Masterarbeit

Morse Code Speller Based on an Auditory Brain-Computer Interface

Isabella Legat

Institut für Semantische Datenanalyse
Technische Universität Graz

Begutachter: Univ.-Doz. Dipl.-Ing. Dr. techn. Gernot Müller-Putz
Betreuer: Univ.-Doz. Dipl.-Ing. Dr. techn. Gernot Müller-Putz

Graz, Mai 2010

Kurzfassung

Das Morse Code Schreibgerät basierend auf einem auditorischen Brain-Computer Interface beschreibt eine Hirn-Computer Schnittstelle mit akustischen Stimuli und P300 Detektion, die Patienten eine Möglichen zur Kommunikation geben soll. Mit dieser “mentalen Schreibmaschine” kann mit Hilfe von Gedanken der Morsecode (Punkt, Strich) und damit Buchstaben und Wörter geschrieben werden.

Zwei Tonreihen, bestehend aus kurzen Tönen mit hoher und tiefer Frequenz, werden abgespielt, einige wenige davon sind abweichend in der Tonhöhe. Damit wird ein Oddball Paradigma erzeugt. Durch das Lenken der Aufmerksamkeit auf die hohen bzw. tiefen abweichenden Töne ensteht ein P300 Potential. Die Hirnsignale werden als Elektroenzephalogramm aufgezeichnet, verarbeitet und anschließend klassifiziert. Aus dem Ergebnis werden die Klassen “Punkt”, “Strich”, “Ende des Morse Wortes” und “Leerzeichen zwischen zwei Buchstaben” berechnet und visuell ausgegeben.

10 Personen nahmen an einem Test, bei dem 80 vorgegebene Morse Zeichen geschrieben wurden, teil. Verschiedene Vorverarbeitungsschritte und Klassifikationsmethoden wurden mit Testdaten verglichen. Die getesteten Klassifikationsmethoden waren: (1) Fisher Diskriminanz Analyse (2) schrittweise lineare Diskriminanz Analyse (3) eine Fisher Diskriminanz Analyse für jedes Feature. Die Fisher Diskriminanz Analyse ist mit einer vorgeschalteten Diskriminanzanalyse (discriminant power) zur Feature Auswahl kombiniert.
Die besten Ergebnisse wurden mit der schrittweisen linearen Diskriminanz Analyse erziehlt - 9 der 10 Teilnehmer erreichten eine Klassifikationsgenauigkeit von 70% oder mehr mit 10-facher Kreuzvalidierung.

Stichwörter: Brain-Computer Interface, BCI, akustisch, P300, Morse Code

Abstract

The Morse code speller based on an auditory brain-computer interface is a mental typewriter that is based on acoustic stimuli and P300 detection. Specific "thoughts" write the Morse code (dot, dash) and thus spell letters and words.

Two tone streams composed of short beep tones of high and low frequencies are created (auditory stream phenomenon); some of the stimuli are deviant in pitch (oddball paradigm). By focusing the attention on either the low or the high deviant tones, a P300 potential is elicited. The brain signals are recorded as electroencephalogram, processed and classified. The target classes "dot", "dash", "end of Morse word", and "space between two letters" are predicted and communicated on a visual feedback screen.

Ten participants took part in a test, where 80 given Morse symbols were spelled. Different signal processing steps, feature extraction and selection methods, and classification methods were tested on the data recorded during the tests. The compared classification methods were: (1) Fisher's discriminant (2) stepwise linear discriminant analysis (SWLDA) (3) one Fisher's discriminant for every feature. The Fisher's discriminant is combined with a preceding discriminant power algorithm.
The SWLDA classifier performed best; 9 of the 10 participants achieved a selection accuracy of 70% or more with a 10-fold cross validation.

Keywords: brain-computer interface, BCI, auditory, P300, Morse code

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz,am ....

(Unterschrift)

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

....

date (signature)

Contents

Chapter 1

Introduction

1.1 Brain-Computer Interfaces

A brain-computer interface (BCI) permits the connection of the human brain with a computer without using muscular activity. All started 1973 with Jacques Vidal thinking about direct brain-computer communication [43]. Since then it is possible for disabled people to steer wheelchairs with their thoughts, the grasping function of spinal cord ingury patients can be restored, and people with locked-in syndrome, who are paralyzed and can not communicate anymore, can spell with thoughts alone[7].

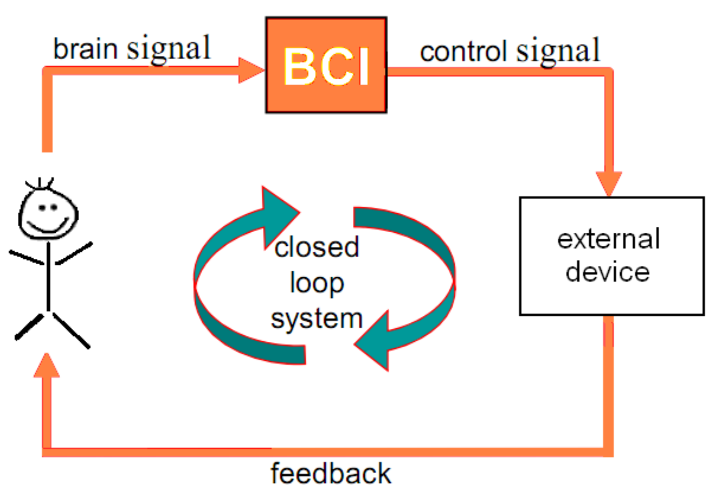


Figure 1.1: Concept of a brain-computer interface adapted from [32].

A BCI establishes a direct communication channel between the brain and the external world. It is a non muscular way to control an external device or effector.

A BCI transforms "thoughts" into control signals by analyzing the brain activity [46] [Fig. 1.1]. The components that come along with a BCI are the signal recording, the type of the brain signals, the signal processing, the mode of operation, the experimental strategy and the feedback.

1.1.1 Signal Recording

The recording of the brain activity (brain signals) is divided into invasive and non invasive methods. Invasive brain recordings include implanted electrodes [11, 14] and (sub) dural electrodes [24] [Fig. 1.2].

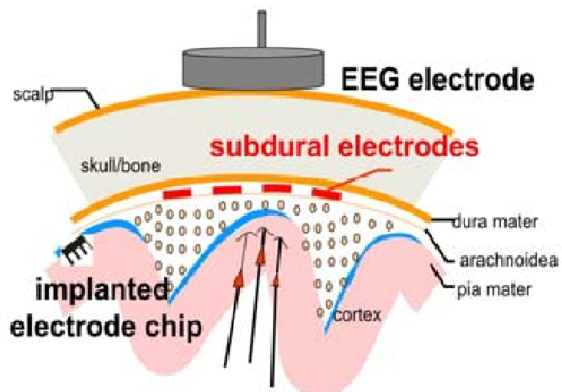


Figure 1.2: Brain signal recording [32].

