Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display, Sydney, Australia, July 6-9, 2004

AUDITORY FEEDBACK OF HUMAN EEG FOR DIRECT BRAIN-COMPUTER COMMUNICATION

Thilo Hinterberger, Gerold Baier, Jürgen Mellinger, Niels Birbaumer*

Institute of Medical Psychology and Behavioral Neurobiology
Gartenstr. 29, University of Tübingen
72074 Tübingen, Germany
* Facultad de Ciencias, Universidad Autónoma del Estado de Morelos, Cuernavaca, Mexico
thilo.hinterberger@uni-tuebingen.de

ABSTRACT

The Thought-Translation-Device (TTD) is a Brain-Computer-Interface (BCI) that enables completely paralyzed patients to communicate by the use of their brain signals only. Self-regulation of brain signals (e.g. the slow cortical potentials) is achieved by a feedback training. Visual impairment of these patients asks for an auditory feedback mode. The TTD can be entirely operated by combined listening and mental activity. It provides auditory feedback of brain signals which can operate a verbal spelling interface. The extension POSER allows for sonified orchestral real-time feedback of multiple EEG parameters for the training of self-regulation. The properties of the system are reported and the results of some studies and experiments with auditory feedback are presented.

1. INTRODUCTION

The Thought Translation Device (TTD, Fig. 1) is a brain computer interface (BCI) developed to enable severely paralyzed patients, e.g. people diagnosed with amyotrophic lateral sclerosis, to communicate using slow potential changes of the electroencephalogram (EEG) [1]. Slow cortical potentials (SCPs) are very low frequency potential changes (below 1 Hz) lasting from 500 ms to several seconds. Negative potential shifts (negativity) represent an increased excitability of neurons (e.g., readiness) while a positive shift (positivity) can be measured during the consumption of resources or during rest. The participants are first trained to produce positive or negative SCP shifts using a visual feedback task presented on a computer screen. This ability can be used to select letters and write messages [2]. For patients with impaired visual abilities, the TTD can be entirely operated by brain signals as a voluntary response to auditory instructions and feedback as presented in the following.

2. BRAIN-COMPUTER COMMUNICATION

2.1. The Thought Translation Device (TTD)

The EEG is acquired with a Psylab EEG8 amplifier which is connected to a PC via an A/D-converter. The recording site for the feedback signal is Cz (International 10-20-system) with the references at both mastoids. The EEG signal is sampled at 256 Hz and digitized with a 16 bit A/D converter (PCIM-DAS1602/16 from Measurement Computing™) in an amplitude range of +/-1 mV. The filters of the amplifier are set to 0.01 Hz (i.e. a time constant of 16 s) as low frequency cut-off and 40 Hz as high frequency cut-off. EEG usually is recorded from 3 to 7 Ag/AgCl-electrodes placed at Cz, C3, C4, Fz and Pz referenced to the mastoids. Additionally, one bipolar channel is used to record the vertical electrooculogram (vEOG) for on-line and off-line artifact correction. Feedback is updated 16 times per second. SCPs are calculated by applying a 500 ms moving average to the EEG signal [3][4].

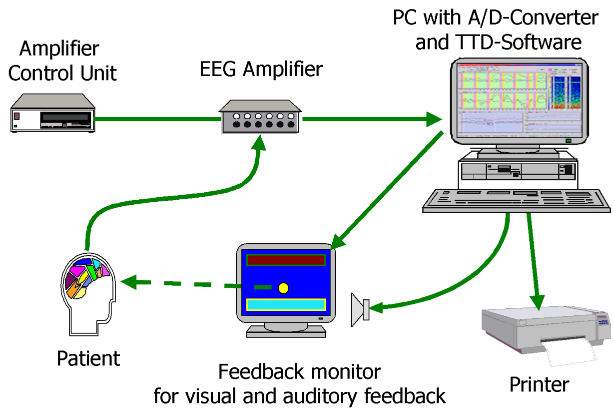

Figure 1. *Experimental Setup of the TTD that serves as a multimedia feedback and communication system. The EEG is amplified, and acquired by the PC with an A/D converter board. The TTD software performs on-line processing, storage, display, and analysis of the ongoing EEG. It provides feedback on a screen for learning self regulation of various EEG components (e.g. SCPs) in a paced paradigm and enables a well-trained person to interface with a variety of tasks, e.g. a visual or auditory speller for writing messages or a web-browser for navigating through the world wide web using brain potentials only. All feedback information can be given auditorily to enable visually impaired patients to communicate with brain signals only.*

