Brain-Operated Typewriter using the Language Prediction Model

Saebyeok Lee and Heuiseok Lim
Dept. of Computer Science Education, Korea University
Seoul, Republic of Korea
[e-mail: saebyeok@blp.korea.ac.kr, limhseok@korea.ac.kr]
*Corresponding author: Heuiseok Lim

Received March 31, 2011; revised July 20, 2011; accepted August 30, 2011; published October 31, 2011

Abstract

A brain-computer interface (BCI) is a communication system that translates brain activity into commands for computers or other devices. In other words, BCIs create a new communication channel between the brain and an output device by bypassing conventional motor output pathways consisting of nerves and muscles. This is particularly useful for facilitating communication for people suffering from paralysis. Due to the low bit rate, it takes much more time to translate brain activity into commands. Especially it takes much time to input characters by using BCI-based typewriters. In this paper, we propose a brain-operated typewriter which is accelerated by a language prediction model. The proposed system uses three kinds of strategies to improve the entry speed: word completion, next-syllable prediction, and next word prediction. We found that the entry speed of BCI-based typewriter improved about twice as much through our demonstration which utilized the language prediction model.

Keywords: Language prediction model, brain-computer interface, Korean typewriter

The part of this paper was presented in the ICONI (International Conference on Internet) 2010, December 16-20, 2010, Philippines. This version includes a concrete analysis and supporting implementation results. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (No. 2011-0018014).

DOI: 10.3837/tiis.2011.10.005

1. Introduction

Computers have become necessary instruments for communicating with the external world. Most people are used to using applications such as messengers and blogs in order to communicate with the computer to transmit their opinions or thoughts. A computer which transmits messages to the external world is not only a sucessful instrument but also extends to communication channels. People who are totally paralyzed, however, cannot communicate by themselves. A brain-computer interface(BCI) is a new communication channel between human and machines that include computers or other devices [1][2]. It allows for communication and control without any neuromuscular ability [1]. Therefore, it is useful for “locked-in” patients: patients who have lost voluntary muscle control through neurological diseases such as ALS or brainstem stroke [3]. Such patients are mentally fit. They are able to hear, see and understand everything, but they are not able to express themselves physically through their muscle control [1].

The role of BCIs is to transfer commands by measuring various mental states using recording methods such as electroencephalography (EEG), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), and electrocorticogram (ECoG) [4]. Non-invasive BCI generally involves EEG-based recording. It is less expensive and easy to use in experiments compared to other invasive BCI methods, but it is still quite expensive to be used for practical purposes. Recently, BCIs have appeared in portable devices which measure EEG, and have arisen as a secondary controller in games and neuro-feedback applications [5][6]. These devices provide comparatively low-priced software development kit (SDK) which includes functions to control EEG raw data.

Typically, the research on BCI has mainly focused on measuring various reliable mental states through the use of signal processing and/or classification methods [7][8]. Nevertheless, most of the practical BCI systems depend on binary decisions [9][10[11]. However, designing an interface that makes use of the binary decisions alone is very limited.

The present study has focused on the practicality of the BCI system by improving its communication rate that can be applied directly in real life which deviates from the traditional BCI research. Therefore, in order to increase the speed of character input, three kinds of language prediction models were used. Since the majority of research has focused on the English language, the present study emphasizes the use of the Korean language and its specific features as there are no typewriters that have considered Korean. Also, in order to overcome the limitations of devices used in BCI researches, a low-cost headset was utilized to acquire EEG signals.

2. Related Work

Though the idea of using brain signals as an input to BCIs has existed since the initial conception of BCIs, actual working BCIs based on EEG input have only recently appeared [11]. BCI applications can be separated into communication and control. BCIs for control mainly deal with invasive recordings in order to control a wheelchair or a robot arm [12][13]. Most of the EEG based BCI systems have a cycle of 4 parts. First, signals from the brain are acquired by electrodes on the scalp. Second is to find features that reflect the user’s intention through the signal processing. Third, the signal is processed and assigned to different classes using the machine learning technique like a linear discriminant analysis (LDA), an artificial neurul network (ANN), and a support vector machine (SVM). Finally, different classes are sent to output computers or other devices for actual commands [1][4][9][10]. The EEG classification method depends on the response of the user to detect commands, which include event-related potential (ERP), steady-state evoked potential (SSVEP), and motor imagery or slow cortical potential (SCP). The detection of these features from EEG signals is determined by the classification of certain feature extraction methods. In this paper, we focus on the non-invasive BCI systems which allows for communication by using text as its input.

In **Fig.1**, (a) shows the first BCI system that was introduced by Farwell and Donchin in 1988 which was a P300 Speller designed with 6x6 matrix. [15]. The P300 is a large positive potential that appears around 300 ms after the stimulus presentation . It is possible to generate P300 through the oddball paradigm. This paradigm provides the presentation of random stimuli that causes a surprise effect on the subject. When the user focuses on the alphabet that they want to write on the cell of matrix, it is possible to detect a P300 which is time-locked to the onset of the cell intensification. For the past two decades since the introduction of the P300 Speller, the field of signal processing and graphical user interface have developed extensively and have obtained its stability [15][16].The current P300 Speller, developed as a commercial application by *g.tec* in 2009 which was named *intendiX*, is shown in **Fig. 1** (b) [17]. Accoring to *g.tec*, performance for the majority of healthy users during their first trial is 5 to 10 cpm (characters per minute).

