Available online at www.sciencedirect.com

ScienceDirect

Procedia Technology 18 (2014) 25 – 31

Procedia Technology

International workshop on Innovations in Information and Communication Science and Technology, IICST 2014, 3-5 September 2014, Warsaw, Poland

Spatial Auditory Two-step Input Japanese Syllabary Brain-computer Interface Speller

Moonjeong Chang[a], Koichi Mori[b], Shoji Makino[a], Tomasz M. Rutkowski[a,c,*]

[a] *Life Science Center of TARA, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8577, Japan*
[b] *Research Institute of National Rehabilitation Center for Persons with Disabilities, 4-1 Namiki, Tokorozawa, Saitama, 359-8555, Japan*
[c] *RIKEN Brain Science Institute, 2-1 Hirosawa, Wako-shi, Saitama, 351-0198, Japan*

Abstract

We present a concept and a pilot study of a two-step input speller application combined with a spatial auditory brain-computer interface (BCI) for locked-in syndrome (LIS) users. The application has been developed for 25 Japanese syllabary (*hiragana*) characters using a two-step input procedure, in order to create an easy-to-use BCI-speller interface. In the proposed procedure, the user first selects the representative letter of a subset, defining the second step. In the second step, the final choice is made. At each interfacing step, the user's intentional choices are classified based on the P300 event related potential (ERP) responses captured in the EEG, as in the classic oddball paradigm. The BCI experiment and EEG results of the pilot study confirm the effectiveness of the proposed spelling method.

Peer-review under responsibility of the Scientific Committee of IICST 2014

Keywords: brain-computer interface (BCI), spatial auditory BCI, spatial auditory perception

---

1. Introduction

A brain-computer interface (BCI) is a technology that uses brain activities to control a computer, or a machine, without any body muscle movements [1]. This technology could allow disabled people, such as those suffering from locked-in syndrome (LIS), to regain communication skills or to manage daily life support related functions. There are several types of BCI based on the brain signal modalities utilized. In our research, we focus on an auditory and non-invasive BCI modality, since it allows application requiring no training even for the naive user [2]. A diagram depicting the most popular BCI modalities is presented in Figure 1 (reproduced with permission from [3]). Non-invasive BCIs are the safest because the brainwave sensors are usually attached to the surface of the human head. These modalities usually demonstrate lower signal quality due to a very poor signal-to-noise ratio compared to invasive systems requiring potentially dangerous surgical interventions. Non-invasive BCIs are divided into stimulus-driven BCIs and imagery BCIs, as shown in Figure 1. Stimulus-driven BCIs provide interactive interfacing solutions in which

---

* Corresponding author. Tel.: +81-29-853-6261; web: http://bci-lab.info/
E-mail address: tomek@bci-lab.info

Peer-review under responsibility of the Scientific Committee of IICST 2014
doi:10.1016/j.protcy.2014.11.007

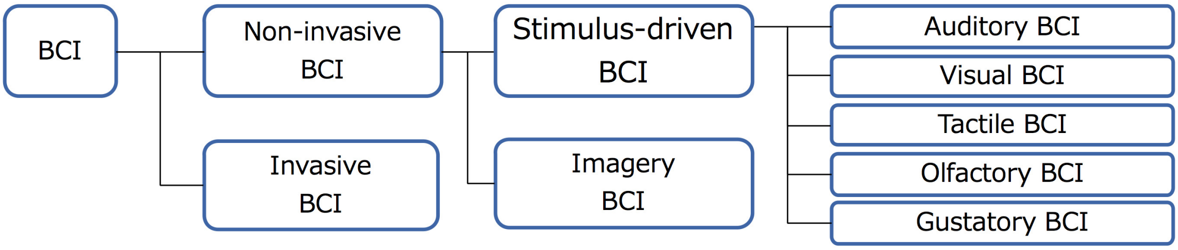

Fig. 1. A tree diagram depicting the most common types of BCI. In the stimulus-driven BCI branch, spatial auditory, visual, and tactile modalities have recently been successfully applied in various neurotechnology applications [2,4,5].

the commands to operate a computer or an application are associated with sensory stimuli delivered to the user. The user generates a signature brainwave captured by electroencephalography (EEG) by focusing attention on (the *target*) or by ignoring (the *non-target*) randomly presented sensory stimuli. Attention focused on randomly presented *targets* in a sequence of *non-targets* evokes a specific event related potential (ERP) pattern called the "aha" or *P*300 response (see results in Figure 4, in Section 4, where averaged brainwaves are depicted), since it is a positive brainwave EEG signal deflection starting 300 ms after a stimulus onset [1].