Implanted electrodes (or intracortical electrodes) are inserted into the grey matter of the brain. It requires invasive surgery to implant the electrodes. The electrodes are placed at the source of the brain signal; they record the spike activity of multiple neurons (multiunit activity). Therefore, implanted electrodes provide the signals with the best signal-to-noise ratio. However, the electrodes lead to a build-up of scar tissue and eventually the brain signal is blocked. [14]

Subdural electrodes are placed on the surface of the brain, beneath the skull on the dura. They are also invasive but implanted farther away from the source than implanted electrodes and thus deliver slightly less accurate signals. This signals are called electrocorticogram (ECoG) signals. [25]

Non invasive brain signal recordings include:

- Electroencephalography (EEG)
  Electroencephalography is the recording of the electrical brain activity; the electrical brain activity is the total neural activity or the sum of the apical post-synaptic potentials. Electrodes are attached to the scalp and measure the electrical potential [Fig. 1.2]; the signals are forwarded to a computer and recorded. [46]

- Magnetoencephalography (MEG)
  Magnetoencephalography measures the magnetic fields and electrical activity that are generated due ionic currents of the neurons. Since these magnetic fields are very small (fT) compared to the magnetic field of the earth, extremely sensitive devices (called SQUIDS) are needed for the recording. [13]

- Functional magnetic resonance imaging (fMRI)
  Functional magnetic resonance imaging plots an image of the brain activity by measuring the contents of oxygen in the blood (BOLD signal) due to neural activity. [44]

- Near infrared spectroscopy (NIRS)
  Near infrared spectroscopy is based on the interaction of near infrared light and tissue. The brain activity is computed with regard to the absorption of the light, which is related to the metabolic changes caused by the neural activity (haemodynamic response). [38]

1.1.2 Brain Signals

The brain activity steers the BCI. For EEG based BCIs, changes in the brain activity that are used to control the BCI include brain oscillations and event-related potentials.

Oscillations are time-locked but not phase-locked responses and can be the result of a change in the functional connectivity within neuronal networks. [33]

- Event-related desynchronization (ERD)
  Event-related desynchronization is a decrease of power in specific frequency

band due to a decrease in the synchrony of the neurons due to activated areas of neurons [Fig. 1.3].

- Event-related synchronization (ERS)
  Event-related synchronization is an increase of power in specific frequency bands due to an increase in the synchrony within neural networks [Fig. 1.3].

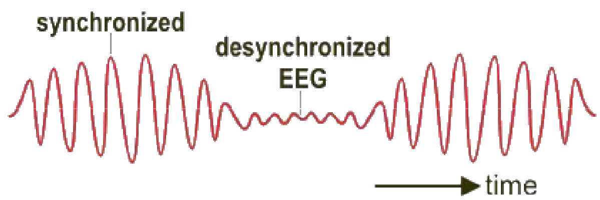


Figure 1.3: Event-related (de) synchronization of brain oscillations [28].

Event-related potentials (ERPs) are time and phase locked brain responses to an internal or external stimulus, i.e. the evoked response has a more or less fixed time delay to the stimulus. The ERP is embedded in the EEG waveform and is small in relation to the EEG (small signal-to-noise ratio). Thus the ERP needs to be extracted from the EEG signal. One popular method is to average multiple ERP responses. [5]

- Slow cortical potentials (SCPs)
  Slow cortical potentials are slow changes in the amplitude (100–200 µV) with a duration of a few seconds [2].

- Evoked potentials (EPs)
  An evoked potential is elicited due to a stimulus and reflects the processing of the physical stimulus. EPs can be the response to a sensory stimulus (auditory, visual, somatosensory, olfactory).
  The response to repetitive sensory stimulation at a rate higher than 6 Hz (e.g. flickering light) is an oscillation with constant frequency and phase called steady-state evoked potential (SSEP) [35].

ERPs that reflect "higher" processing, which involve attention and memory, are:

- Mismatch negativity (MMN, N2a)
  The mismatch negativity depicts a negative component that is elicited about

200 ms after the stimulus. It occurs if the presented cue differs from the previous stimuli. It is assumed to deflect an automatic (unconscious) detection of physical deviation.

- P300
  The P300 is a deflection in the elicited response with a latency of 300 ms (up to 900 ms) to a rare target stimuli in an oddball paradigm. It is maximal over the parietal or central area. One explanation is that the P300 reflects an update of the current environment model due to incoming information. It can be elicited by a novel stimulus or by silently counting the occurrence of a task relevant stimulus.

- N400
  The N400 is a negative deviation with a latency of about 400ms. It is elicited by stimuli that are deviant in meaning, e.g. a non meaningful sentence, for example "'Martin eats a sandwich with stone".

1.1.3 Signal Processing

In order to extract control information for the BCI from the recorded brain signals, the signals need to be processed. Signal processing includes preprocessing of the raw brain signals, feature extraction, and transformation (classification) of the signal [Fig. 1.4].

The intention of preprocessing is to improve the signal-to-noise ratio, e.g. with filtering or artefact removal. Feature extraction aims to extract significant features from the signal and therewith reduces the large amount of data that is received from the measurement. Finally, classification transforms the signal so that the control signal can be extracted. Several different methods for classification can be used, e.g. thresholds, Fisher's discriminant, stepwise linear discriminant analysis, neural networks, support vector machines,... [19, 36].

1.1.4 Mode of Operation

The BCI can either be used at predefined times (cue-based, synchronous BCI) or any time (uncued, self-paced, asynchronous BCI). The cue-based BCI classifies the

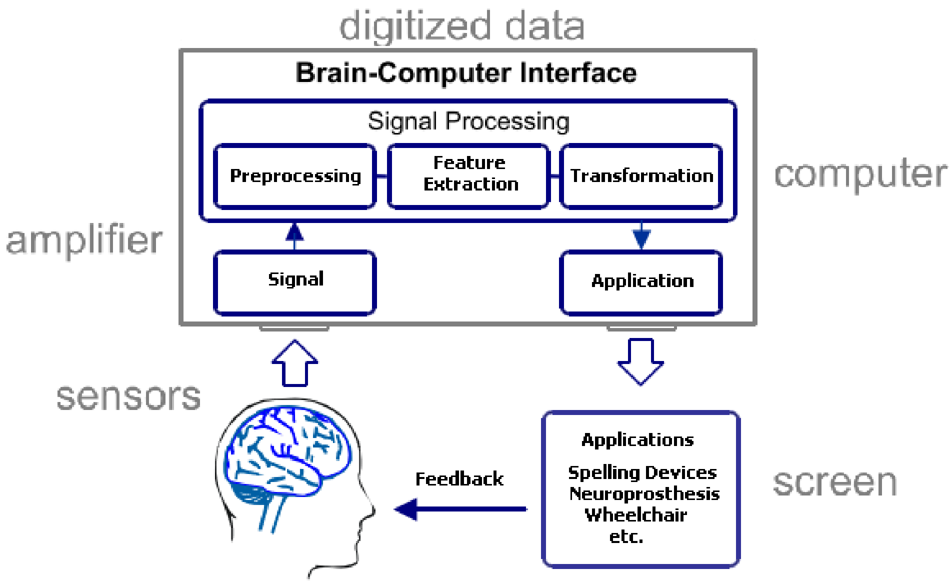


Figure 1.4: Principle of a BCI system [28].

brain signal only at certain time slots, whereas the uncued BCI does a continuous analysis of the brain signal.

1.1.5 Experimental Strategy

The experimental strategy deals with the question how to elicit brain signals that are distinguishable, so that the control signal can be generated.

- Operant conditioning by neurofeedback
  The user learns to change brain signals. An example is the volitional increase or decrease of slow cortical potentials [2, 3].

- Mental strategy – type of "thought"
  Specific "thoughts" are used to activate different neurons and generate different brain signals. Motor imagery is an example where the user imagines for instance hand or foot movements [34].

- Focused attention
  The user focuses attention to specific visual, auditory or somatosensory events; for example the user listens to one specific tone in a series of tones [9, 27, 29]. One strategy where the user focuses attention is the oddball paradigm. The

oddball paradigm is a sequence of stimuli of two or more different categories whereof the stimuli of the target category are usually rare. The task to categorize these events elicits an ERP. One component of the ERP is the P300 [Section 1.1.2] which is often used to classify the different events.