The participants or patients view the course of their SCPs by the vertical movement of a feedback cursor on the screen that should be moved towards a prescribed direction. Each attempt (trial) lasts for 4 to 14 s. The first 2 s prepare the person for the feedback task by showing the targets while from second 2 until 0.5 s before the end of the trial feedback is presented. Finally, a smiley face combined with a sound of chimes reinforces a correct performed trial. When a rate of 75% correct responses is obtained, they can be trained to select letters and write messages using their self-regulative abilities. In this speller task, letters or blocks of letters are presented to the patient and the feedback cursor has to be moved towards the desired letters for selection (Figure 2a). The effectiveness of the device was demonstrated in several locked-in patients [1]. A limitation was soon evident with this initial TTD. In patients in an advanced stage of the disease, focusing gaze to sufficiently process the visual feedback or read the letters in the verbal communication paradigm is no longer possible. In this case, a non-visual feedback modality such as auditory or tactile feedback has to be implemented.

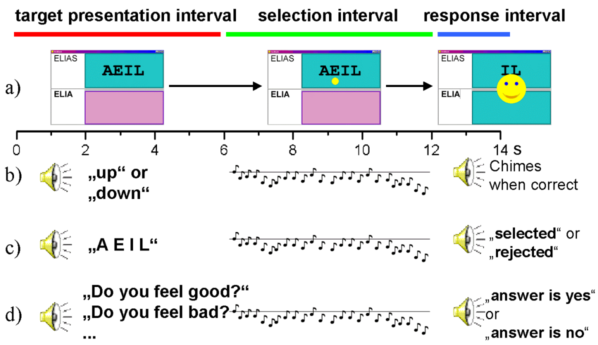

Figure 2. *To train a completely "locked-in" patient with the copy-spelling mode a predefined word has to be spelled. The communication process requires three intervals in a trial: 1) 2 to 6 seconds for presentation of the letter set which is spoken by a computer voice and/or displayed in the target rectangle on the screen. 2) 2-6 seconds of feedback presentation. Self regulation of SCP-amplitudes is used to select or reject the letter set. 3) a response interval that informs the user about the outcome of a trial. a) shows the visual stimuli for spelling. b) shows the stimuli for the auditory training of self-regulation of auditory displayed SCPs. c) depicts the stimuli in an auditory spelling system for brain-computer communication. In each trial a single letter or a set of letters can be selected or rejected by a binary brain response which corresponds to a cortical negative or positive potential shift. A voice informs the user at the end of a trial by saying "selected" or "rejected". In the auditory mode, a patient can spell words by responding to the suggested letter sets trial by trial. d) The question-answering paradigm allows for receiving yes-no answers even in less skilled patients.*

2.2. Auditory Feedback for Brain-Computer Communication

The following describes how all visual feedback information was transferred to the auditory channel to operate the TTD completely auditorily. For permanent auditory feedback, the SCP amplitude shifts are coded in the pitch of MIDI sounds which are presented with 16 touches per second. High pitched tones indicate cortical negativity, low pitched tones cortical positivity. The task was presented by a computer-generated voice (.wav-file) that said "high" or "low" to indicate that the participant had to increase or decrease the pitch of the feedback sound. If the overall result was correct a harmonious jingle was presented at the end of the feedback period as positive reinforcement (Figure 2b). Figure 2c) presents the sequence of stimuli in the letter selection mode. The letter sequence to be selected is delivered by a pre-recorded, computer-generated voice at the beginning of a initial preparation interval. After feedback, the selection or rejection response is confirmed [AuditorySpeller.mp3].