Among recently developed BCIs, the visual modality is the most successful because the evoked *P*300 responses usually have the largest amplitudes allowing for the easiest classification [2]. However, it is known that some LIS users (e.g. advanced stage amyotrophic lateral sclerosis patients) cannot use visual modalities because they gradually lose intentional muscle control, including eye movements, focusing or intentional blinking. Our research hypothesis is that the auditory modality BCI shall be a more suitable solution for establishing communication with LIS users.

Previously our research group already proposed an auditory BCI-speller application with spatial sounds generated through a vector based amplitude panning approach and delivered to the user via headphones [2,6]. In the research project reported in this paper, we aim to expand the spatial auditory BCI speller in order to test the possibility of inputting a larger set of Japanese syllabary characters. We propose to utilize the two-step input procedure in order to create an easy-to-use interface with limited numbers of choices (commands) in each step. We report results from a series of experiments in which the users were asked to spell short Japanese words using the two-step input scenario, as illustrated in Figure 2.

The remainder of the paper is organized as follows. In the next section, we describe the online BCI experiment set-up, EEG signal acquisition, pre-processing and classification procedures, as well as the two-step input Japanese syllabary speller paradigm. Then the results of the online BCI experiments are discussed, together with the ERPs, and especially the *P*300 response latencies. Finally, conclusions are drawn and future research directions outlined.

2. Materials and Methods

In the experiments reported in this paper, five healthy users took part (mean age of 27.2 years old, standard deviation (SD) of 9.5 years). All the experiments were performed at the Life Science Center of TARA, University of Tsukuba, Japan (experimental permission no. 2013*R*7). The online (real-time) EEG experiments were conducted in accordance with the *WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects*, and the procedures were designed according to the guidelines and with the approval of the ethics committee of the Faculty of Engineering, Information and Systems at University of Tsukuba, Japan.

In the proposed stimulus-driven spatial auditory BCI paradigm, 25 Japanese syllabary characters were used (half of the full set of *hiragana* characters). The sound recordings were obtained from a sound dataset constructed of female natural voice recordings and available for research purposes from [7]. The auditory stimuli were delivered from five distinct spatial locations using loudspeakers positioned in a row in front of the user's head, as depicted in Figure 3. In order to manage the spatial sound distribution and precise presentation to evoke the time aligned *P*300 responses, an in-house multimedia application was developed in a *MAX* environment [8]. The application communicated with the EEG acquisition, processing and classification environment *BCI2000* [9] using a User Datagram Protocol (UDP) network.