1.1.6 Feedback

When controlling an external device with the BCI, feedback information is needed to give the user information about the progress or the current state of the BCI [Fig. 1.4]. Feedback can either be realistic (e.g. actual hand movement with a neuroprosthesis) or abstract (e.g. bargraph on the screen indicating the direction of the movement). The feedback can be provided visually, auditory or tactile.

1.2 State of the Art

This section describes the development of spelling devices and selected studies about auditory BCIs.

1.2.1 Spelling Devices

People that suffer from almost complete paralysis while sensory and cognitive functions work are in a horrible position – they have no control over their muscles and are not able to use conventional means of communication. Farwell and Donchin were the first to design a spelling device, a P300-based visual matrix speller, in 1988 [9] (abbreviated with visual matrix speller). The user is confronted with a matrix arrangement of characters [Fig. 1.5]. The matrix size is typically 5x5 or 6x6. The rows and columns of the matrix are randomly intensified; this creates an oddball paradigm. The user is encouraged to focus attention to a character. This categorizes the overall intensifications into those intensifications of rows and columns that contain the target character (rare) and those that do not. With the help of the elicited P300 component the target character can be extracted. Farwell and Donchin only tested this approach offline.

The first ray of hope came with the development of Birbaumer's spelling device for the paralyzed. A BCI for spelling records volitional changes in the brain activ-

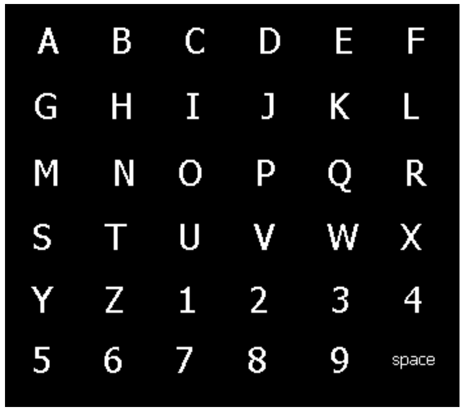


Figure 1.5: Visual speller matrix adapted from [6].

ity and processes them to write letters to communicate. Birbaumer et al. showed that patients could learn to change their SCP to select letters presented on a visual screen; the changes in brain activity were recorded with EEG. Communication was possible, but only at a rate of 0.5 characters per minute. [3]

Also, the offline speller of Farwell and Donchin [9] was further developed into an online system. The usability and typing speed was assessed in several studies. The communication rate ranges from 7.8 characters per minute with an offline accuracy of 80% for healthy participants [6] to 1.2 characters per minute with an accuracy of 82% for people with amyotrophic lateral sclerosis (ALS) [31]. Sellers et al. found that 7 of 16 ALS patients could not write successfully with a visual matrix speller and proposed a simpler communication system for answering yes/no questions [Section 1.2.2] [41].

The Berlin Brain-Computer Interface group developed a mental typewriter based on motor imagery called the Hex-O-Spell [4]. Six hexagonal fields are arranged around a circle, each field contains initially 5 symbols [Fig. 1.6a]. With the imagination of a right hand movement the arrow in the center can be rotated clockwise and imagination of a right foot movement extends the arrow to select characters. Firstly, the user moves the arrow to the field that contains the target letter, extends the arrow and therewith selects the field [Fig. 1.6b]. Now, the hexagonal fields are reloaded with the 5 symbols of the selected hexagon [Fig. 1.6c]. The same procedure is repeated to select the target symbol [Fig. 1.6d].

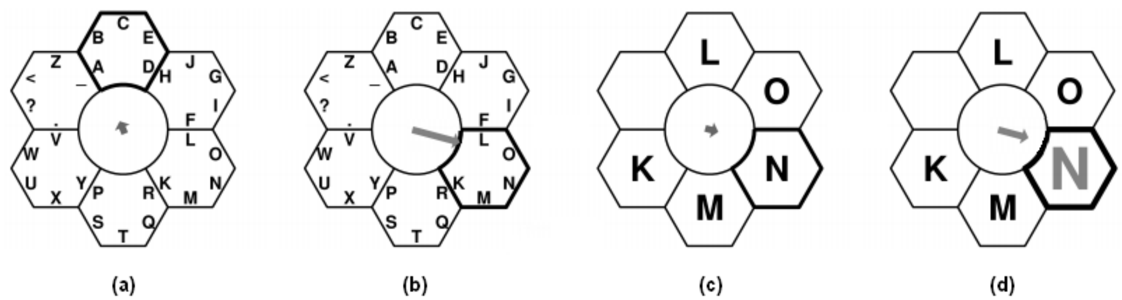


Figure 1.6: Mental typewriter Hex-O-Spell adapted from [4].

Two healthy users demonstrated for one day the operation of the system at the CeBIT[1] 2006 with a typing speed up to 7.6 characters per minute.

The target group for spelling devices are people who are no longer able to communicate in conventional ways. These are people with severe neurological or muscular diseases, e.g. locked-in patients, ALS patients or paralyzed people. Some of these conditions come along with a visual impairment, for example, problems to focus the gaze. In such cases visual based BCIs are not usable, auditory based BCIs on the contrary are feasable [30]. Hence, an auditory correspondent to the visual matrix speller was developed [12, 21]. A 5x5 speller matrix was adapted; the rows are represented with the acoustic signals 1 to 5, the columns with 6 to 10 [Fig. 1.7]. Letters are selected by concentrating on the acoustic numbers for the target row and column. Firstly, the user attends to the stimuli for the target row. Next a stimulus sequence for the columns is presented and the user again focuses attention to the target stimulus.

Studies showed that the performance of the system improved with training. However, in comparison to the visual matrix speller, it performed worse. Nine out of 14 healthy participants achieved a selection accuracy over 70% with a spelling rate of 1.5 minutes per character [12]. Kübler et al. found that four ALS patients achieved a selection accuracy above chance level; the participants reported difficulties to concentrate on the task. [21]

Klobassa et al. created an auditory matrix speller with initial visual cues [17]. The

[1]The world's largest computer fair in Hannover, Germany

Figure 1.7: Auditory matrix speller adapted from [21].

study investigated whether additional visual stimuli can be helpful. The participants were divided into two groups: (A) that received only auditory stimuli, and another group (AV) received auditory plus gradually decreasing visual stimuli. The auditory stimuli for row and column selection were the environmental sounds bell, brass, ring, thud, chord, buzz; the target sound was displayed in brackets. The AV group additionally received intensified rows and columns in the beginning [Fig. 1.8]. The study showed that early visual stimuli did not improve the performance after the visual cues had been removed; the two groups performed equally well. The accuracy of the auditory only group increased with training. Eight out of ten participants achieved a classification accuracy above 50% for the auditory modality; the average mean bit rate was about 2 bits/min. Moreover, the study found that environmental sounds were feasible and no sound bias existed.

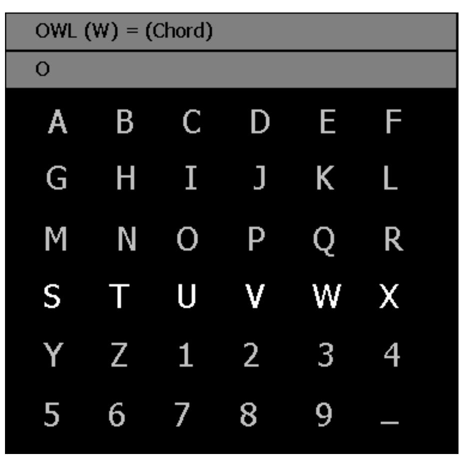


Figure 1.8: Matrix speller with auditory and visual stimuli adapted from [17].