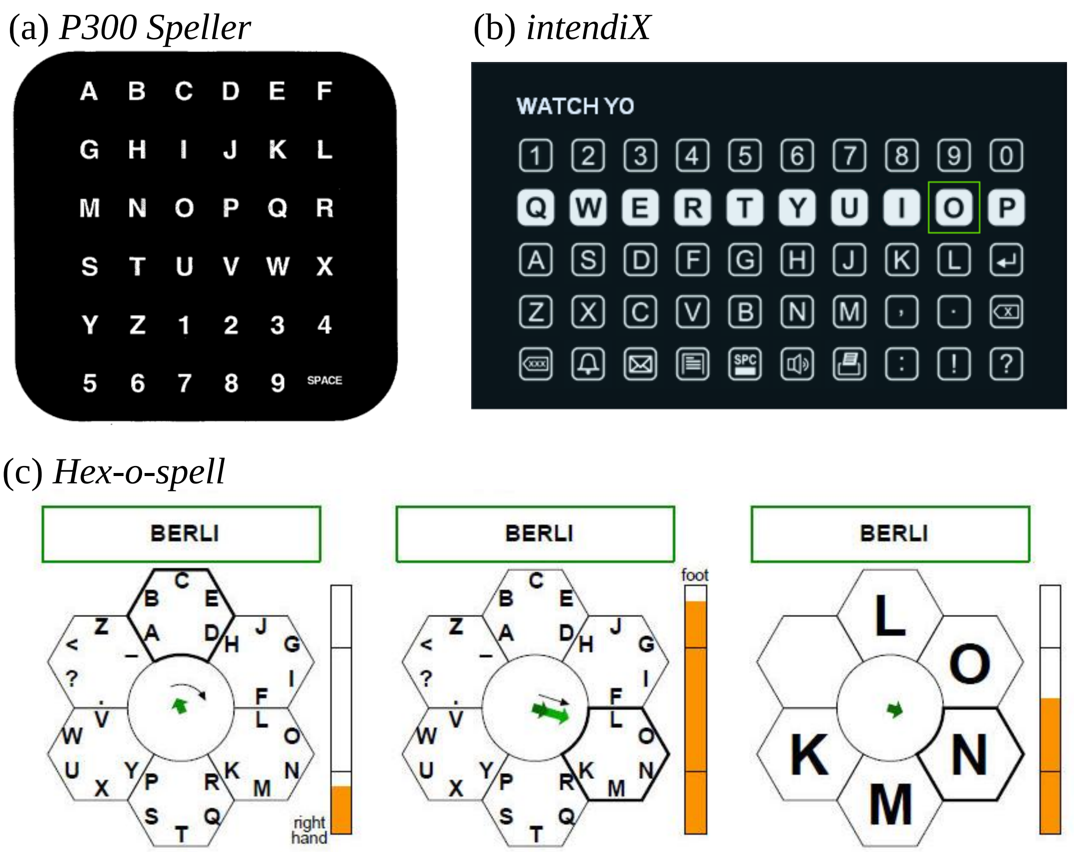

Fig. 1. Various BCI communication applications

A BCI research group from Fraunhofer FIRST IDA, Berlin, Germany has proposed the Berlin BCI called Hex-o-spell based on motor imagery [9]. This is controlled by two mental states: imagined right hand movement and imagined foot movement, as depicted in **Fig. 1** (c). When the user imagines a right hand movement, the arrow turns to the right. When the user imagines a foot movement, the rotation stops and the arrow starts extending to the desired field. The speed of the Hex-o-spell was found to be between 2.3 and 7.6 cpm when tested on the two volunteers at CeBIT 2006 in Hannover, Germany [9].

Both the P300 Speller and the Hex-o-spell facilitate communication without any physical movements. Each application, however, has its drawbacks. Both require a training session for the calibration of the system. The Hex-o-spell which is based on the use of motor imagery, suffers from BCI illiteracy. Performance of this system highly depends on the subject's ability to adapt to the system [10].

The purpose of this work is thus twofold: to improve efficiency of the BCI-based typewriters by using a language prediction model and to improve its usability by utilizing a portable and low-priced device which measures EEG signals. In this paper, we propose a brain operated Korean typewriter using the language prediction model by using a commercial device which measures user's EEG signals for Korean individuals who suffers from total paralysis.

3. Language Prediction Model

Languages have their sequences. Language modeling is done to solve these phenomena. It can predict the next word when given the previous words. This task is used in speech or optical character recognition, statistical machine translation, and these tasks are often referred to as a *shannon game* [18]. Here, we deal with a theoretical language model using the n-gram which is described in [19] and how it can be applied to Korean prediction modeling.