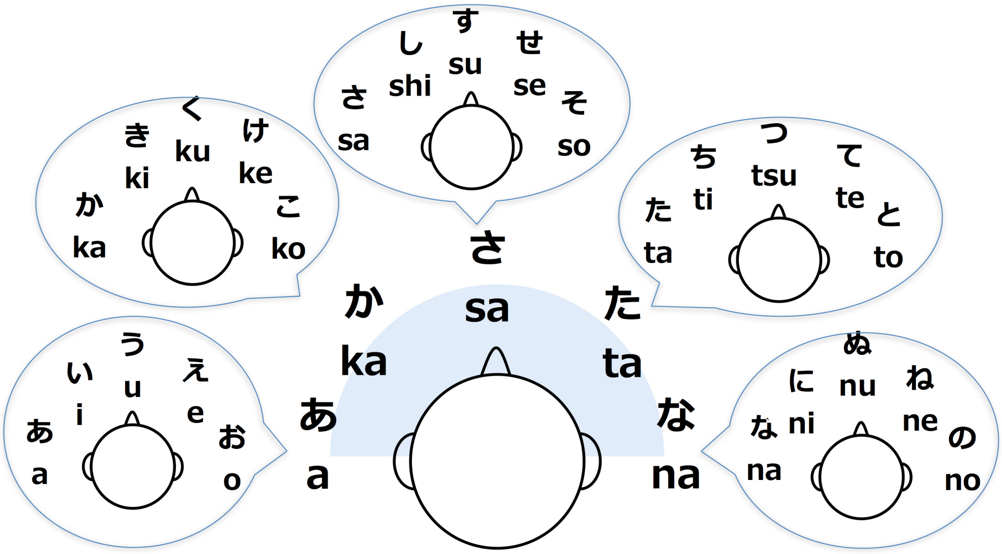

Fig. 2. The two-step input Japanese syllabary (*hiragana*) BCI-speller paradigm diagram. In the first step, the user chooses one of the subsets by attending to spatially distributed sounds. In the second step, the sounds in the spatial locations are replaced with those in the new subset.

Each user first conducted a short psychophysical test with a button press behavioural response to confirm understanding of the experiment set-up. Next, the EEG online BCI protocol experiments were conducted. In the EEG experiments, each user was instructed to spell a sequence of Japanese syllabary (*hiragana*) characters presented in random order, as in the classic *P*300 response-based oddball paradigm [1]. The following two types of experiment sessions were conducted with each user in the EEG experiments:

Training session: Using only the five Japanese vowels (*a, i, u, e,* and *o*), in a simple single-step spelling set-up, EEG data was collected in order to train the BCI classifier to discriminate the P300 responses (a binary classification of *target* versus *non-target* brainwave evoked responses).

Copy spelling test session: The research target two-step input Japanese syllabary speller paradigm test sessions were conducted online using the classifier trained in the preceding single-step run. As a task in the test experiments, five Japanese words (*sakana, kisetu, tesuto, nekutai,* and *shiitake*) were used for spelling in syllabary mode (*sa–ka–na, ki–se–tsu, te–su–to, ne–ku–ta-i,* and *shi–i–ta–ke*).

The single-step spelling mode was used in the training session in order to collect more quickly the necessary set of *P*300 responses for the successful training of the BCI classifier. The user only spelled a single string of the vowels, *a, i, u, e,* and *o* in 20-trial averaging mode in order to collect a representative set of responses for the classifier training. The final, 25 character-based two-step input procedure was based on the initial selection of a representative letter of a subset, defining the second step. In the second step, the final choice was made from the subset, as shown in the system diagram depicted in Figure 2.

During the online BCI experiments, the EEG signals were captured with eight active electrodes *g.LADYbird* by g.tec Medical Instruments GmbH, Austria, connected to an EEG amplifier *vAmp* by Brain Products GmbH, Germany. The electrodes were attached to the following head locations *Cz, CPz, P1, P2, P3, P4, F3* and *F4,* as in the 10/10 extended international system [10] (see Figure 3 with a user wearing the EEG cap connected to the amplifier during an experiment). The ground and reference electrodes were attached at *FCz* and the left earlobe head location, respectively.

The captured online EEG signals were processed by the in-house enhanced *BCI2000* application [11,12], using an SWLDA classifier with features drawn from 0 ∼ 1000 ms ERP intervals. The sampling frequency was set at 500 Hz, a high-pass filter at 0.1 Hz and the low-pass filter at 40 Hz, with an electric power line interference rejection notch filter set in a band of 48 ∼ 52 Hz.

Fig. 3. The experiment environment in which the spatial auditory BCI-speller paradigm trials were conducted using the *BCI2000* software [9] to capture and classify the brainwaves. The spatial accosting environment was created using the loudspeakers depicted in the photograph and managed with a *MAX* programme developed in-house. The BCI user depicted is wearing an EEG cap with active electrodes *g.LADYbird* by g.tec Medical Instruments GmbH, Austria, connected to an amplifier *vAmp* by Brain Products GmbH, Germany.