2.3. Comparison of Auditory versus Visual Feedback

An experiment is presented which was carried out to investigate the usability of auditory feedback for controlling a brain-computer interface. Three groups of healthy persons (N=3*18) were trained over three sessions to learn SCP self-regulation by visual, auditory and both visual and auditory feedback. The task to produce cortical positivity or negativity was randomly assigned. Each session comprised 10 runs with 50 trials each. Each trial of 6 seconds duration consists of a two seconds preparation interval, a 3.5 seconds feedback interval followed by 0.5 seconds for presentation of the result and the reinforcing smiley associated with a jingle sound. Similarly to Fig. 2a) and b), the task was presented either by an emphasized rectangle into which the feedback cursor should be moved to or by a voice telling whether the feedback sound (the pitch reflected the SCP-amplitude) should be high or low ("up" or "down").

The performance of the third run was analyzed for each subject for each feedback modality. All groups showed significant learning for each modality for the majority of the subjects. More than 70 % correct responses in the third session were achieved by 6 (out of 18) subjects with visual feedback, by 5 subjects with auditory and only by two with combined feedback. The average correct response rate in the third session was 67 % in the visual condition, 59 % in the auditory, and 57 % in the combined condition. Overall, visual feedback is superior to the auditory and combined feedback modality [5].

2.4. Answering Questions with the auditory TTD

The auditory (letter selection) communication paradigm was tested with a completely paralyzed patient without any other means of communication. Despite the fact that his performance for SCP self-regulation was at average only about 60%, he could spell some intended words using a letter set of eight letters. However, to achieve a reliable answer from the less-skilled patients a question-answering paradigm was set up which presented questions instead of letters (Figure 2d). Repetition of the same question allows for detection of a statistically significant brain response and thus a reliable answer. Presentation of almost 500 questions to this patient showed that even with unreliable brain control significant answers can be obtained after averaging the responses of all identical questions.

3. PARAMETRIC ORCHESTRAL SONIFICATION OF EEG RHYTHMS (POSER)

3.1. Methodology

SCPs are not the only feature of the EEG that can be used to control a BCI. Different kinds of the rhythmic activity of the EEG such as the μ-rhythm [6]-[8] can also be used. A comparative study showed that human performance depending on multivariate information (simultaneous input from various sources) was significantly better with an auditory display as compared to a visual display or a mixed visual and auditory display [9]. This encouraged us to enhance the simple single parameter auditory feedback to a multi-parametric sonification approach [9]-[13]. A number of modules were added to the

TTD for presentation of EEG-rhythms in real time which form the POSER system (Parametric Orchestral Sonification of EEG in Real-time) [14]. Previous studies with artificial irregular oscillations (generated by chaotic differential equations) have demonstrated that rhythmic properties can optimally be studied in aural representation if smooth waveforms are transformed to spike-like (discrete) signals first [15]. For POSER, this idea is applied to the polyrhythmic EEG signal. As the human auditory system is able to distinguish sounds in a complex superposition, different rhythms are projected to different instruments and pitches and can be played simultaneously. The POSER approach satisfies the following criteria for a realistic sonification in real-time:

1. rhythmic activity below 15 Hz should be audible as such
2. intensity of a rhythm should be audible
3. frequency of a rhythm should additionally modulate the pitch of the rhythmic presentation to allow for the detection of possible harmonies (e.g., a rhythm of 8 Hz may be played one octave higher than a rhythm of 4 Hz).
4. where possible, frequency relationships should be maintained

These criteria have been implemented according to Fig. 3. as follows.

Data processing is performed at a rate of 128 per second (i.e. half the sampling frequency). This is also the pacing of the sonification providing a sufficient time resolution for the rhythms. A 1.6 GHz Pentium IV machine running under MS Windows XP showed excellent real-time behavior. The response time is mainly determined by the phase shift of the FIR filter (here 100 to 250 ms).

