

Brain-Computer Interface Control of Smartphone Messaging Applications

Francisco Velasco-Álvarez, Álvaro Fernández-Rodríguez, Ricardo Ron-Angevin

Departamento de Tecnología Electrónica

Universidad de Málaga

Málaga, Spain

e-mail: {fvelasco, afernandezr, rron}@uma.es

Abstract— This work-in-progress paper presents an implementation of a Brain-Computer Interface (BCI) system focused on the control of the most common messaging applications of a smartphone: WhatsApp, Telegram, e-mail and Short Message Service (SMS). The control of these applications is achieved through the use of a virtual assistant running in the smartphone. The BCI system is based on the visual Row-Column Paradigm (RCP), which allows users to select several control commands and to spell messages that are converted to synthesized voice and received by the mentioned virtual assistant in the smartphone.

Keywords- Brain-Computer Interface (BCI); P300; assistive technology; virtual assistant; messaging applications.

I. INTRODUCTION

Brain-Computer Interfaces (BCI) are a type of Assistive Technology (AT) that uses the brain signals of users to establish a communication and control channel between them and an external device (usually a computer) [1]. BCI systems may be a suitable tool to restore communication skills in severely motor-disabled patients, as BCI do not rely on muscular control. There are several diseases that cause severe impairment of motor skills in affected patients, such as Amyotrophic Lateral Sclerosis (ALS). AT can be used to control multiple devices, such as a wheelchair, a home automation system, or a verbal communication system [2]. AT should be able to be controlled through those output channels that the patient still has preserved, such as the voice, the eye gaze, movements of a finger, the head, the cheek or the tongue. However, in severe and progressive motor limitations (as is the case in ALS), most of these examples of AT may no longer be useful because they depend on some type of muscular channel that may be affected in the patient. In these cases, BCI may be a suitable option for those people who have completely lost the ability to move their muscles. The neuroimaging technique most used by BCIs is electroencephalography (EEG), possibly due to its relatively low cost and high temporal resolution [3]. One of the most extended EEG signal used in BCI system is the P300 evoked potential. P300 is a positive potential generally located in the parieto-occipital areas that appears about 250-500 ms (although the range can vary depending on numerous factors) after the presence of both an expected stimulus and a rare one [4]. Usually, the P300 is evoked through an oddball paradigm, in which the available items are highlighted pseudo-randomly while the user pays attention only to the desired item, thus resulting in a P300

potential after the stimulation of this desired item. After a predetermined number of iterations, the system averages the resulting EEG and determines which item the user wanted to select. This concept was adopted by [5] to propose a paradigm to control a text speller.

The loss of communication (mainly with family and caregivers) is considered by ALS patients as even more negative than the loss of physical aspects [6][7]. Therefore, this work will focus on the use of a BCI that could allow patient communication through some of the most common messaging applications on a smartphone: WhatsApp, Telegram, e-mail and Short Message Service (SMS). Other researchers have focused on the BCI control of daily use and domestic applications, for example the work in [8] uses visual event-related potentials to control a TV, an air conditioner and to make an emergency call. The authors in [9] propose a BCI control of Telegram and Twitter in a smartphone. A hybrid BCI (that uses as well electrooculography as input) presented in [10] allowed to control a web browser and an e-mail client. To our knowledge, the present work in progress is the first BCI proposal that allows to control some of the previously used applications by other BCI works (Telegram, e-mail) and WhatsApp and SMS as well.

There are some works that have already explored the idea of using voice commands sent to virtual assistants to facilitate integration between applications. Outside the field of BCI, the work in [11] can be highlighted. The authors used a proximity sensor on the fingers, feet or head (depending on the patient) to select commands in an application with a graphical interface which later allowed text to be converted to speech to verbalise the users' selections. The voice assistant used for this work was Google Assistant. This system allowed to control WhatsApp and YouTube. This work was based on a reduced set of possible actions: for WhatsApp, three contacts to choose from and three predetermined possible messages to send; for YouTube, three possible music/videos alternatives and four alternative-related events. In relation to the works that have used a BCI, the work in [12] has been the only one, to our knowledge, that has used a voice assistant (Amazon's Alexa) to control two devices (a light bulb and a fan).