The sound stimuli were presented randomly with an inter-stimulus interval (ISI) set at 300 ms. Each sound stimulus had a length of 300 ms. The brainwave response signals (ERPs) in each classification step were averaged 20 times (in the classifier training sessions) and only 10 times (in the two-step input spelling paradigm tests).

3. Results

This section presents and discusses the results obtained from the EEG experiments conducted with the five participating users in the proposed two-step input spatial auditory BCI-speller paradigm.

The grand mean averaged ERP responses obtained from the five users are depicted in Figure 4 in the form of a time series, and in Figure 5 as area under the curve (AUC) and head topography plots. The P300 response is clearly depicted for *target* responses (purple lines in the figures) in the ranges of 400 ∼ 600 ms, as can be seen in Figure 4. In the AUC score graph in Figure 5, the *target* versus *non-target* response distributions analysis resulted in clear and possibly separable differences, including in the latency range of 400 ∼ 600 ms. These latencies clearly support BCI classification using the SWLDA classifier.

The two-step input BCI-speller accuracy results of the five participating users are summarized in Table 1. The maximum achievements even result in 100% scores for three of the five users. However, 0% scores are also observed, which are clearly a sign of the interface usage difficulty. The averaged scores of all the subjects are above the two-step input BCI-speller theoretical chance level of 4% (see bottom row in Table 1).

4. Conclusions

In this paper, we report the results obtained with a two step input Japanese syllabary BCI-speller paradigm using 25 characters as a pilot study before the application of the full 50-character *hiragana* set.