Preprocessing: After calibration of the digitized EEG-signal into µV a spatial filter offers the user to arrange the EEG-channels by a linear combination of incoming channels. In the standard setting three channels of EEG are used: Cz-mastoids, C3-mastoids, C4-mastoids. A rough correction of eye movement artifacts is realized by subtraction of a fixed portion of the vertical Electrooculogram (vEOG) (the factor is about 0.12).

Band-pass filtering: The EEG is then split into various frequency bands by a band-pass filter using an FIR (finite impulse response) filter algorithm with either 127 or 63 coefficients. A moving average of a window of 127 coefficients results in the slow-wave activity (SCP) from 0 to 1 Hz. The next five frequency bands comprise 1 to 4 Hz (delta band), 4 to 7 Hz (theta band), 7 to 12 Hz (alpha band), 12 to 30 Hz (beta band), and 30 to 40 Hz (gamma band). While the first four bands up to 12 Hz use 127 filter coefficients, the beta and gamma bands are filtered with 63 coefficients to achieve faster response times.

Extrema detection: Characteristic rhythms of the EEG in the frequency range below 12 Hz are sonified by triggering the touches of a note at the maxima of a wave. For this purpose, maxima are detected in the filtered signals of the three frequency bands from 1 to 12 Hz. As a maxima can only be detected after it occurred (one processing step=1/128 s afterwards) an additional latency of about 8 ms arises. In addition, the potential differences between subsequent extrema (maxima minus previous minima or minima minus previous maxima) are calculated. The three output signals of this filter carry the potential differences together with the times where the extrema were detected, otherwise they are zero.

Extrema frequency: The inverse time difference between consecutive maxima of a band-pass filtered signal serves to estimate the "instantaneous" frequency of a signal.

Down sampling: To lower the update rate of the sounds to half the sampling rate, all output signals are down-sampled.

Calculation of band power: As 15 Hz is about the maximum frequency at which two consecutive events are resolved as distinct by the human ear, the touch triggering of sounds is not appropriate for the beta and gamma band. In these frequency bands we chose the "instantaneous" band power as a representative measure of activity. The corresponding outputs of the FIR-filter are therefore squared and low-pass filtered to obtain the progressional band power.

The parameters extracted from the EEG are assigned to voices of a MIDI device. A voice, in turn, is assigned to a MIDI channel. The MIDI device of a common PC soundcard provides 16 MIDI channels. Each channel is defined by three parameters for an instrument, its volume and balance. Thereby, each instrument is spatialized by using intensity panning on the stereo sound signal. Other specific parameters of a voice are the note (pitch) and its velocity. An EEG parameter can therefore modulate either pitch or velocity, or both, pitch and velocity. The EEG parameters carry signal amplitudes with a certain amplitude range. Appropriate modulation of note and/or velocity additionally requires the definition of a baseline note (resp. velocity) and a scaling factor. A threshold parameter suppresses continuous touching [POSER_Example.mp3].

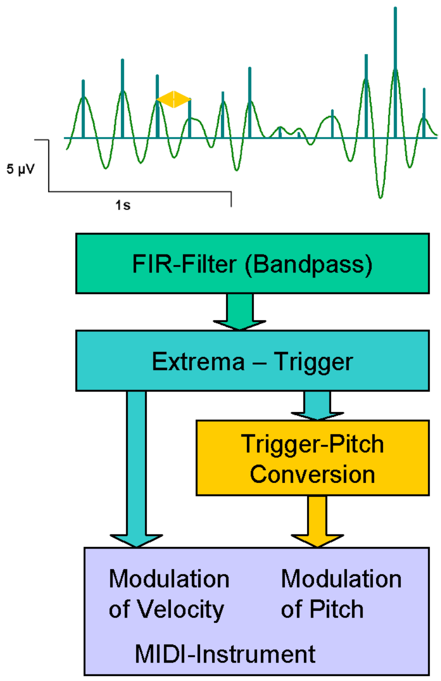

Figure 3. *Sonification of rhythmic brain activity: A band-pass filter extracts the activity in a frequency band. The curve gives an example of the extracted theta band. The wave maxima serve as trigger for the touch of a MIDI instrument. The amplitude between maxima and minima determine the velocity of a touch. The pitch is calculated by the time between two wave maxima.*