The objective of this work is to present a communication bridge between the UMA-BCI Speller platform [13] (a BCI software developed by UMA-BCI group of the University of Malaga, UMA) and the messaging services of WhatsApp, Telegram, e-mail and SMS, through the use of Google Assistant on a smartphone. This paper presents a short

version of a full-length work that is currently under review [14].

This paper is organized as follows: section 2 and section 3 describe the system implementation and the control paradigm, respectively. The preliminary results are presented in section 3, followed by the conclusion and future works in section 4.

II. SYSTEM IMPLEMENTATION

The aim of the BCI system was to generate voice commands that could be interpreted by a virtual assistant running in a smartphone. These voice commands were intended to read and send messages through various messaging services. In order to achieve that, a BCI system was implemented that could generate these commands in text form and convert them into voice. On the one hand, a laptop ran the software that presented the stimuli and registered and analysed the EEG. This software was the UMA-BCI Speller [13], a free tool that wraps BCI2000 [15] and simplifies its configuration and use. This software was used to spell and convert the control commands to voice. On the other hand, a virtual assistant was running on a smartphone. This virtual assistant was Google Assistant and it received and interpreted the voice commands, performing the corresponding action. The system implementation is shown in Figure 1. Through the EEG, the user can select commands from a computer control (in the example, “Read SMS”) and spell messages to be sent. These commands and messages are converted to voice and received by the virtual assistant running in the smartphone. The assistant interprets them and performs the corresponding action (in the example, it informs the user that there is a SMS and it reads it).

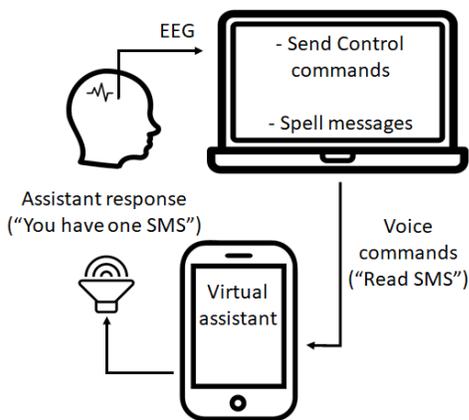


Figure 1. System implementation.

The UMA-BCI Speller includes a text prediction function that may help users when spelling words. As users choose the characters of a word (starting with the first character), the system proposes several predicted words based on the characters already written and the probability of occurrence based on a Spanish language specific corpus.

Once users had completed the spelling of the text command to send, they had to select a confirmation item in the interface so that the system could convert this command into speech. The Windows 10 Narrator (a text-to-speech feature) was used, particularly the voice named “Microsoft Helena” from the Spanish voice catalogue.

To avoid the influence of ambient noise on the understanding of the command by the virtual assistant, the voice commands were sent to the smartphone via a cable connection (using a mini-jack audio cable) connecting the laptop audio output with the smartphone microphone input. The output volume of the laptop was fixed throughout the whole experiment, so the assistant always received the same level of audio.

As the virtual assistant used in the experiment was Google Assistant, each command started with the words “Ok Google...”, which is one of the wake-up keywords of the assistant. Two main types of voice commands were used: i) commands asking the assistant to read the received messages, e.g., “Ok Google, read my WhatsApp messages”, and ii) commands asking the assistant to send a message using one of the messaging services installed in the smartphone, e.g., “Ok Google, send a Telegram to [contact], [message]”. In addition to these, other voice commands were used that were needed to confirm or cancel actions, as we will explain below.

III. CONTROL PARADIGM

In order to send a command to the virtual assistant, users had to select items from different menus. The selection of an item followed the usual procedure in a P300 Row-Column Paradigm (RCP): users had to pay attention to the desired item (within a matrix of possible items) and mentally count the number of times it was highlighted. The timing of each selection for all the menus was the same, as all the interfaces consisted of a 7×7 matrix, even though in three of them there were dummy items (items that had no effect when selected). An item was selected after all the seven rows and columns were highlighted a certain number of times.