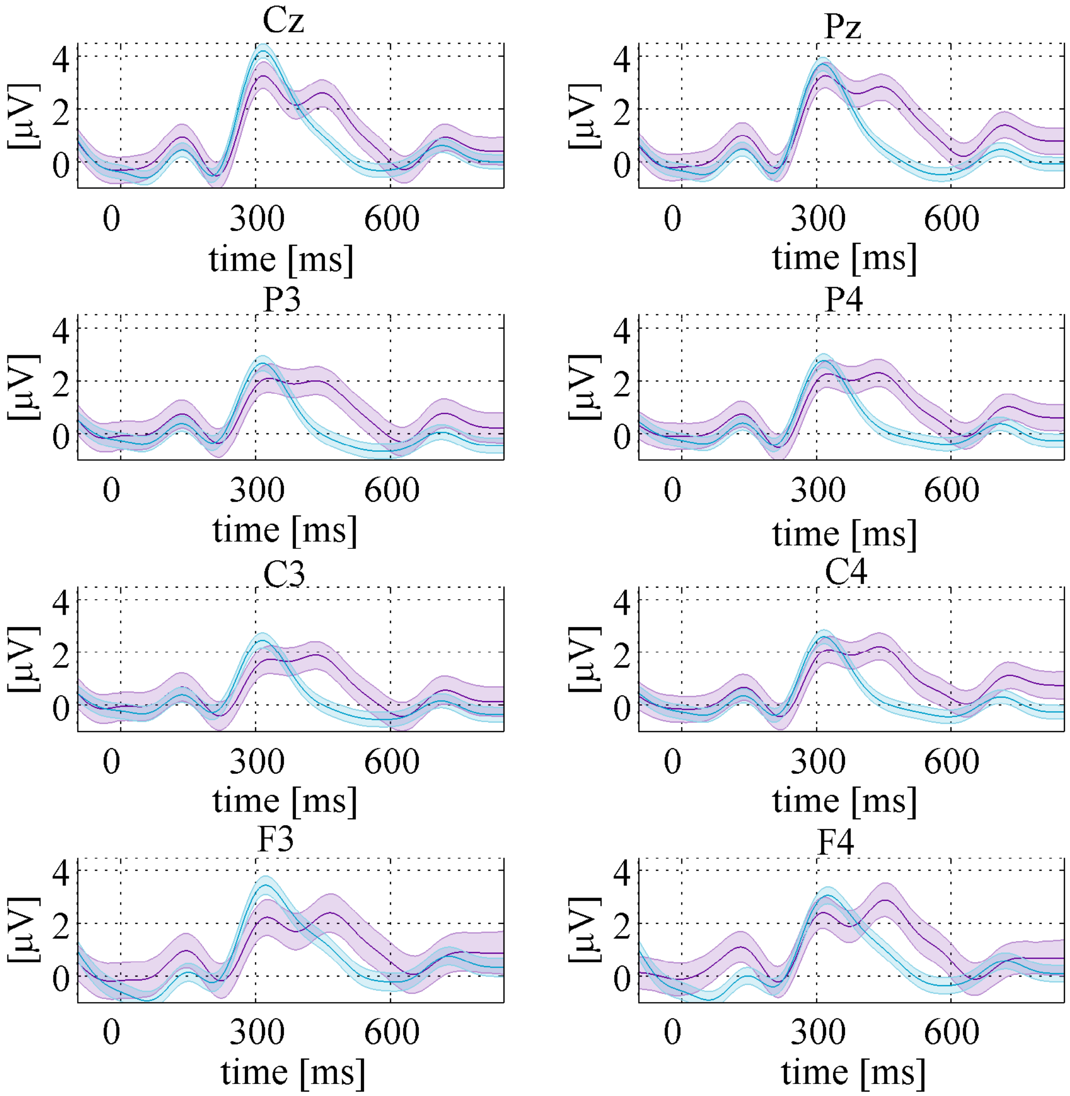

Fig. 4. The grand mean averaged ERP brain responses of all the users participating in the online BCI experiments, collected in the training sessions. The purple lines depict the ERPs for *targets*; (note the clear *P*300 positive deflections after about 400 ms and lasting until about 600 ms in this study.) The blue lines represent the averaged brainwaves for *non-targets*.

The results obtained allowed the five participating users to spell short words using two steps, as reported in the Table 1. The mean results of all the users were above the theoretical chance levels, and three of the five managed to obtain perfect results. However, zero accuracies were also observed.

The preliminary yet encouraging results presented call for more research into two-step input spatial auditory BCI-speller paradigms. The next research steps will include the spatial auditory stimulus optimization for handicapped or bedridden subjects, who cannot utilize a full surround sound acoustic environment. We will also extend the interface to the full set of 50 characters.

We plan to continue this research in order to apply the method in an online BCI application for patients suffering from locked-in syndrome.