1.2.2 Auditory Evoked Potential Based BCIs

Auditory evoked potential (AEP) based BCIs are a relatively new field of research compared to visual BCIs. Nevertheless, the importance is growing since a lot of severely paralyzed people are not able to control their eye muscles and therefore can not fixate the gaze. Recent studies include an auditory BCI for answering yes/no questions, a BCI based on auditory stream segregation and a spatial BCI.

A BCI for Answering yes/no Questions

Sellers and Donchin evaluated the effectiveness of a simple P300-based BCI with visual, auditory and concurrent visual and auditory stimuli for answering yes / no questions. [39]
Three ALS patients and three healthy participants were part of three online experiments. The stimuli YES, NO, PASS, END were presented (stimulus length: 600 ms, interstimulus interval: 1400 ms); two additional stimuli to yes and no where implemented to decrease the probability for each stimulus from 50 % (not distinguishable) to 25 %. The users attended to the yes or no stimuli to answer questions. The users were presented with random feedback to ensure that the results were not influenced by the performance. The EEG was recorded monopolarly at the locations Fz, Cz, Pz. The signals were digitized at a rate of 256 Hz, bandpass filtered between 0.1 and 50 Hz, moving average filtered (4 samples) and decimated by 4. A stepwise linear discriminant analysis (SWLDA) classification for 0–900 ms data segments followed (entry criterion $p<0.1$, removal criterion $p>0.15$); the classification coefficients were derived from either the first run in every experiment, data of the previous experiment or data of the two previous experiments.

The resulting classification rates for two of the three ALS patients were comparable to those of the healthy participants. ALS patients performed best for auditory stimuli when the SWLDA coefficients were derived from data of the first run in a session; whereas healthy participants achieved the best accuracy in the auditory + visual modality with SWLDA coefficients computed from data of the previous two sessions.

A BCI Based on Auditory Stream Segregation

Kanoh et al. tested a two choice BCI system that worked with selective attention to one of two segregated tone streams [Fig. 1.9]. [16]

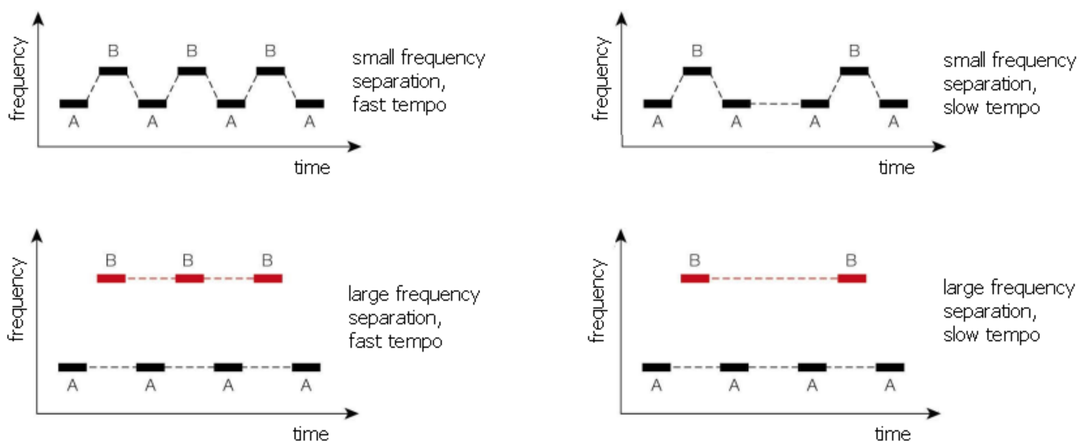


Figure 1.9: Perception of auditory stimuli adapted from [1].

The perception of two alternating tones (ABABAB) as two distinct, simultaneous, segregated tone streams (auditory stream phenomenon) depends on the frequencies of the tones and the presentation tempo. If the frequency difference is small and/or the tempo is slow one coherent stream of tones alternating in pitch is perceived. If the frequency difference is large and/or the tempo is fast the sequence is perceived as two distinct streams, one with frequency A and one with frequency B [Fig. 1.9]. [1]

Six participants with normal hearing ability took part in an experiment of about an hour. Two alternating oddball sequences with short tone bursts of different frequencies were presented to the participant's right ear with earphones (60 ms duration, 200 ms period, 1000 Hz (deviant 700 Hz) and 2000 Hz (deviant 2600 Hz)). The sequences were perceived as two segregated tone streams [Fig. 1.9]. The user counted the deviant stimuli in one tone stream and therewith communicated a binary decision. The EEG was recorded monopolarly at the locations Cz and F3. The signals were amplified, sampled at 1 kHz and bandpass filtered between 1–40 Hz. A linear discriminant analysis (LDA) classification for 10 s blocks was carried out; data segments of 100 ms before to 500 ms after stimulus onset were referenced to the mean amplitudes in the pre stimulus baseline (-50 ms–0 ms).

Deviant auditory stimuli elicited a P300 and MMN response. Using 10-fold cross validation 60 % to 70 % correct detections were found.

A Spatial Auditory BCI

Schreuder et al. tested a BCI paradigm based on spatially distributed acoustic stimuli. [37]
Ten healthy participants took part in the experiments where three different paradigms were tested:

- Paradigm C1000: Spatial noise (75 ms) was used as stimulus with an interstimulus interval of 1000 ms. Eight speakers were used.
- Paradigm C300: Spatial noise and different tones (40 ms) were used as stimulus with an interstimulus interval of 300ms. Five speakers were in use.
- Paradigm C175: Spatial noise and different tones (40 ms) were used as stimuli with an interstimulus interval of 175 ms. Five speakers were used.

The participants were asked to count the number of stimulus occurrences from a target direction. A varying number of electrode locations (60–115) was monopolarly recorded and the 20 best channels were used for the analysis. The recorded signals were analog filtered between 0.01 Hz–250 Hz, sampled at 1 kHz, low pass filtered with 10 Hz, subsampled at 100 Hz, decimated by 5 (with computation of the mean) and normalized to zero mean and unit variance. Signals greater than 70 µV were rejected as artefacts. Data segments (-150 ms–800 ms) were used for a Fisher's discriminant classification with regularization, where the pre stimulus part served as baseline.

With a 10-fold cross validation a classification accuracy greater than 70 % was achieved. The best classification accuracy was achieved with the C1000 paradigm, where also the most prominent P300 component occurred. For the faster conditions, the positive deflection was superimposed on the rhythmic response to the stimuli and produced a discontinuity in the periodic wave pattern. Faster stimulus rates also showed the trend that the maximal difference in the response to a target/non target stimulus shifted to more frontal areas. Shorter interstimulus intervals produced reasonably better information transfer rates and were also reported to make the task easier. The optimal number of stimulus sequences presented in a run was discovered to be individual.

1.3 Assessing a Brain-Computer Interface

When building a spelling device, inevitably the need to assess the performance and quality of the BCI arises. Possible benchmarks to evaluate the communication with a BCI, are the classification score, the bit rate and in a wider sense the spelling rate.

The classification score states the number of correct classifications in a BCI and can be divided into classification accuracy and selection accuracy [37]:

- Classification accuracy
  The classification accuracy states the number of correct classifications (selections) in percent. This is for example correct spelled letters in the visual matrix speller.