The n-gram model is a simple and strong probability model which adopts the modeling of actual language. *n* is the size of the previous word or syllable. When the words $w_1, \dots, w_m$ are given, the sequence for the probabilty in an n-gram model is as follows:

$$\begin{aligned} \mathrm{P}(w_1, \dots w_m) &= \prod_{i=1}^{m} P(w_i | w_1, \dots, w_{i-1}) \\ &\approx \prod_{i=1}^{m} P(w_i | w_{i-(n-1)}, \dots, w_{i-1}) \end{aligned} \quad (1)$$

Here, it is assumed that the probabillty of observing the n^{th} word w_i in the context history of the preceding $\mathrm{i}-1$ words can be approximated by the probabillity of observing it in the shortened context history of the preceding $\mathrm{n}-1$ words that follows the n^{th} order Makov property. For example, in a bigram(n=2) language model, the probabillity of the sentence "I saw the red house" is approximated as $\mathrm{P(I, saw, the, red, house)} \approx \mathrm{P(I|} < s >) \mathrm{P(saw|I)P(the|saw)\ P(red|the)\ P(house|red)}$. Each probability of n-grams can be calculated from the frequency counts in the corpus as shown below:

$$P(w_i | w_{i-(n-1)}, \dots, w_{i-1}) = \frac{Count(w_{i-(n-1)}, \dots, w_{i-1}, w_i)}{Count(w_{i-(n-1)}, \dots, w_{i-1},)} \quad (2)$$

When predicting the next word or next syllable, we most likely need w_i. This can be calculated as follows:

$$w_i = argmax_{w_i} P(w_1, \dots, w_i) \approx argmax_{w_i} P\left(w_{i-(n-1)}, \dots, w_i\right) \quad (3)$$

This expression occurs earlier in the corpus by calculating the frequency of syllables and words receiving the highest probability of the current syllable or word as recommended by the system. Accordingly, the number *n* refers to uni-gram, bi-gram and tri-gram. We use the word uni-gram, the syllable bi-gram, and the word bi-gram in order to predict the syllable. Further details of implementation are presented in section 4.3.

4. Korean Typewriter using the Language Prediction Model

When you are using a system to write, and if the proposed system is able to predict the next syllable, then you can speed up the process of writing. Generally, a computer using keyboard 101/103 cannot use this system because writing is faster than prediction. Several mobile phones, however, can predict input more quickly because they have a limited interface of 9 to 12 keys. Most BCI systems have a much more limited interface than mobile phones because they use binary decisions. Therefore, we propose a Korean typewriter using the language prediction model in order to overcome these interface limitations.

4.1 Overall Architecture

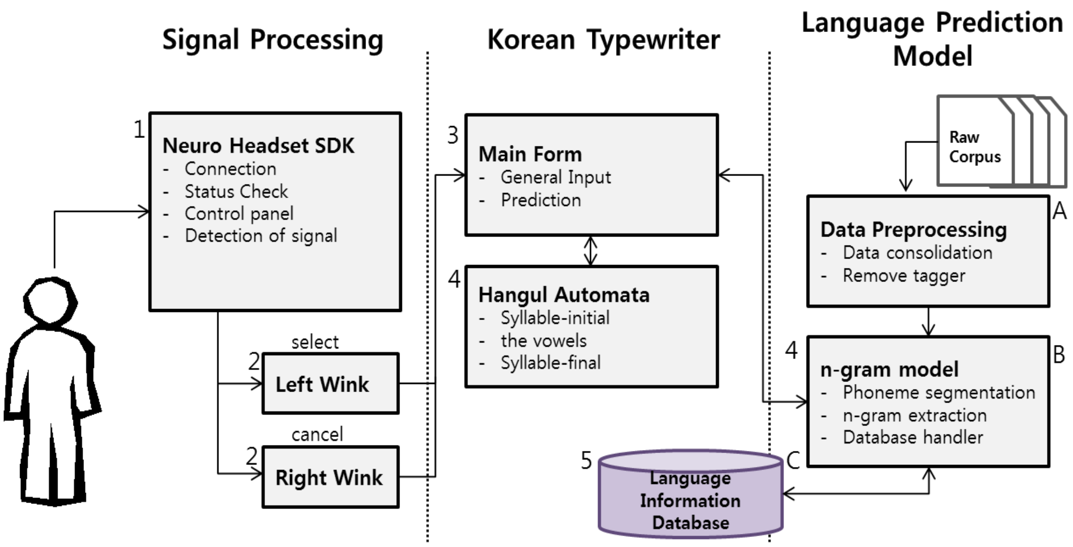

Fig. 2. Brain-operated Korean typewriter

As shown in **Fig. 2**, the whole system consists of three parts: a Korean typewriter, a language prediction model, and signal processing. With the Korean typewriter, user interface is provided with a general entry method and the result from the prediction model when given a syllable or a word. In addition, Hangul automata are used to combine characters (syllable-initial, medial vowels, syllable-final). The language prediction model includes n-gram models for prediction and a language information database. In section 3, we will describe the language prediction model. In this section, we will describe how the language prediction database is constructed. The signal processing part in **Fig. 2** deals with the selection of features and brain-operated signals using a neuro-headset from *Emotiv© System.*