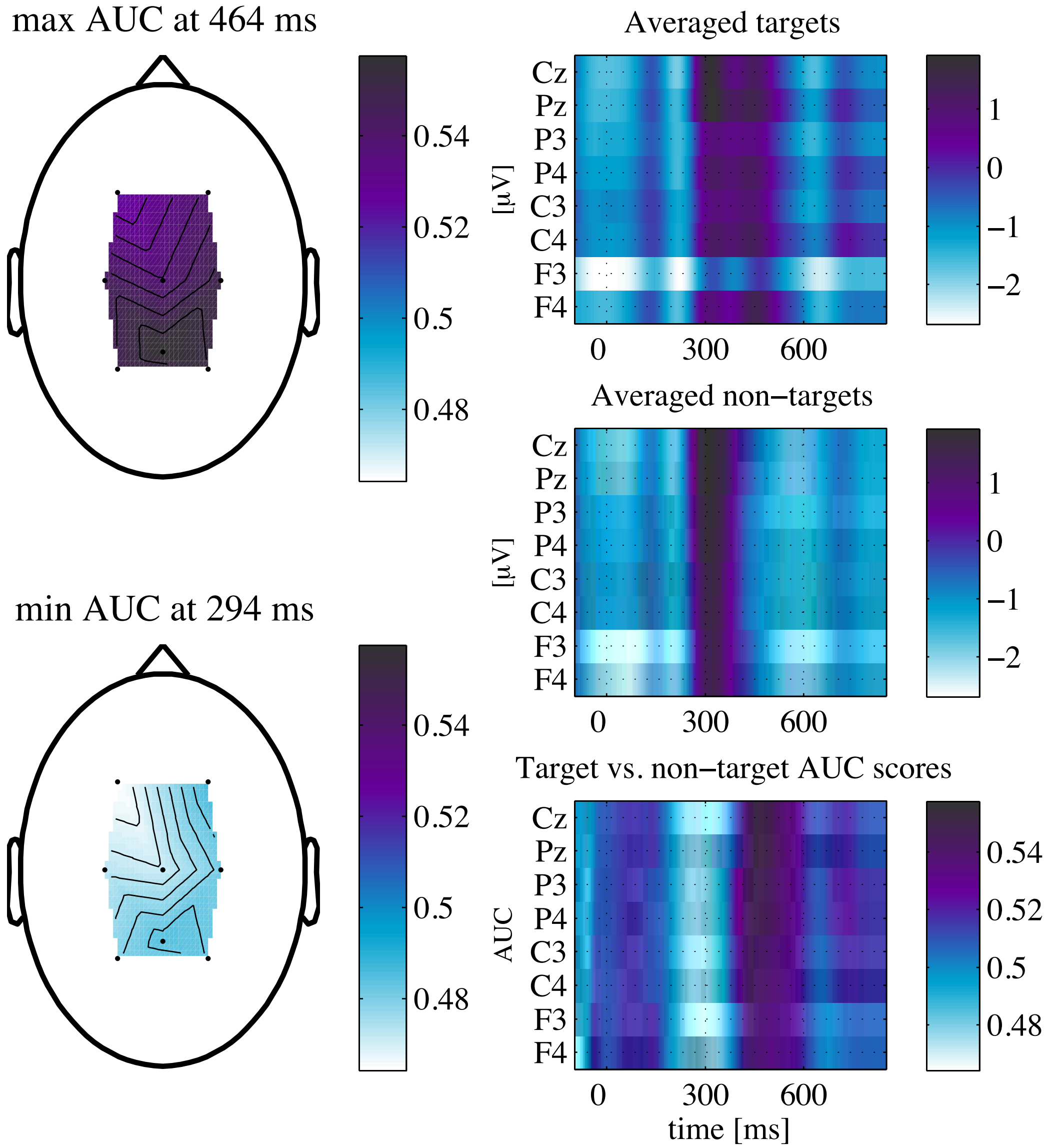

Fig. 5. Grand mean averaged ERP and AUC scores leading to final classification results for the participants. The panels in the left column present the head topographies at the maximum and minimum AUC scores as obtained from the bottom right panel. The top right panel presents averaged ERP responses to the *target* and the middle panel to the *non-target* stimuli, respectively. The bottom right panel depicts the AUC analysis results of the *target* versus *non-target* response distribution differences.

Acknowledgments

We would like to express our gratitude to Yoshihiro Matsumoto, BCI lab group and Multimedia Laboratory alumnus, who supported the project in the early stages with programming of *BCI2000* updates.

This research was supported in part by the Strategic Information and Communications R&D Promotion Programme no. 121803027 of The Ministry of Internal Affairs and Communication in Japan.

We also acknowledge the technical support of YAMAHA Sound & IT Development Division in Hamamatsu, Japan.

Table 1. Spelling accuracies of the two-step input Japanese syllabary paradigm obtained from the online BCI experiments with a theoretical chance level of 4%

		Spelling accuracy of the users				
Spelled syllables	Run	#1	#2	#3	#4	#5
sa-ka-na	#1	0.0%	33.3%	33.3%	0.0%	33.3%
	#2	0.0%	0.0%	33.3%	0.0%	**100.0**%
ki-se-tsu	#1	66.7%	0.0%	33.3%	0.0%	66.7%
	#2	33.3%	0.0%	66.6%	33.3%	66.7%
te-su-to	#1	66.7%	33.0%	33.3%	33.3%	33.3%
	#2	66.7%	33.0%	**100.0**%	0.0%	66.7%
ne-ku-ta-i	#1	25.0%	50.0%	**100.0**%	0.0%	50.0%
	#2	**100.0**%	0.0%	75.0%	0.0%	50.0%
shi-i-ta-ke	#1	50.0%	0.0%	**100.0**%	25.0%	75.0%
	#2	75.0%	25.0%	75.0%	0.0%	75.0%
	Averaged scores:	**48.3**%	**17.5**%	**65.0**%	**9.2**%	**61.2**%

References

[1] Wolpaw, J., Winter Wolpaw, E., editors. Brain-Computer Interfaces: Principles and Practice. Oxford University Press; 2012.