- Selection accuracy
  The selection accuracy gives the number of correct sub-classifications in percent; a sub-classification alone is usually not meaningful. In the examples of the visual matrix speller this is the number of correctly classified rows and columns.

Meaningful communication with a BCI requires a selection accuracy of at least 70 %, otherwise the high number of errors slurs the communication [22].

The bit rate or information transfer rate states the amount of information that is conveyed. For target classes with the same probability the bit rate is

$$\text{bit rate} = \log_2 n + p \log_2 p + (1-p) \log_2 \left(\frac{1-p}{n-1}\right)$$

according to Wolpaw et al. [47]. n depicts the number of targets, ans p the probability that the target is correctly classified.
Kronegg et al. found that Wolpaw's bit rate does not give reliable results for a larger number of target classes and recommends Nykopps bit rate instead [18]:

$$\text{bit rate} = H(\hat{W}) - H(\hat{W}|W)$$

$$H(\hat{W}) = -\sum_{j=1}^{M} p(\hat{w}_j) \cdot \log_2 p(\hat{w}_j)$$

$$p(\hat{w}_j) = \sum_{i=1}^{N} p(w_i) \cdot p(\hat{w}_j|w_i)$$

$$H(\hat{W}|W) = -\sum_{i=1}^{N}\sum_{j=1}^{M} p(w_i) \cdot p(\hat{w}_j|w_i) \cdot \log_2 p(\hat{w}_j|w_i)$$

$p(w_i)$ states the probability of the target class w_i and $p(\hat{w}_j|w_i)$ depicts the probability that the target class w_i is recognized as the target class w_j.

The bit rate alone does not provide a realistic measure of the communication speed because the conveyed information can be erroneous (e.g. a wrong letter). Erroneous information does not contribute to true communication in a useful way.

1.4 Motivation

A spelling device benefits from being non visual because it pays attention to visual impairments such as problems to focus or fixate the gaze. Another issue concerning ALS patients is that the P300 responses can be variable or abnormal [41]. This introduces difficulties to detect the P300 components.
Kübler et al. looked into the feasibility of an auditory matrix speller for ALS patients and found problems with concentrating on the task [21]. Thus the task was too taxing. This leads to the conclusion that a simple system with only a few stimuli should be robust and easy to use. A system based on these considerations is tested here: the auditory BCI based Morse code speller. It follows the idea to acoustically code characters with the Morse code [Fig. 2.2] and thus takes into account the requirements of the target group, people with neurological or muscular impairments.

1.5 Objective

The aim of this work is to test an auditory based spelling device that is robust, easy to use, and that has a high classification accuracy. The speller should be based on the Morse code, auditory stimuli and an oddball paradigm with P300 detection. The aim is to implement the system and test it with 10 participants. Different feature selection methods and classifiers are employed in an offline study on the data from the experiments and the best settings are to be determined.

Chapter 2

Methods

2.1 Auditory BCI Based Morse Code Speller

2.1.1 The Idea

The auditory BCI based Morse speller (abbreviated with Morse speller) has the following set up [Fig. 2.1]:

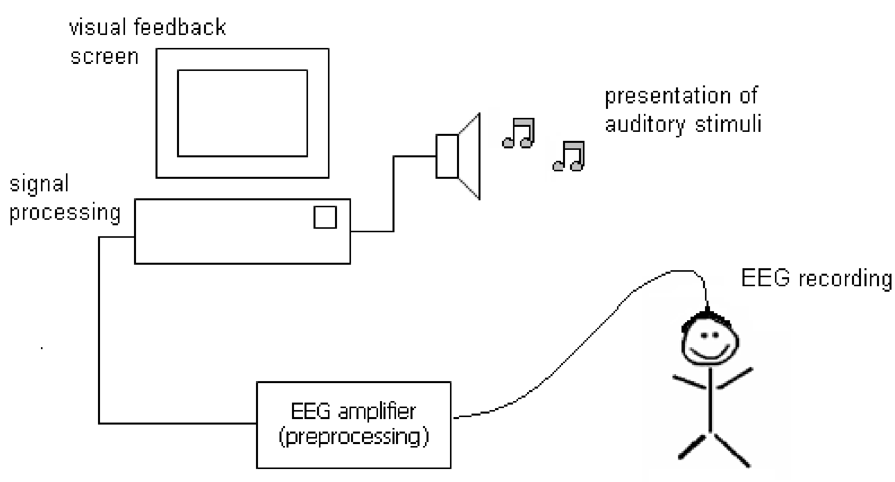


Figure 2.1: Principle of the Morse code based speller.

The auditory stimuli are low and high beep tones which are presented in an oddball paradigm. Focusing attention on the different tones generates different brain responses. The brain potential is recorded with EEG, preprocessed and forwarded to a computer. There, the signal processing is done (detection and classification of P300 deflections), which transforms the brain signal into a control signal. The control signal is used to pilot the spelling system; spelled symbols are shown on a visual feedback screen.

The Morse speller uses a common approach in assistive technology and coded access, where letters are represented with the Morse code. [Fig. 2.2]. The Morse code is composed of the two symbols dot (•) and dash (–).

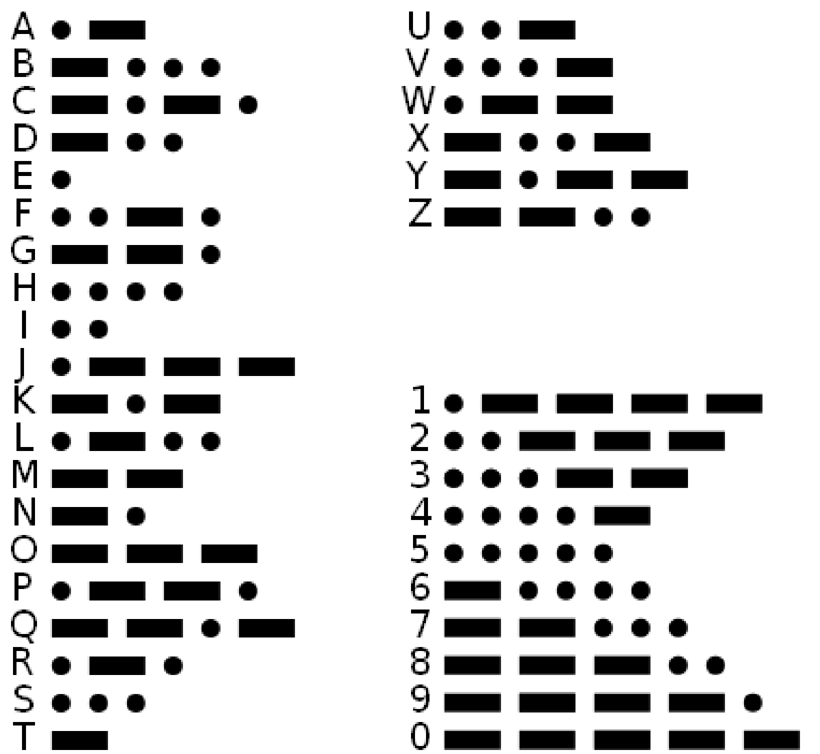


Figure 2.2: Morse Code [45].

2.1.2 System Design

The system design describes the auditory stimuli and their presentation, how letters are coded with the Morse symbols, and gives an overview of the paradigm and the system model.

