The Musical BCI: Control by Selective Listening

M. S. Treder[1,2], H. Purwins[1,2], D. Miklody[1], I. Sturm[1,3], B. Blankertz[1,2]
[1]Neurotechnology Group, Berlin Institute of Technology, Berlin, Germany; [2]Bernstein Focus: Neurotechnology, Berlin, Germany; [3]Berlin School of Mind and Brain

Correspondence: M. S. Treder, Neurotechnology Group, Sekr. MAR 4-3, Marchstrasse 23, 10 587 Berlin, Germany.
E-mail: matthias.treder@tu-berlin.de

Abstract. The predominant approach in Brain-Computer Interface (BCI) research to reading out bits of the user's intent is to deliver a stream of isolated visual or auditory stimuli and to detect the modulation of event-related potentials (ERPs) by the user's attention. While such BCI systems provide relatively good control, the stimulation is highly artificial and may become unpleasant. In contrast, music allows for a more intuitive way of listening, since the separate information streams integrate into a coherent and hedonistically appealing entity. Here, we explore an alternative approach to BCI control by employing polyphonic pieces of music as stimuli. As control paradigm, the user shifts attention selectively to one of the musical instruments. Since the musical pieces are composed such that each instrument plays tones deviating from a regular repetitive pattern, modulated ERP responses can be used to infer the user's intent in a similar fashion as for existing ERP-based BCIs.

Keywords: ERP, auditory, P300, music

1. Introduction

It is among the supreme skills of a musical performer of a fugue by Bach that its different instruments can be perceived as independently articulated auditory streams [Bregman, 1990] of their own right so that the listener can deliberately focus her attention to the musical beauty of one instrument or the other. We present work on how selective attention in music listening can be used as a new paradigm in BCI, following earlier BCI systems based on visual or auditory attention [Schreuder et al., 2010; Höhne et al., 2011; Hill and Schölkopf, 2012]. We exploit the fact that polyphonic music integrates several instruments into an aesthetic entity and that listeners are able to follow individual instruments while still being immersed in a holistic listening experience. In Western major-minor tonal music, repetition and variation of patterns is an essential part of the structure that plays with the listeners' expectations. Taking advantage of that we implement a 'musical' oddball paradigm by repeating a characteristic pattern for each instrument (standard stimulus) and varying it infrequently (deviant stimulus) without violating the characteristics of the musical idiom. Our hypothesis is that if listeners focus on one instrument a deviant pattern in this instrument triggers an ERP response. The occurence of this response reveals which instrument has been attended to.

2. Material and Methods

In this study ($N = 11$), a minimalistic version of *Just can't get enough* (Depeche Mode) consisting of synthesized bass, keyboard, and drums was used. Each instrumental part is composed of frequent repetitions of a standard one-measure-long pattern, once in a while interrupted by a deviant bar-long pattern. Then these three instrumental parts were overlaid. The stimuli were presented in audio clips of 40 s length, containing between 3–7 deviants per instrument[1]. Ten randomized versions of such clips were generated. In random order, 21 of these clips form a block. On the whole, three of these blocks were presented to the subject. In order to have the attention of the subject focused on the particular instrument, the subject was asked to count the number of deviants per instrument in each clip and report it afterwards using the computer keyboard. Concurrently, we used a 64-channel active-electrode setup with a standard montage to record the brain activity associated with the musical stimuli.

Since the deviants in each instrument had different physical characteristics and, hence, the evoked response differed across instruments, we trained a separate binary classifier for each instrument using regularized linear discriminant analysis. Each classifier was trained to discriminate between attended deviants and unattended deviants of that instrument. We then modelled the classifier outputs as Gaussian probability distributions using maximum likelihood estimates on the training data. In the test phase, we obtained posterior probabilities for each deviant indicating the

[1]Sound example: soundcloud.com/hpurwins/depmod1

DOI:10.3217/978-3-85125-260-6-70

probability that the instrument was attended. For each clip, posterior probabilities were averaged within each instrument and the instrument with the highest mean posterior probability was selected as the attended instrument.