3.2. Results of Listening to Multi-parametrically Sonified EEG

Ten healthy participants (aged from 20 to 44 years; 6 female, 4 male) received full orchestral feedback of their brain activity from one channel of EEG (Cz versus C5). A ProComp+ portable amplifier was used with an EEG-sensor providing a frequency range from 0.01 Hz to 40 Hz. The following setting was used for sonification of the brain rhythms:

- SCPs were played by instrument no. 80 (ocarina) in the General MIDI instrument set.
- The delta band was represented by low pitches of instrument no. 47 (harp) from both speakers.
- Theta and alpha were played by instrument no. 12 (vibraphone) with the same base note leading to a pitch proportionally higher according to its speed. Theta was assigned more to the left channel, alpha more to the right.
- Easy to distinguish high-pitched instruments no. 98 and no. 101 were used for beta and gamma activity, respectively. Beta was assigned to the left and gamma to the right speaker position.

Initially, participants were instructed concerning to the meaning of the different sounds. For this purpose, each instrument was played separately. After that, all instruments were played together and their volumes were balanced according to the participants' individual spectral power distribution. All subjects appreciated the characteristic timbres of the instruments and found the stereo presentation helpful for distinction, e.g., between theta and alpha rhythms. None of them reported problems with sound complexity.

After a short period of adaptation the participants were introduced to the attention task. They were instructed to focus their attention alternately on two different sounds. The randomized sequence of the tasks was defined by the computer in a balanced order and symbolized on a screen by two vertically arranged rectangles. A red colored upper rectangle asked participants to focus attention on the rhythmic vibraphone sounds (alpha, theta, and gamma). A red colored lower rectangle asked them to focus attention on the smooth sounds (SCP, delta, and beta). Each task symbol was presented for 8 seconds. After a two-second resting interval, the next trial began. 50 trials comprised one block of trials. Four blocks separated by short breaks were carried out in one session.

We evaluated task-dependent variations of the parameters used for control of the POSER device. A two-tailed t-test revealed significant differentiation between the tasks at least in one parameter for nine of ten participants.

The highest variations were induced by one subject in the delta-band ($t(200)=9.21$, $p<0.01$). In a classification of these delta amplitudes 80 % of all trials could be classified in accordance with the task. This regulative ability could be replicated in a second session with the same participant with 85 % correct responses and a t-value of $t(150)=12.2$. Three participants revealed regulatory effects in more than one of the rhythmically presented frequency bands. Significant changes of the amplitudes in the alpha-band could be observed in four of the ten participants. The theta-band was modulated significantly by four participants also. Each, beta and gamma power revealed significant differences in three of the subjects.

Significant regulation of the frequency of a rhythm was observed in two participants additionally to at least one significantly modified amplitude.

No significant differentiation of SCP amplitudes was seen in this pilot study. More details are reported in [14].

4. CONCLUSIONS

With the TTD, a Brain-Computer-Communication system is available that can be operated without visual abilities by presenting all relevant information across the auditory channel to the user. Physiological regulation of SCPs can be learned with auditory and combined auditory and visual feedback although the performance is significantly worse than with visual feedback alone. We demonstrate that auditory feedback of EEG parameters allows for acquisition of self regulation skills and that these skills can be utilized to communicate with brain activity. Even patients with unreliable EEG control can give reliable answers with the presented questioning paradigm. Further, the POSER approach for simultaneous feedback of multiple EEG parameters offers a fast way for finding controllable EEG parameters.