4.2 Detection of Brain States

The main topic in BCI research is concerned with the detection of the brain state when the user is operating on a computer. However, in exploring the questions of measuring brain states, this paper will be limited to the consideration of using the equipment and use of different features. The epoc (shown **Fig. 3**) that we have used is SDK provided by neuro-headset which includes the detection of affective states, cognitive states, and facial expressions. The affective states cannot be used because 'affect' reflects the user's emotional states. The cognitive state includes thirteen different actions: six directional movements(push, pull, left, right, up and down); six rotations (clockwise, counter-clockwise, left, right, forward and backward); and one additional action: disappear. In order to use these states, the user needs an extensive amount of training. Therefore, we have used the facial expressions which do not require extensive training.. The facial expressions employed by epoc provide eleven different states: blink, right wink, left wink, look right, look left, raised brow, furrowed brow, smile, clenched mouth, right smirk, left smirk, and laugh. The left/right wink can be made artificially. Therefore, we selected the left/right wink. For binary decision, we defined the left wink as "selection" and right wink as "cancel."

Fig. 3. Epoc neuro-headset (right) and its electrode location(left)

4.3 Language Information Database

Implementing a language model must include the construction of a language information database. This requires a large raw corpus. Therefore, we built a database that extracted n-gram and its frequency by using the Sejong corpus. The Sejong corpus was compiled from the raw text in the 21st Century Sejong project carried out by the National Institute of the Korean Language. It consists of approximately 54 million words. As shown in **Fig. 4**, we built the database though preprocessing in order to extract the n-gram in the corpus. Constructing a database that includes n-gram language information is similar to constructing the "bag of words" in information retrieval systems. The first step is the data consolidation. It entails creating a single file from the individual files in the raw corpus. At the same time, data is removed as tag or metadata. After the completion of the first step, a file with a bag of words is assembled and is made into counts of the word uni-grams, syllable bi-grams, and word bi-grams. Here, we defined "word" as "Eojeol" which is the Korean-based spacing between continuous words. We also defined the unit of a syllable in Korean as "Jaso" which consists of three parts: syllable-initial, medial vowel, and syllable-final. The syllable bi-gram is to predict the next syllable. The word 'uni-gram' has to do with the completion of the word, and the word 'bi-gram' is concerned with the prediction of next word when given an input word.

To implement the databse, we used Microsoft Visual Studio 2010 and the C# programing language. A hash table was used to count each n-gram in order to speed up the calculation.At this step, memory overflow had to be considered. Hence we divided the database into different trials. Finally, the value which included the hashtable was inserted into the database system using the SQL query. Before inserting the database, low-frequency words were removed because it slowed down the query response time when all of the 54 million words were used. Therefore, a threshold for the removal had to be determined. **Fig. 5** shows the ratio of frequency coverage in the Sejong corpus.The A (blue) line is the frequency rate of unique words and B (red) line indicates the overall frequency rate. For example, if the overall frequency rate of the a word is 90% in the corpus and this word is included in the top 12.5% of unique words frequency rate, then this particular word will appear over 5 times in the corpus. Therefore, we removed the words that appeared less than 5 times(in frequency).