Auditory Stimuli

The acoustic realization of the Morse speller combines the auditory stream phenomenon [Fig. 1.9] with the oddball paradigm. Auditory streams ease the focus of attention, i.e. the attention can easily flip between the two tone streams. The two tone streams are composed of short beep tones (40 ms) of two-lined and three-lined notes respectively with an interstimulus interval of 160 ms [Fig. 2.3].

- Stream 1: 880 Hz (A'') deviant cue: 784 Hz (G'')
- Stream 2: 1760 Hz (A''') deviant cue: 1568 Hz (G''')

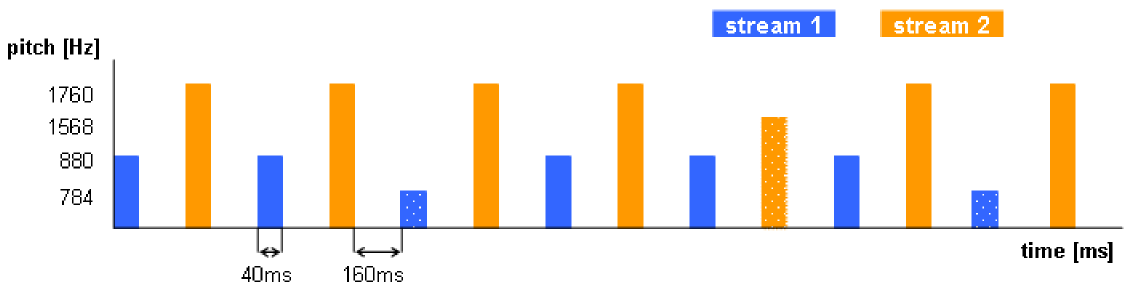


Figure 2.3: Schematic illustration of the tone sequence.

10 % of the stimuli in the Morse speller are deviant in pitch. The amount of deviant cues is the same for the two tone streams. The first and the last tone of the tone stream are not deviant; there is at least one non deviant tone between two deviant stimuli in a stream.

Letter Coding

With the Morse symbols dot and dash only a sequence of Morse code can be written. To write words, the end of a Morse word and a space have to be included in the system design. Therefore, an additional symbol is introduced: the space symbol (⋄). If the space symbol follows a Morse symbol it depicts the end of the Morse word, if it succeeds a space symbol it is interpreted as a space between two letters [Fig. 2.4].

Thus three classes have to be distinguished: (1) dot (2) dash (3a) end of Morse Word (3b) space between two letters. The classes (3a) and (3b) correspond to the same symbol, the space, and are differentiated with respect to the preceding symbol.

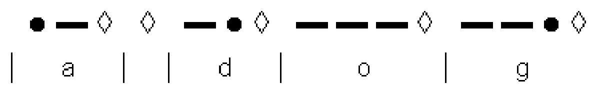


Figure 2.4: Example of Morse coding.

Mapping of the Auditory Stimuli to the Morse Symbols

The deviant stimuli of stream 1 (low deviant tones) are assigned to the Morse symbol dash, the deviant stimuli of stream 2 (high deviant tones) to the dot. That is, focusing attention to the low deviant tones yields a dash, focusing attention to the high deviant tones gives a dot, and not focusing on any of the deviant stimuli results in a space symbol.

Paradigm

The task to spell one Morse symbol is called a trial. At the beginning of a trial (second 0) a spelling instruction appears, which tells the user which deviant stimuli to pay attention to. Two seconds later the presentation of the tone sequence of 30 s length starts. 150 tones are presented, 75 low and 75 high tones with 8 deviant tones each. In the following 1s break an online classification could be done and a feedback could be presented on a screen [Fig. 2.5]. This gives a trial duration of 33 s.

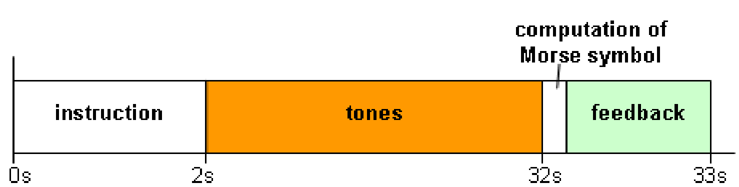


Figure 2.5: Paradigm for writing one Morse symbol.

System Model

The real time model presents the tones, records the preprocessed EEG data, allocates the data with regard to the tone streams, extracts significant features, probes the low and the high tone stream for a P300 deflection (classification), and computes the Morse symbol using the classification result [Fig. 2.6].

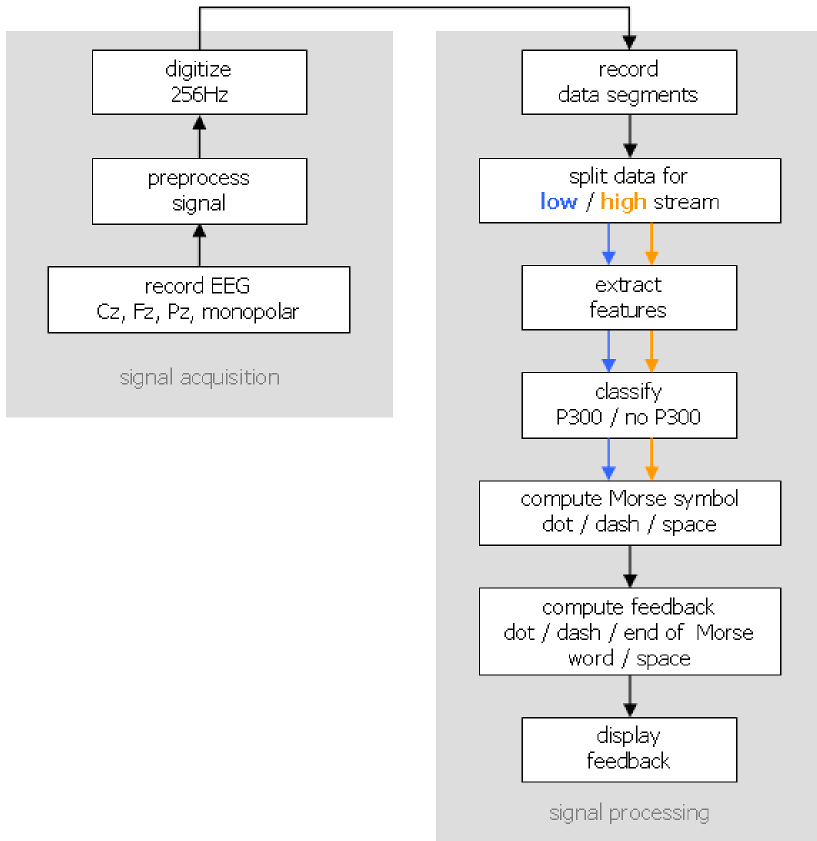


Figure 2.6: Morse speller - model.
This illustration shows the acquisition of the EEG data (left) and the signal processing (right). The model presents also the tones, which is not included here.

It also comprises a visual feedback screen to present the results. The used graphical user interface [Fig. 2.7] is composed of four elements:

1. The symbols of the current Morse word
2. Characters that can be written with the current Morse code (possible characters)
3. Total written text of the session
4. Morse code instruction for the copy spelling mode. The instruction states which deviant stimuli (low, high, none) to pay attention to.