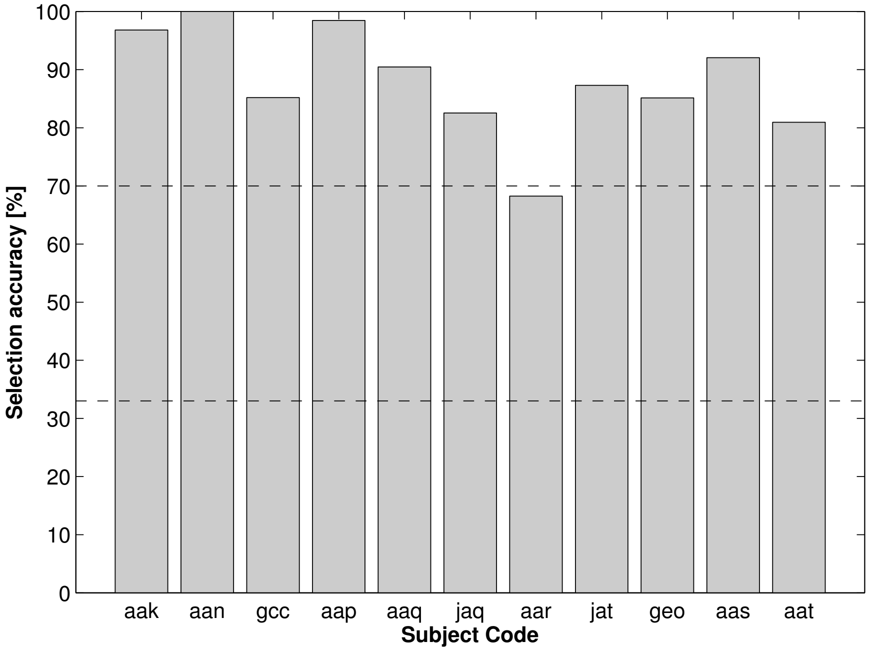

Figure 1: *Instrument classification accuracy for each subject. The lower dashed line indicates chance level, the upper dashed line indicates the 70 % benchmark.*

3. Results

In an offline analysis, we performed a leave-one-clip-out cross-validation, wherein classifiers were trained on all clips but one and then validated on the held out clip. Figure 1 depicts the accuracy of detecting the attended instrument in a clip. Ten out of eleven subjects have a performance above 80 %, the lowest performance is 68 %.

4. Discussion

In this study we addressed the problem of transferring the typical BCI setting to a more enjoyable setting using musical pieces. Our approach is built on the everyday listening ability to follow an individual instrument in polyphonic music. Although the number of possible concurrent streams is limited and our stimulus sequence is not the fastest way to present an oddball paradigm, our results show that it is possible to design a BCI that is linked to an important source of joy for individuals, namely music listening. In future work, we intend to improve the information rate of this paradigm, compare various musical stimuli, and investigate its performance in an online setting. Furthermore, this approach opens an avenue for investigating selective auditory attention to music and how it relates to stream segregation [Bregman, 1990], supported by differentiating the streams with respect to timbre, pitch range, and rhythmical structure. Also this approach could be used to investigate which signatures in a complex musical score involuntarily call the attention of the listener.

References

Bregman, A. S. (1990). *Auditory Scene Analysis*. MIT Press, Cambridge, MA.

Hill, N. J. and Schölkopf, B. (2012). An online brain-computer interface based on shifting attention to concurrent streams of auditory stimuli. *J Neural Eng*, 9 (2).

Höhne, J., Schreuder, M., Blankertz, B., and Tangermann, M. (2011). A novel 9-class auditory erp paradigm driving a predictive text entry system. *Front Neurosci*, 5:99.

Schreuder, M., Blankertz, B., and Tangermann, M. (2010). A new auditory multi-class brain-computer interface paradigm: Spatial hearing as an informative cue. *PLoS ONE*, 5 (4).