Four menus were implemented that allowed subjects to gradually form a sentence that would finally be converted to speech by the Windows 10 Narrator voice synthesiser (from now on, this conversion will be denoted as “speak”). The sentence to be spoken was present in the interface so that subjects decided when to indicate to the system to speak it. Some items added several predetermined words to this sentence, while other items were present to add individual letters to spell a message. The four menus are described next:

A. No Control (NC) menu.

This was a 7×7 matrix in which only one item was a valid command, and the other 48 items (“X”) were dummy commands. The objective of this menu was to allow subjects to remain in a state where they could rest without generating control commands; the term “no control” is generally used in asynchronous systems to refer to such a state. The only valid command was a control command named “IC” whose selection changed the menu to an Intentional Control (IC) menu.

B. Intentional Control (IC) menu.

This was the main menu of the system, where subjects could choose what action they wanted to select. In a 7×7 matrix, ten valid options were available (the remaining 39 options were dummy non-visible items). This menu is shown in Figure 2. This menu presented ten options that can be grouped into three categories:

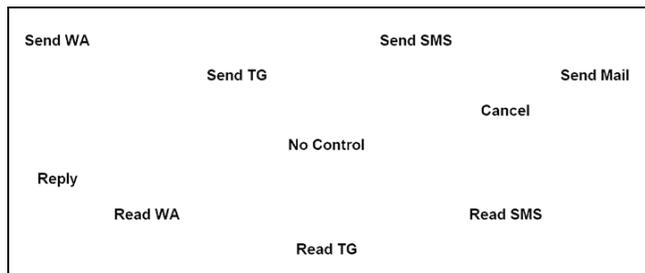


Figure 2. Intentional Control (IC) menu.

- **Send messages.** This group consisted of four commands: “Send WA”, “Send TG”, “Send SMS” and “Send Mail”, that enabled users to send a message using WhatsApp, Telegram, SMS or e-mail, respectively. Once one of these commands was selected, the system wrote part of the sentence to be spoken, “Ok Google, send a WhatsApp to”, “Ok Google, send a Telegram to”, “Ok Google, send a SMS to” or “Ok Google, send an e-mail to” and then changed to a Spelling menu (Figure 3), so that the user could next spell out the receiver of the message and the message itself.
- **Read messages.** Three commands formed this group: “Read WA”, “Read TG” and “Read SMS” used to read the messages received through WhatsApp, Telegram and SMS, respectively. The selection of one of these commands made the system speak the corresponding sentence: “Ok Google, read my WhatsApp messages”, “Ok Google, read my Telegram messages” or “Ok Google, read my SMS messages”. After the sentence was spoken, the system deleted it and automatically changed to the NC menu so that subjects could listen to the received messages, if any. After the virtual assistant read each received message, it asked the users if they wanted to reply to it or not. To do this, subjects first had to change to the IC menu where two commands were available related to replying to messages. These commands will be explained in the third group. After cancelling or replying to each message, the system continued to read the remaining messages, if any.
- **Other commands.** In this group, three commands were included: “Reply”, “Cancel” and “No Control”. The “Reply” command allowed users to reply to a WhatsApp, Telegram

or SMS received message after the system read them. Once this command was selected, the system wrote the sentence “Ok Google, reply” and changed to the Spelling menu so that users could complete the sentence with the desired response (in a similar way to what was done with the “Send” commands); please note that in this case it is not necessary to specify the receiver of the message to be sent. The “Cancel” command was used to indicate to the system that the user did not want to reply to a received message. Once it was selected, the system wrote and spoke the sentence “Ok Google, cancel” and then deleted it and changed to the NC menu. Finally, the “No Control” command was presented in order to allow subjects to voluntarily change to the NC menu, in case they wanted to take a rest.