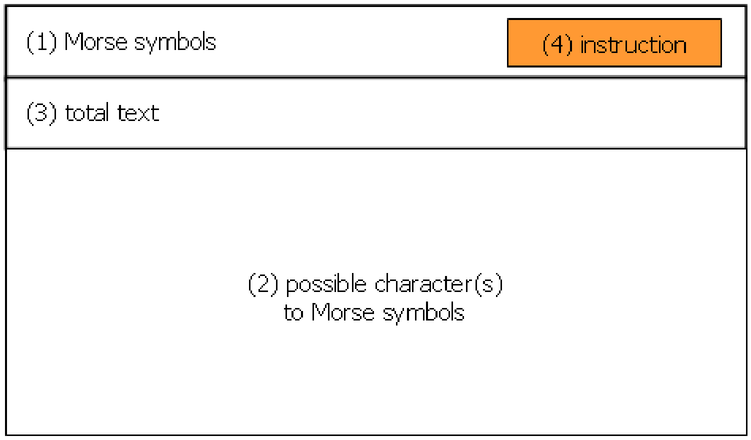


Figure 2.7: Feedback screen.

2.2 Experimental Design

The participants took part in one experiment where 80 Morse symbols in form of the predefined letter sequence

xpzyeduvplqbcyqfj

with 17 letters were copy spelled. The used letters are composed of 3 or 4 Morse symbols, except the letter ‘e’ which is only composed of 1 Morse symbol. The sequence comprises 31 dahes, 32 dots, and 17 space symbols.The task was to focus attention to the deviant stimuli that were determined in the tone instruction. The instruction LOW (HIGH) asked for attention to the deviant stimuli in the low (high) tone stream. It was suggested to the participant to either silently count the stimuli or to intentionally recognize the occurrence of the target cue. The instruction NONE demanded to concentrate on nothing; it is allocated to the space symbol. Moreover, the user was advised to avoid artefacts, i.e. to keep movement, eye blinks, swallowing and muscle tension during the signal recording to a minimum.

The experimental session comprised 80 trials (Morse symbols). The trials were grouped in 8 blocks with 10 Morse symbols each. A 1 s pause preceded every block and a 2 s pause finished it. Within a block, the pause between two trials is of a random length between 1.5 s to 2.5 s. The length of the pause between two blocks was determined by the participant [Fig. 2.8].

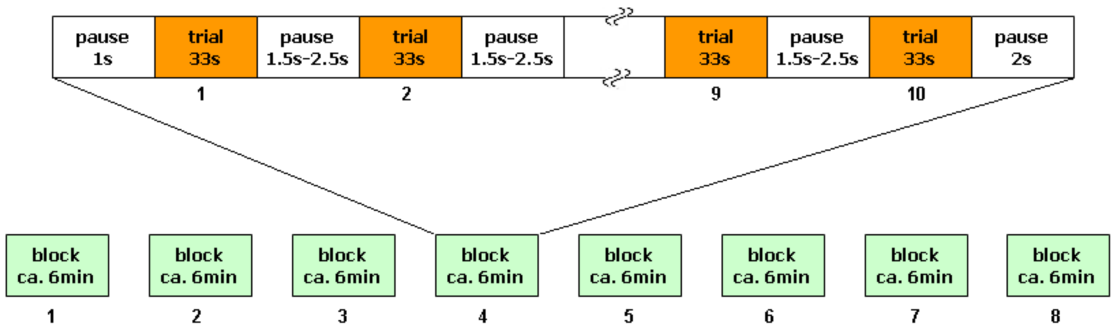


Figure 2.8: Procedure of the test.

The duration of one trial is 33 s; one block with 10 Morse symbols lasts approximately 6 minutes.

$$t_{\text{block}} = 1s + 10 \cdot t_{\text{trial}} + 9 \cdot (1.5s - 2.5s) + 2s = 5\text{min } 51s \approx 6\text{min}$$

This yields approximately one hour for the spelling of 80 Morse symbols. The overall time for the test with electrode fixation, explanation of the procedure and adaptation of individual settings was around 1.5 hours.

No online classification was performed during the experiments, and therefore random feedback was shown to the participants.

2.3 Experiment

The Morse speller was tested with participants to assess the feasibility of the system and to collect data for the offline comparison of different signal processing steps, feature selection and classification methods.

Ten healthy participants tested the Morse speller. Nine had normal hearing ability; one reported to suffer from tinnitus (subject 8). Age ranged from 24 to 48 years (mean age: 29.1±7.3 years). Six participants were male, four female. The participants were explained the functionality of the Morse speller and informed about the procedure of the test. All participants gave informed consent prior to the test.

The subjects sat comfortably in a chair facing the feedback screen. The auditory stimuli were presented with headphones. The distance of the feedback screen and the volume were adapted to the users' comfort. It was made sure that the participants could hear the two tone streams with the deviant stimuli and that the task was clear.

2.4 Data Acquisition and Preprocessing

The EEG is recorded using a fixation cap and sintered Ag/Ag Cl electrodes of the company Easycap[1]. The cap has 64 electrode locations [Fig. 2.9] embedded over the scalp based on the International 10-20 system [15]. The 10-20 system distributes the electrodes at positions with 10 % or 20 % of the distance from inion to nasion and the left to right preauricular point.

The classical channel set with the electrode positions Fz, Cz and Pz [Fig. 2.9] is chosen for the EEG recoding. All electrodes are referenced to the left mastoid and grounded to the right mastoid.

The EEG signals are amplified with the biosignal amplifier g.BSamp of the company g.tec[2]. The amplifier has a built-in low pass filter, high pass filter, notch filter, and sensitivity control. The low pass filter was set to 100 Hz, the high pass filter to 0.5 Hz, the 50 Hz notch filter turned on and the sensitivity set to 0.1 mV.

The data acquisition and digitalization is provided by the data acquisition card NI 6033E of the company National Instruments[3]; the sample frequency (f_s) is set to 256 Hz.

2.5 Signal Processing

All data processing and analysis was accomplished with Matlab and Simulink[4] on a standard PC. The toolboxes BioSig and rtsBCI from the BioSig Project[5] were used. The BioSig toolbox provides biomedical signal processing tools; with the rtsBCI

[1] EASYCAP GmbH, Herrsching-Breitbrunn, Germany, www.easycap.de
[2] g.tec Guger Technologies OEG, Graz, Austria, www.gtec.at
[3] National Instruments Corporation, Austin, Texas, United States, www.ni.com
[4] Mathworks Inc., Natick, Massachusetts, United States, www.mathworks.com
[5] The BioSig Project, biosig.sourceforge.net

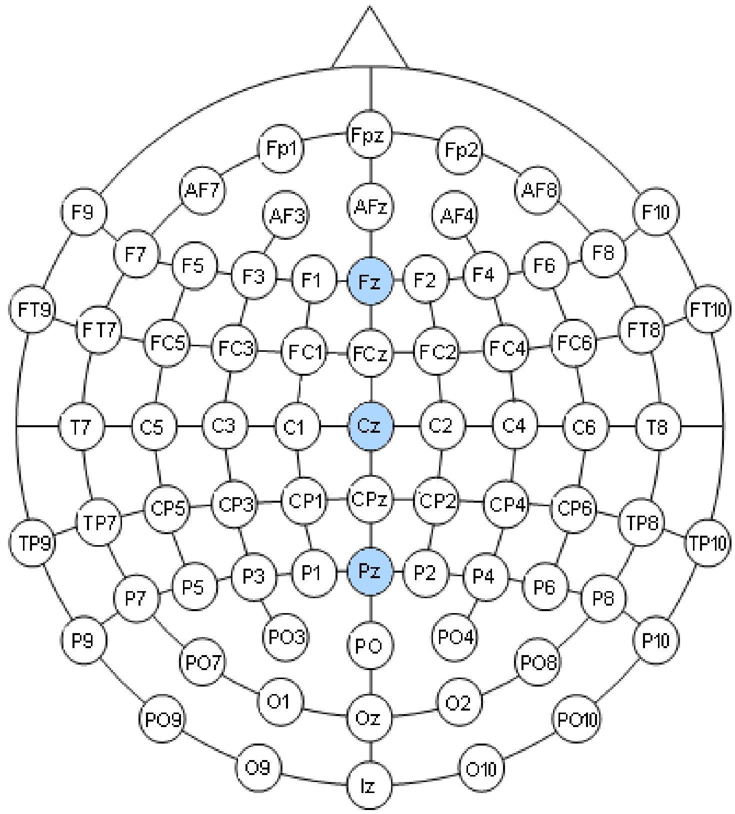


Figure 2.9: EEG positions and recorded locations (blue).

module real time data processing is possible. The SWLDA classifier of the BCI2000 software[6] was incorporated. Additionally, statistical tools from the Bioinformatics toolbox were employed for the offline analysis.