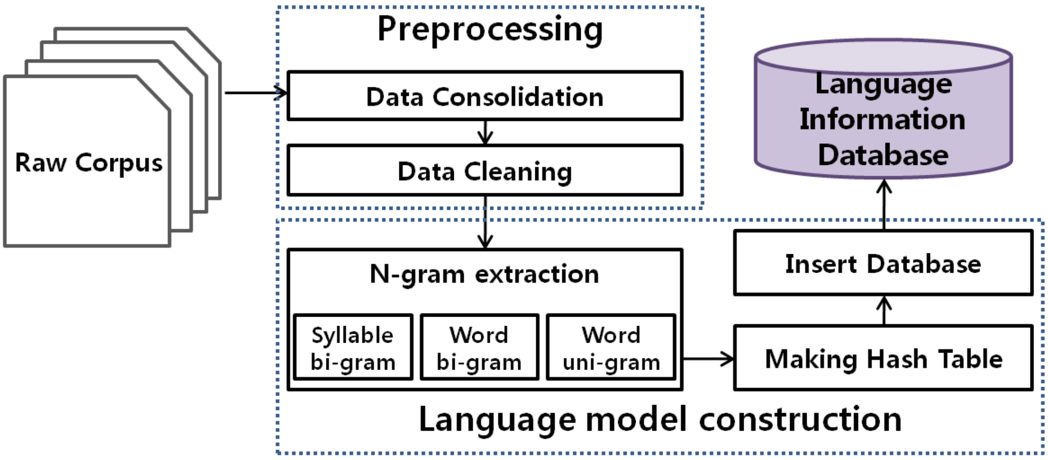

Fig. 4. Construction process of language information database

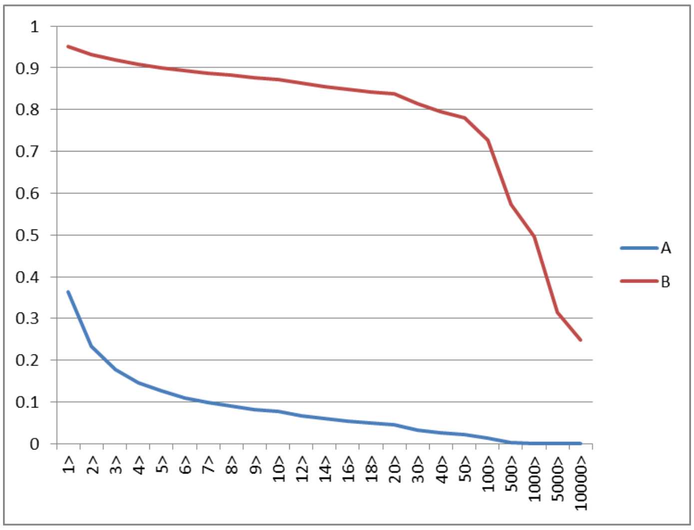

Fig. 5. This figure describes the rate of word frequency which occurred in the whole corpus.The A (blue) line is the frequency rate of unique words and the B (red) line is the overall frequency rate.

4.4 Hangul Entry Modules

The Korean Hangul writing system is a phonogramic system. A syllable might consist of any combination of 19 syllable-initials, 21 medial vowels, and 27 syllable-finals. Accordingly, all the possible notations are available within the set of 11,172 total possible syllables. Thereby, input is made by combining the characters. Generally, this process is called "Hangul Automata". It determines the position occupied by the syllable-initial position, medial vowel, and syllable-final position by the inputed character. The input of the text with limited interface is frequently used in mobile phone environments. Therefore, we constructed a similar inputing mobile interface using the Hangul Automata as shown in **Fig. 6**. For example, if you want to input "고려대학교", the following keys should be selected: 6→3→ 9→2→2→ 8→1→5→ 15→1→ 6→18→ 6→3→3. This sequence is complicated when binary decision is used. Accordingly, we represent flashing the a row and columns of a row which is a set of keys following the sequence shown in **Fig. 7**. First, flashing rows containing the desired syllable will be selected based on the user's intention. If the user selects the row, the next selection is a column followed by the blue arrow in **Fig. 7**, with the same method as the previous step. If the user orders a cancellation, he/she reverts to the previous step through the red arrow shown in **Fig. 7**.

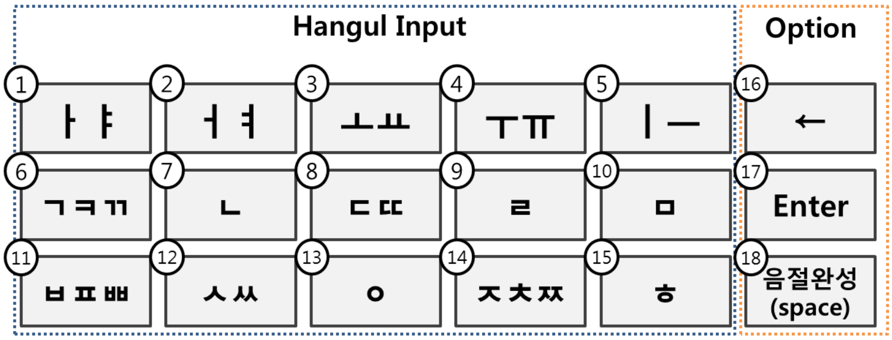

Fig. 6. Key arrangement for the entry of Hangul

5. Demonstration of the Proposed Korean Typewriter

5.1 Demonstration

We conducted a demonstration to verify that communication rate of the proposed method is increased compared to those from non-prediction systems. For the demonstration, we implemented both prediction system in **Fig. 8** (b) and non-prediction system in **Fig. 8** (a). The non-prediction system only uses the key selection module in **Fig. 7** for text input. The prediction system uses three kinds of prediction results from next-syllable prediction using the syllable bi-gram, entry word prediction using the word uni-gram, and next-word prediction using the word bi-gram. The predicted results are aligned from left to right by the probability of their occurrence. The most likely result is positioned at the far left side

The demonstrations were given to 8 healthy subjects using both applications. The subjects first used the non-prediction application, then used the prediction application. Before the demonstration, we identified the signal states and effective detection of left/right wink using the emotive control panel. During the session, applications were recorded in a log file that included entry text and time, which is shown in **Table 1**.