C. Spelling menu.

When users selected one of the “Send” or “Reply” options from the IC menu, the system changed to the Spelling menu (after adding some predetermined text). This menu is shown in Figure 3. Here, the users could spell out

Ok_google_Send_Whatsapp_to_F						
A	B	C	D	E	F	francisco
G	H	I	J	K	L	fue
M	N	O	P	Q	R	forma
S	T	U	V	W	X	fueron
Y	Z	1	2	3	4	frente
5	6	7	8	9	0	fin
SPC	OK	Del.	IC	Del. W	,	fuera

Figure 3. Spelling menu with 38 characters to spell, four control commands and seven available predicted words (last column, in Spanish)

the receiver and the message to send (or just the message in the case of the option “Reply”). It is worth mentioning that the message had to be spelled right after the receiver, only separated by a space. This menu consisted of a 7×7 matrix with spelling and control commands. The first six columns and rows corresponded to specific characters to be added (English alphabet letters and numbers). The last column was used to provide subjects with seven predicted words. The last row contained two characters (“SPC” (space) and “,”), two delete commands (“Del.” to delete a single character and “Del. W” to delete a complete word) and two control commands (“OK” and “IC”). The command “OK” was used to indicate to the system that the receiver (if needed) and the message to send were complete so that the written sentence could be spoken and interpreted by the virtual assistant. The “IC” command was used to return to the IC menu without generating any voice command (this was useful if a subject entered this menu unintentionally). The selection of “OK” or “IC” caused the current written sentence to be deleted (after speaking it in the case of “OK”), so a confirmation menu

was offered to subjects in order to avoid undesired selections of these two commands.

D. Confirmation menu.

Two valid commands were available (among other 47 non-visible dummy options) in a 7×7 matrix, “Confirm” and “Back”. On the one hand, the “Confirm” command was used to corroborate the previous selection in the Spelling menu (that is, “OK” or “IC”). In the case of confirming an “OK” command, the system spoke (and deleted) the complete sentence so it could be interpreted by the virtual assistant, and it changed to the NC menu. In the case of confirming an “IC” command, the system deleted the written sentence and changed to the IC menu. On the other hand, the “Back” command was used to return to the Spelling menu in order to continue writing the sentence to be sent to the virtual assistant.

IV. RESULTS

Some preliminary tests have been carried out. In these, some healthy volunteers were asked to perform four tasks related to the four messaging applications: i) send a WhatsApp message with a predetermined message; ii) read the incoming SMS; iii) read an incoming Telegram message and reply to it with a free answer; and iv) send a free text e-mail to a contact chosen by them. The preliminary online results obtained, as well as the results of some questionnaires related to the subjective experience controlling the interface, support the viability of the proposed system.

V. CONCLUSION AND FUTURE WORK

The use of a virtual assistant to control the smartphone makes it possible to easily extend the functionality to other applications beyond messaging services, for example to control domotic devices through the smartphone. However, the use of a virtual assistant also presented some misinterpretation problems when the synthesized voice was not correctly understood by the assistant. Other issue related to the use of the virtual assistant arose when the assistant responded to a command in an unexpected way, as the system was based in a one-way communication (from the laptop to the smartphone), taking for granted that the virtual assistant would interpret this command correctly.

The future work based on this study is related to the extension of the functionality to domotic features, the improvement of the one-way communication from the BCI application to the virtual assistant, and the possibility of testing the system with motor-disabled patients.

ACKNOWLEDGMENT

This work was partially supported by the project SICCAU: RTI2018-100912-B-100 (MCIU/AEI/FEDER, UE) and by the University of Malaga (“Universidad de Málaga”).

REFERENCES

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, “Brain-computer interfaces

for communication and control,” *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, 2002, doi: 10.1016/S1388-2457(02)00057-3.

[2] R. Jamwal, H. K. Jarman, E. Roseingrave, J. Douglas, and D. Winkler, “Smart home and communication technology for people with disability: a scoping review,” *Disabil. Rehabil. Assist. Technol.*, vol. 0, no. 0, pp. 1–21, 2020, doi: 10.1080/17483107.2020.1818138.

