Single-trial Classification of Inner Speech and the No-control State Using Electroencephalography

A. Rezazadeh Sereshkeh[1,2*], R. Trott[1,3], B. Schultz[1,3], T. Chau[1,2]
[1]Bloorview Research Institute, Toronto, Ontario, Canada; [2]University of Toronto, Toronto, Ontario, Canada; [3]Flinders University of South Australia, Adelaide, Australia

* Holland Bloorview Hospital, 150 Kilgour Road, Toronto, Ontario, M4G 1R8, Canada. E-mail: a.rezazadeh@utoronto.ca

Introduction: The main goal of brain-computer interface (BCI) research is to provide communication capabilities for people with severe motor impairments who are unable to communicate conventionally. However, a major drawback for most BCIs is the fact that they make use of non-intuitive mental tasks such as motor imagery, mental arithmetic, or mental reaction to external stimuli to make a selection. These control schemes usually have no correlation with normal communication methods making them difficult to perform by the target population. The goal of the work presented is to investigate the reliability of electroencephalography (EEG) signals in detecting inner speech (also known as covert speech or silent vocalization) against an unconstrained “no-control” state. Previous EEG-based inner speech studies have been limited to the silent vocalization of vowels or syllables rather than complete words [1]. To our knowledge, this study is the first report of using EEG measurements to detect covert articulation of a complete meaningful English word. Also, this is the first EEG study of inner speech performed over multiple sessions for each participant.

Material, Methods and Results: The study was conducted on ten able-bodied participants (five males) with a mean age of 26.8 ± 4.1. EEG signals were recorded using 64 electrodes placed across the scalp in accordance with the International 10-20 system. Each participant undertook two sessions, each consisting of 60 trials, on two different days. Each trial starts with a blank screen with a fixation cross in the center which remains unchanged during the trial. In the “inner speech” trials (half of the trials) participants were asked to answer a perceptual yes/no question by iteratively repeating the answer mentally without any vocalization and motor movement. The question was always the same, “Is this word in uppercase letters? word” with a different word every trial but always in lowercase, hence the answer was “no” for all inner speech trials. There were several reasons for choosing the word “no” for the mental speech task including being useful, intuitive and having a nasal letter, “n”, and vowel of /o/ both of which have been shown to be detectable with reliable accuracy during imagined and articulated speech using EEG [2]. In the remaining trials, called “no-control” trials, participants were told to allow normal thought processes to occur without restriction, rather than being instructed to control their mental activity in a particular way. The trials were presented in pseudorandom order. The duration of each trial, regardless of their class, was 12 s. However, the first 2 s of each trial was removed for analysis since any reactive brain signal to the visual cue at the beginning of trial is not of interest in this study. Acquired EEG data were preprocessed to reduce the effects of artifacts and noise. Then, autoregressive coefficients and wavelet transform features were used to form the feature vector for each trial. Data collected across two sessions were pooled together for classification, and accuracies were evaluated offline using 100 iterations of 10-fold cross-validation. A fast correlation based filter method [3] was used for feature selection and a support vector machine model with linear kernel was used for classifier training. During each fold, only the training data was used for feature selection and classifier training while the accuracy was calculated and reported based on the test data (Table 1).

Table 1. Per-participant classification accuracies. The upper confidence limit is 66.7% for α = 0.01 (confidence level of 99%)

Participant	1	2	3	4	5	6	7	8	9	10	Average
Accuracy	76.1	62.4	68.2	82.5	84.2	78.5	76.8	53.7	51.5	90.3	72.41±13.1

Discussion and Significance: An average classification accuracy (sensitivity, specificity) of 72.41% ± 13.1% (77.4%, 73.1%) between the inner speech and no-control trials was achieved across all participants with 7 participants exceeding 66.7% accuracy (the confidence limit for α = 0.01) and 6 participants exceeding 70%, the threshold necessary for effective BCI communication [4]. The results suggest the potential of a two-choice EEG-BCI based on inner speech tasks. The suggested paradigm can be used intuitively by the target population for communication purposes involving yes/no queries regarding to states of pain, hunger and need of help. Also, since this study used an unconstrained “no-control” state as one of the mental tasks, the presented method is suitable for, and will later be used in, an asynchronous BCI with inner speech as an intuitive activation switch.

References

[1] DaSalla CS, Kambara H, Sato M, Koike Y. Single-trial classification of vowel speech imagery using common spatial patterns. *Neural Networks*, 22(9): 1334-1339, 2009.
[2] Zhao S, Rudzicz F. Classifying phonological categories in imagined and articulated speech.In *ICASSP*, 992-996, 2015.
[3] Yu L, Liu H. Feature selection for high-dimensional data: A fast correlation-based filter solution. In *ICML*, 856-863, 2003.
[4] Kubler A, Mushahwar VK, Hochberg LR, Donoghue JP. BCI meeting 2005-workshop on clinical issues and applications. *IEEE Transactions on neural systems and rehabilitation engineering*. 14(2): 131, 2006.