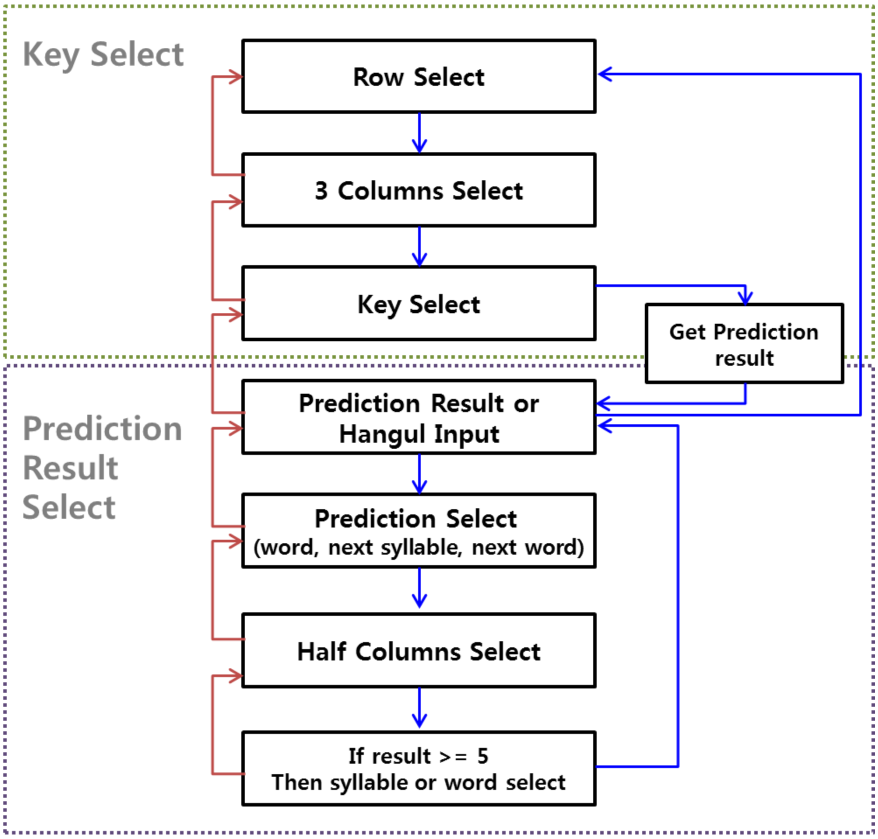

Fig. 7. Text entry process

Table 1. Entry texts

Session	Entry texts (translate english text)
1	고려대학교 (Korea University)
2	반갑습니다. (Nice to meet you.)
3	안녕히가세요. (Good bye.)
4	피곤해요. (I am tired.)
5	감사합니다. (Thank you.)
6	화장실에가고싶어요. (I have to go to the bathroom.)

5.2 Results

We used the Korean character "Jaso" mentioned in section 4.3 to calculate the character per minute rate. As shown in **Fig. 9** and **Fig. 10**, using the prediction system was faster compared to using the non-prediction system. Additionally, selection count was lower than those obtained from non-prediction. For example, input session 1 ("고려대학교") took 130 seconds with non-prediction whereas it only took 79 seconds with prediction system. Selection times also showed a remarkable reduction. The average text entry speed was 3.3 to 4.6 characters per minute (cpm) with non-prediction system and 6 to 12 cpm with prediction system.

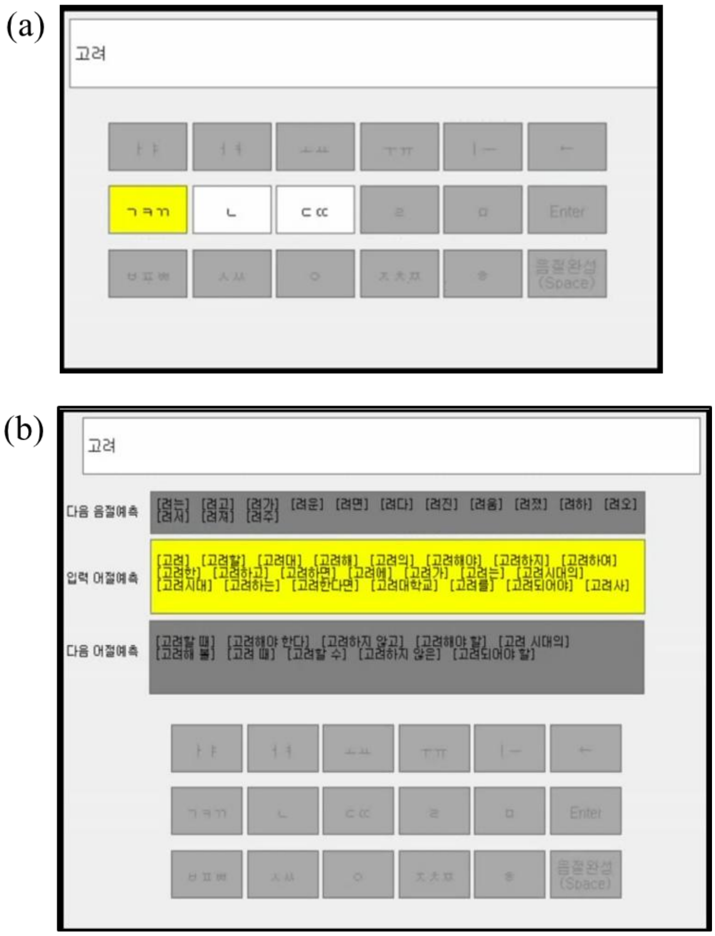

Fig. 8. Text entry system implementation: without prediction on the top (a), with prediction on the bottom (b)