[3] L. F. Nicolas-Alonso and J. Gomez-Gil, “Brain computer interfaces, a review,” *Sensors*, vol. 12, no. 2, pp. 1211–1279, 2012, doi: 10.3390/s120201211.

[4] J. Polich, “Updating P300: An integrative theory of P3a and P3b,” *Clin. Neurophysiol.*, vol. 118, no. 10, pp. 2128–2148, Oct. 2007, doi: 10.1016/J.CLINPH.2007.04.019.

[5] L. A. Farwell and E. Donchin, “Talking off the top of your head: toward a mental prosthesis utilizing event-related potentials,” *Electroencephalogr. Clin. Neurophysiol.*, vol. 70, no. 6, pp. 510–523, 1988.

[6] Z. Simmons, B. A. Bremer, R. A. Robbins, S. M. Walsh, and S. Fischer, “Quality of life in ALS depends on factors other than strength and physical function,” *Neurology*, 2000, doi: 10.1212/WNL.55.3.388.

[7] S. H. Felgoise, J. L. Stewart, B. A. Bremer, S. M. Walsh, M. B. Bromberg, and Z. Simmons, “The SEIQoL-DW for assessing quality of life in ALS: Strengths and limitations,” *Amyotroph. Lateral Scler.*, vol. 10, no. 5–6, pp. 456–462, 2009, doi: 10.3109/17482960802444840.

[8] K. T. Sun, K. L. Hsieh, and S. R. Syu, “Towards an accessible use of a brain-computer interfaces-based home care system through a smartphone,” *Comput. Intell. Neurosci.*, vol. 2020, pp. 16–18, 2020, doi: 10.1155/2020/1843269.

[9] V. Martínez-Cagigal, E. Santamaría-Vázquez, J. Gomez-Pilar, and R. Hornero, “Towards an accessible use of smartphone-based social networks through brain-computer interfaces,” *Expert Syst. Appl.*, vol. 120, pp. 155–166, 2019, doi: 10.1016/j.eswa.2018.11.026.

[10] S. He et al., “EEG- And EOG-Based Asynchronous Hybrid BCI: A System Integrating a Speller, a Web Browser, an E-Mail Client, and a File Explorer,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 2, pp. 519–530, 2020, doi: 10.1109/TNSRE.2019.2961309.

[11] G. E. Lancioni et al., “Mainstream technology to support basic communication and leisure in people with neurological disorders, motor impairment and lack of speech,” *Brain Inj.*, vol. 00, no. 00, pp. 1–7, 2020, doi: 10.1080/02699052.2020.1763462.

[12] V. K. K. Shivappa, B. Luu, M. Solis, and K. George, “Home automation system using brain computer interface paradigm based on auditory selection attention,” *I2MTC 2018 - 2018 IEEE Int. Instrum. Meas. Technol. Conf. Discov. New Horizons Instrum. Meas. Proc.*, pp. 1–6, 2018, doi: 10.1109/I2MTC.2018.8409863.

[13] F. Velasco-Álvarez, S. Sancha-Ros, E. García-Garaluz, Á. Fernández-Rodríguez, M. T. Medina-Juliá, and R. Ron-Angevin, “UMA-BCI Speller: An easily configurable P300 speller tool for end users,” *Comput. Methods Programs Biomed.*, vol. 172, pp. 127–138, Mar. 2019, doi: 10.1016/j.cmpb.2019.02.015.

[14] F. Velasco-Álvarez, A. Fernández-Rodríguez, F. J. Vizcaíno-Martín, A. Díaz-Estrella, and R. Ron-Angevin, “Brain-computer interface (BCI) control of a virtual assistant in a smartphone to manage messaging applications.”, unpublished.

[15] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, “BCI2000: A general-purpose brain-computer interface (BCI) system,” *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1034–1043, 2004, doi: 10.1109/TBME.2004.827072.