[2] Chang, M., Nishikawa, N., Struzik, Z.R., Mori, K., Makino, S., Mandic, D., et al. Comparison of P300 responses in auditory, visual and audiovisual spatial speller BCI paradigms. In: Proceedings of the Fifth International Brain-Computer Interface Meeting 2013. Asilomar Conference Center, Pacific Grove, CA USA: Graz University of Technology Publishing House, Austria; 2013, p. Article ID: 156. doi:`10.3217/978-3-85125-260-6-156`.

[3] Matsumoto, Y., Nishikawa, N., Makino, S., Yamada, T., Rutkowski, T.. Auditory steady-state response stimuli based bci application - the optimization of the stimuli types and lengths. In: Signal Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific. 2012, p. 1–7. URL: `http://www.apsipa.org/proceedings_2012/papers/285.pdf`; paper ID 285.

[4] Rutkowski, T.M., Cichocki, A., Mandic, D.P.. Spatial auditory paradigms for brain computer/machine interfacing. In: International Workshop On The Principles and Applications of Spatial Hearing 2009 (IWPASH 2009) - Proceedings of the International Workshop. Miyagi-Zao Royal Hotel, Sendai, Japan; 2009, p. P5. doi:`10.1142/9789814299312_0025`.

[5] Rutkowski, T.M., Mori, H.. Tactile and bone–conduction auditory brain computer interface for vision and hearing impaired users. Journal of Neuroscience Methods 2014;:(available online 21 April 2014)URL: `http://dx.doi.org/10.1016/j.jneumeth.2014.04.010`. doi:`10.1016/j.jneumeth.2014.04.010`.

[6] Chang, M., Makino, S., Rutkowski, T.M.. Classification improvement of p300 response based auditory spatial speller brain-computer interface paradigm. In: TENCON 2013 - 2013 IEEE Region 10 Conference (31194). 2013, p. 1–4. URL: `http://dx.doi.org/10.1109/TENCON.2013.6718454`. doi:`10.1109/TENCON.2013.6718454`; paper ID 111.

[7] Amano, S., Sakamoto, S., Kondo, T., Suzuki, Y.. NTT-Tohoku University familiarity controlled word lists (FW03). Speech Resources Consortium; National Institute of Informatics; Japan; 2006.

[8] Max 6. 2012. URL: `http://cycling74.com/`.

[9] Schalk, G., Mellinger, J.. A Practical Guide to Brain–Computer Interfacing with BCI2000. Springer-Verlag London Limited; 2010.

[10] Jurcak, V., Tsuzuki, D., Dan, I.. 10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems. NeuroImage 2007;34(4):1600 – 1611. URL: `http://www.sciencedirect.com/science/article/pii/S1053811906009724`. doi:`10.1016/j.neuroimage.2006.09.024`.

[11] Matsumoto, Y., Makino, S., Mori, K., Rutkowski, T.M.. Classifying P300 responses to vowel stimuli for auditory brain-computer interface. In: Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2013 Asia-Pacific. 2013, p. 1–5. URL: `http://www.apsipa.org/proceedings_2013/papers/388_Classifying-P300-Matsumoto-2933943.pdf`. doi:`10.1109/APSIPA.2013.6694341`; paper ID 388.

[12] Matsumoto, Y.. Application of stimulus-based ensemble classifier to spatial auditory BCI speller. Master thesis; Graduate School of Systems and Information Engineering, University of Tsukuba; Tsukuba, Japan; 2014.