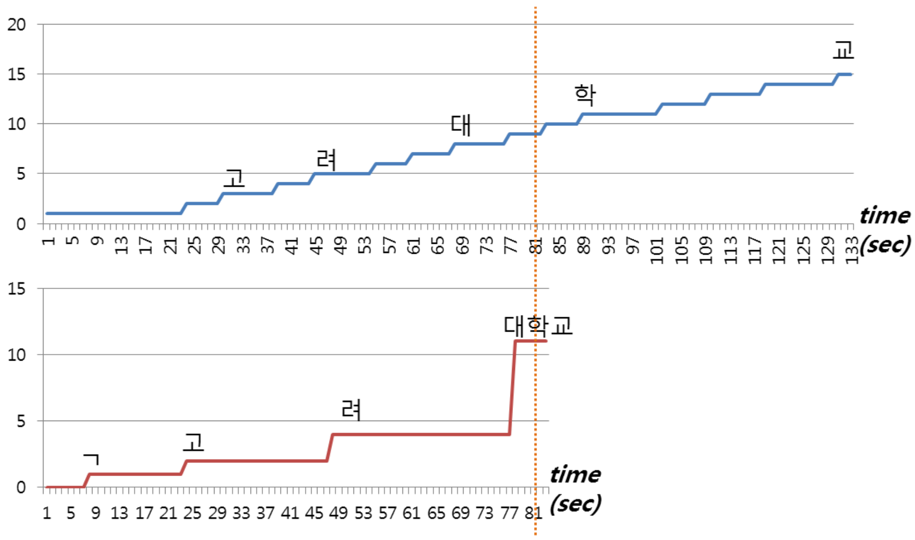

Fig. 9. Average of syllables against time, non-prediction at top, prediction at bottom.

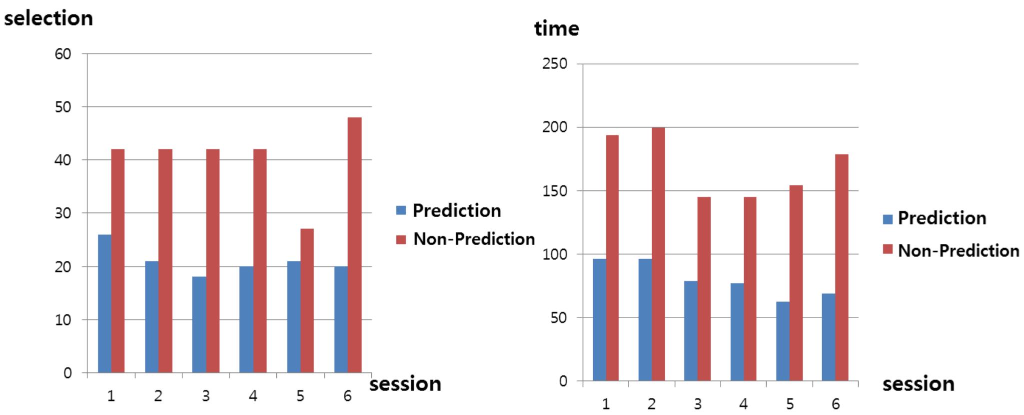

Fig. 10. Prediction vs. non-prediction foreach session

6. Discussion

When inputting a text, if the system can predict or otherwise reflect your intention, the process of writing can be sped up and achieved more naturally. Previous studies have focused on signal processing, which included feature extraction and classification. Prediction using the language model has often been noted but, surprisingly, has been rarely dealt in previous studies. Languages such as English, Spanish, German and Chinese have been adopted in studies with BCI Speller, but BCI which uses Korean typewriter has not been examined in detail if not recognized at all. Written Korean employs syllabic characters and has many syllables which can be combined. Therefore, word completion and syllable prediction are particularly helpful when inputting text.

In this paper, we have used the *Emotiv© systems* epoc. This is oriented for BCI applications, but little research using this device has been conducted. For one aspect, using this device or signals such as left/right wink might not be useful for BCI. Nevertheless, there is significance to this research in BCI field, because it utilizes binary decisions for text entry. Additionally, using these devices means that BCI research is carried out in a laboratory environment.

7. Conclusion

Communication through brain signals is still one of the main challenges in the field of BCI research. Writing even the simple messages that we use in everyday life remains to be a difficult task for people with severe disabilities. For this reason, BCIs are a necessity for improving their quality of life. Through this study, we provided a Korean typewriter using the BCI technique and sped up text entry using the language prediction model. Nevertheless, further experimentation using other languages with the outlined methods would be worthwhile to investigate.

Many studies have been conducted on how to improve the information transfer rate [8][21][22]. It mainly deals with the multi-classification of mental states and feature selection of various spectral or spatial filtering. These researches have contributed to the development of BCIs. However, there were only a few of studies that have dealt with the smart interfaces which boost the communication rate in a fixed bit rate. The smart interfaces include the awareness of the user's context, extraction of the user's intention, and the user experience design. If these techniques are applied to the BCI, it will increase the possibility of achieving the ultimate goal of BCI which will bring the users one step closer to interacting with the machine through 'mind-reading' or 'telepathy'.

