our preliminary evidence provides some encouraging results for continuing to explore this BCI application in real-time with additional participants with and without neurological impairments. In the second example, we validated our approach for decoding continuously varying two-dimensional formant frequencies from sensorimotor rhythm modulations. Our modeling results are compatible with past studies of SMR-based BCIs for 2D cursor control, and the offline prediction of formant frequencies is reliable enough for additional study of online control of a speech synthesizer via BCI.

References

1. Plum, F., Posner, J.B.: The diagnosis of stupor and coma. Contemp. Neurol. Ser. **10**, 1–286 (1972)

2. Bauer, G., Gerstenbrand, F., Rumpl, E.: Varieties of the locked-in syndrome. J. Neurol. **221**(2), 77–91 (1979)

3. Beukelman, D., Fager, S., Nordness, A.: Communication support for people with ALS. Neurol. Res. Int. **2011**, 714693 (2011)

4. Fager, S., Beukelmanm, D.R., Fried-Oken, M., Jakobs, T., Baker, J.: Access interface strategies. Assistive Technol. Official J. RESNA **24**(1), 25–33 (2012)

5. Donchin, E., Spencer, K.M., Wijesinghe, R.: The mental prosthesis: assessing the speed of a P300-based brain-computer interface. IEEE Trans. Rehabil. Eng. **8**(2), 174–179 (2000)

6. Acqualagna, L., Blankertz, B.: Gaze-independent BCI-spelling using rapid serial visual presentation (RSVP). Clin. Neurophysiol. **124**(5), 901–908 (2013)

7. Krusienski, D.J., Sellers, E.W., McFarland, D.J., Vaughan, T.M., Wolpaw, J.R.: Toward enhanced P300 speller performance. J. Neurosci. Meth. **167**(1), 15–21 (2008)

8. Nijboer, F., Sellers, E.W., Mellinger, J., Jordan, M.A., Matuz, T., Furdea, A., Halder, S., Mochty, U., Krusienski, D.J., Vaughan, T.M., Wolpaw, J.R., Birbaumer, N., Kübler, A.: A P300-based brain-computer interface for people with amyotrophic lateral sclerosis. Clin. Neurophysiol. **119**(8), 1909–1916 (2008)

9. Oken, B.S., Orhan, U., Roark, B., Erdogmus, D., Fowler, A., Mooney, A., Peters, B., Miller, M., Fried-Oken, M.B.: Brain-computer interface with language model electroencephalography fusion for locked-in syndrome. Neurorehabilitation Neural Repair **28**(4), 387–394 (2014)

10. Pfurtscheller, G., Neuper, C.: Motor imagery and direct brain-computer communication. Proc. IEEE **89**(7), 1123–1134 (2001)

11. Blankertz, B., Dornhege, G., Krauledat, M., Müller, K.R., Kunzmann, V., Losch, F., Curio, G.: The Berlin brain-computer interface: EEG-based communication without subject training. IEEE Trans. Neural Syst. Rehabil. Eng. **14**(2), 147–152 (2006)

12. Neuper, C., Müller-Putz, G.R., Scherer, R., Pfurtscheller, G.: Motor imagery and EEG-based control of spelling devices and neuroprostheses. Prog. Brain Res. **159**, 393–409 (2006)

13. Ramoser, H., Müller-Gerking, J., Pfurtscheller, G.: Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Trans. Rehabil. Eng. **8**(4), 441–446 (2000)

14. Wolpaw, J.R., McFarland, D.J.: Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. Proc. Nat. Acad. Sci. U.S.A **101**(51), 17849 (2004)

15. Kleih, S.C., Kaufmann, T., Zickler, C., Halder, S., Leotta, F., Cincotti, F., Aloise, F., Riccio, A., Herbert, C., Mattia, D., Kübler, A.: Out of the frying pan into the fire-the P300-based BCI faces real-world challenges. In: Schouenborg, J., Garwicz, M., Danielsen, N. (eds.) Progress in Brain Research, vol. 194, 1st edn, pp. 27–46. Elsevier B.V., New York (2011). Chap. 2

16. Vaughan, T.M., McFarland, D.J., Schalk, G., Sarnacki, W.A., Krusienski, D.J., Sellers, E.W., Wolpaw, J.R.: The Wadsworth BCI research and development program: at home with BCI. IEEE Trans. Neural Syst. Rehabil. Eng. **14**(2), 229–233 (2006)

17. Walter, W.G., Cooper, R., Aldridge, V.J., McCallum, W.C., Winter, A.L.: Contingent negative variation : an electric sign of sensori-motor association and expectancy in the human brain. Nature **203**(4943), 380–384 (1964)

18. Rohrbaugh, J.W., Syndulko, K., Lindsley, D.B.: Brain wave components of the contingent negative variation in humans. Science (New York, N.Y.) **191**(4231), 1055–1057 (1976)

19. Funderud, I., Lindgren, M., Løvstad, M., Endestad, T., Voytek, B., Knight, R.T., Solbakk, A.K.: Differential Go/NoGo activity in both contingent negative variation and spectral power. PLoS ONE **7**(10), e48504 (2012)

20. Guenther, F.H., Brumberg, J.S.: Brain-machine interfaces for real-time speech synthesis. In: Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2011), Boston, MA (2011)