5. ACKNOWLEDGEMENTS

This work was supported by the Deutsche Forschungs-gesellschaft (DFG) within the SFB 550 (B5), the National Institutes of Health (NIH), and the University of Morelos by CONACyT Mexico (project no. 40885-F).

6. REFERENCES

[1] N. Birbaumer, H. Flor, N. Ghanayim, T. Hinterberger, I. Iverson, E. Taub, B. Kotchoubey, A. Kübler, and J. Perelmouter, "A Spelling Device for the Paralyzed," *Nature*, vol. 398, pp. 297-298, 1999.

[2] T. Hinterberger, J. Kaiser, A. Kübler, N. Neumann and N. Birbaumer, "The Thought Translation Device and its Applications to the Completely Paralyzed," in H. Diebner, T. Druckrey & P. Weibel (eds.): *Sciences of the Interfaces*, Genista-Verlag Tübingen, pp. 232-240, 2001.

[3] T. Hinterberger, J. Mellinger, and N. Birbaumer (2003). The Thought Translation Device: Structure of a multimodal brain-computer communication system. *Proceedings of the 1st International IEEE EMBS Conference on Neural Engineering*, Capri Island, Italy, March 20-22, 2003, pp. 603-606.

[4] T. Hinterberger, „Entwicklung und Optimierung eines Gehirn-Computer-Interfaces mit langsamen Hirnpotentialen*", Dissertation in der Fakultät für Physik an der Eberhard-Karls-Universität Tübingen*, Schwäbische Verlagsgesellschaft, 1999.

[5] T. Hinterberger , N. Neumann, M. Pham, A. Kübler, A. Grether, N. Hofmayer, B. Wilhelm, H. Flor, and N. Birbaumer, "A multimodal brain-based feedback and communication system," *Experimental Brain Research*, vol. 154, pp. 521-526, 2004.

[6] M.B. Sterman, "Sensorimotor EEG Operant Conditioning: Experimental and Clinical Effects", *Pavlovian Journal of Biological Sciences,* vol. 12, pp. 63-92, 1977.

[7] J. R Wolpaw and D.J. McFarland, "Multichannel EEG-based Brain-Computer Communication", *Electroencephalography and Clinical Neurophysiology*, vol. 90, pp. 444-449, 1994.

[8] G. Pfurtscheller, D. Flotzinger, M. Pregenzer, J.R. Wolpaw, and D. McFarland, „EEG-Based Brain Computer Interface", *Medical Progress through Technology,* vol. 21, pp. 111-121, 1996.

[9] W.T. Fitch, G. Kramer, "Sonifying the Body Electric: Superiority of an Auditory Display over a Visual Display in a Complex, Multivariate System", in [10], p. 307.

[10] G. Kramer (ed.), *Auditory Display – Sonification, Audification, and Auditory Interfaces*, Addison-Wesley, Reading 1994.

[11] G. Kramer, "Some Organizing Principles for Representing Data with Sound", in [10], p. 185.

[12] G. Mayer-Kress, "Sonification of multiple electrode human scalp electroencephalogram", *Poster presentation demo at*

ICAD '94. http://www.ccsr.iuic.edu/People/gmk/Projects/EEGSound/, 1994.

[13] T. Hermann, P. Meinicke, H. Bekel, H. Ritter, H.M. Muller, S. Weiss, "Sonifications for EEG data analysis", *Proc. Int. Conf. on Auditory Display*, publ. Int. Community for Auditory Display, pp. 37-41, Kyoto, Japan, July 2002.

[14] T. Hinterberger and G. Baier, “POSER: Parametric Orchestral Sonification of EEG in Real-Time for the Self-Regulation of Brain States”, *Proceedings of the International Workshop on Interactive Sonification*, Bielefeld, Germany, January 2004.

[15] G. Baier and S. Sahle, „Listening to Chaos: the Aural Representation of Irregular Rhythms”. In: H.H. Diebner, T. Druckery and P. Weibel (eds.), *Sciences of the Interface*, Genista Verlag, Tübingen, pp. 255-260, 2001.