References

[1] J.R Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtcheller, T.M. Vaughan, "Brain Computer Interface for Communication and Control," *Clinical Neurophysiology*, vol. 113, pp. 767-791, 2002. Article (CrossRef Link)
[2] M.A. Lebedev, Miguel A.L. Nicolelis, "Brain-Machine Interfaces: Past, Present and Future," *Trends in neurosciences*, vol. 29, no. 9, 2006. Article (CrossRef Link)
[3] H. Hayashi, S. Kato, "Total Manifestations of Amyotrophic Lateral Sclerosis: ALS in the Totally Locked-in State." *Journal of the Neurological Sciences*, vol. 93, no. 1, 1989. Article(CrossRefLink)
[4] K. Andrea, K.-R. Müller, "An Introduction to Brain-Computer Interfacing," *Toward Brain-Computer Interfacing*, MIT Press, 2008.
[5] Emotive Systems, "Emotiv - Brain Computer Interface Technology," http://www.emotiv.com/.
[6] NeuroSky Inc., "NeuroSky - What We Do," http://www.neurosky.com/People/WhatWeDo.aspx.
[7] H. Ramoser, J. Müller-Gerking, G. Pfurtscheller, "Optimal Spatial Filtering of Single Trial EEG During Imagined Hand Movement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4. 1998. Article (CrossRef Link)
[8] G. Dornhege, B. Blankertz, G. Curio, K.-R. Müller, "Increase Information Transfer rates in BCI by CSP Extension to Multi-class," *Neural Information Processing Systems*, 2004.
[9] J. Williamson, R. Murray-Smith, B. Blankertz, M. Krauledat, K.-R. Müller, "Desinging for Uncertain, Asymmetric Control: Interaction Design for Brain-Computer Interfaces," *International Journal of Human-Computer Studies*, vol. 67, 2009. Article (CrossRef Link)
[10] H. Cecotti, "Spelling with Brain-Computer Interfaces Current Trends and Prospects," in *Proc. of 5th French Conference on Computational Neuroscience*, 2010.
[11] J.R. Wolpaw, D.J. McFarland, "Control of a Two-Dimensional Movement Signal by a Noninvasive Brain-Computer Interface in Humans," *PNAS* vol. 101, no. 51, 2004. Article (CrossRef Link)
[12] Daniel J. Szafir, "Non-Invasive BCI through EEG: An Exploration of The Utilization of Electroencephalography to Create Thought-Based Brain-Computer Interfaces," *Boston College, Bachelors Theses*, 2010.
[13] Michael J. Black, John P. Donoghue, "Probabilistically Modeling and Decoding Neural Population Activity in Mortor Cortex," *Toward Brain-Computer Interface*, MIT Press, 2008.
[14] Dawn M. Taylor, "The Importance of Online Error Correction and Feed-Forward Adjustments in Brain-Machine Interfaces for Restoration of Movement," *Toward Brain-Computer Interface*, MIT Press, 2008.
[15] L. Farwell, E. Donchin, "Talking Off the Top of Your Head: Toward a Mental Prosthesis Utilizing Event-related Brain Potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, 1988. Article (CrossRef Link)
[16] H. Cecotti, A. Gräser, "Convolutional Neural Networks for P300 Detection with Application to Brain-Computer Interfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 3, 2011.
[17] N. Liang, L. Bougrain, "Averaging Techniques for Single-Trial Analysis of Oddball Event-related Potentials," in *Proc. of 4th International Brain-Computer Interface workshop*, 2008.
[18] g.tec medical engineering GmbH, http://www.intendix.com/.
[19] C.D. Manning, H. Schütze, "Statistical Inference: N-gram Models over Sparse Data," *Foundations of statistical natural language processing*, MIT Press, 1999.
[20] Wikipedia, "Language Model - Wikipedia, the Free Encyclopedia," http://en.wikipedia.org/wiki/Language_model .

[21] D.J. McFarland, W.A. Sarnacki, J.R. Wolpaw, "Brain–computer interface (BCI) Operation: Optimizing Information Transfer Rates," *Biological Psychology*, vol. 63, no. 3, 2003. Article (CrossRef Link)
[22] B. Obermaier, C. Neuper, C. Guger, G. Pfurscheller, "Information Transfer Rate in a Five-classes Brain-Computer Interface," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 9, no. 3, 2001. Article (CrossRef Link)

Saebyeok Lee is a Ph.D student in the Department of Computer Science Education at Korea University in Seoul, Korea. He received his B.S. in Department of Computer Software from Hanshin University in 2008 and M.S. in Department of Computer Science Education from Korea University in 2011. His research interests are brain-computer interface and affective computing to recognize human affects.

Heuiseok Lim received his B.S., M.S. and Ph.D. degrees in Deparment of Computer Science from Korea University, Seoul, in 1992, 1994, 1997, respectively. He is a professor in the Department of Computer Science Education at Korea University in Korea. He workd at SITEC of Samsung for two years after his Ph. D. He is a member of ACL and many other research groups. His research interests are in natural language processing, computational neuroscience and computer science education to enhance people's computational thinking.