2.5.1 Data Segments

With the onset of every tone a data segment is collected for the following feature extraction and classification. The Morse speller operates on a sample based time; the conversion of a time value in seconds to samples is

$$t[\text{samples}] = \text{floor}(f_s \cdot t[s])$$

The floor operator ensures that t[samples] is a natural number.

Four different data segments are compared in the offline analysis:
(1) 0–800 ms (2) 0–1200 ms (3) 200–800 ms (4) 200–1200 ms

[6] BCI2000, www.bci2000.org

0 indicates the onset of the stimulus.

A data segment of 800 ms equals 204 samples for a sample frequency of 256 Hz. Therewith, the data population of one trial is 150 tones x 3 channels x 204 time points [Fig. 2.10]. The input data of every trial is split into sets for the low, low deviant, high and high deviant tones. Every trial consists of 67 normal low tones, 8 deviant low tones, 67 normal high tones and 8 deviant high tones.

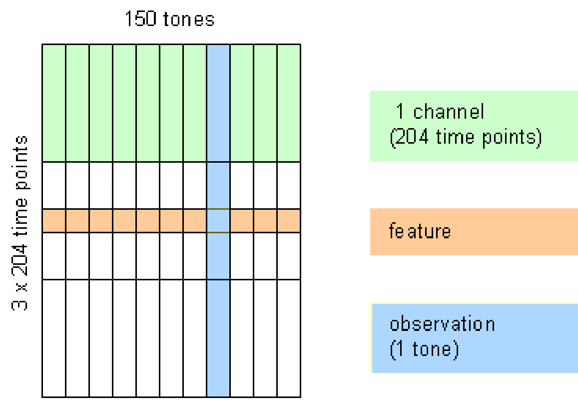


Figure 2.10: Input data.
(one trial, 800 ms data segment)

2.5.2 Feature Selection

To reduce the dimensionality of the input data a feature selection is applied. Many features are not relevant for classification, the feature extraction aims to reduce or transform the input space to those features that give the best classification.

Different signal processing steps and feature extraction methods are investigated:

- digital low pass filter: none, 10 Hz, 60 Hz
- moving average filter: none, 4, 8, 12
- reference: monopolar, bipolar
- discriminant power algorithm: 4, 10, 20 points
  (only for LDA classification)

Filtering

The low pass filter is a digital Butterworth filter of order 3; cut-off frequencies of 10 Hz and 60 Hz are implemented. The moving average filter is implemented as a digital FIR filter in direct form II. The moving average of 4, 8 and 12 samples is computed which reduces the sampling frequency of 256 Hz to 64 Hz, 32 Hz and 21.3 Hz respectively.

Bipolar Channels

Three bipolar channels are computed out of the monopolar referenced channels and are used for the classification.
channel 1 = Fz - Cz
channel 2 = Fz - Pz
channel 3 = Cz - Pz

Discriminant Power Algorithm

A very simple and effective feature selection method for EEG data is a variation of the discriminant power (DP) method for noisy data [10]. The discriminant power function rates a feature on its ability to discriminate between the classes. This property is reflected in a probability density function (pdf) of the feature, which is different for each class, i.e. the distributions for each class are well separable and not overlapping.

The basic discriminant power function actually only computes the minimum and the maximum values of a feature for every class and counts the amount of samples that do not lie in the overlapping zone [Fig. 2.11]. For a two class problem the discriminant power of one feature is:

$$d_1 = \sum_{i=1}^{N_1} (s_{f1}(i) > \max(s_{f2})) + (s_{f1}(i) < \min(s_{f2}))$$

$$d_2 = \sum_{i=1}^{N_2} (s_{f2}(i) > \max(s_{f1})) + l(s_{f2}(i) < \min(s_{f1}))$$

$$DP(f) = \frac{d_1 + d_2}{N_1 + N_2} \qquad \in [0;1]$$

where s denotes a sample; f depicts a feature, and the indices 1/2 note the class membership. d_1, d_2 depict the discriminant fraction of class 1 and 2. N_1, N_2 are the number of samples in class 1 and 2.

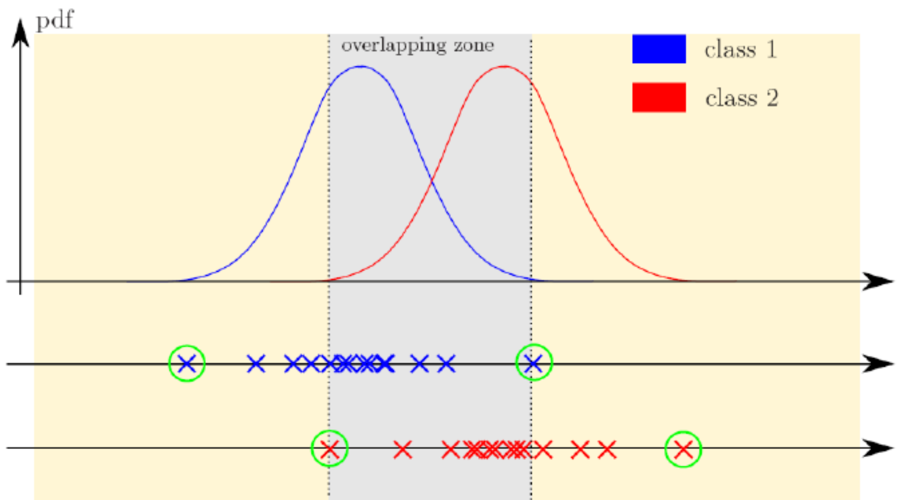


Figure 2.11: The basic DP algorithm [10].

A problem of the basic discriminant power algorithm is that outliers distort the overlapping zone in the calculation, thus the result is not representative of the separating ability anymore. To overcome the problem of noisy data, the distributions of the classes are truncated in a preprocessing step. This ensures that only data within a certain distance (confidence interval) from the mean are kept, and outliers are discarded [Fig. 2.12].

For the normal distribution a confidence interval of $[\mu - 1.77 \cdot \sigma; \mu + 1.77 \cdot \sigma]$ is known to correspond to 76 % of the original population (μ is the mean, σ is the standard deviation). This confidence interval is used in the Morse speller.

The discriminant power algorithm selects the features that are best classifiable. The best classification results are achieved with a small number of discriminant power points [42]. Thus 4, 10 and 20 discriminant power points were computed and used as features in the following classification. Since the stepwise linear discriminant analysis assesses the input features and uses only significant ones itself, it is reasonable to use the DP algorithm only in combination with linear discriminant analysis (LDA) classification. The discriminant power algorithm is applied to the input data to

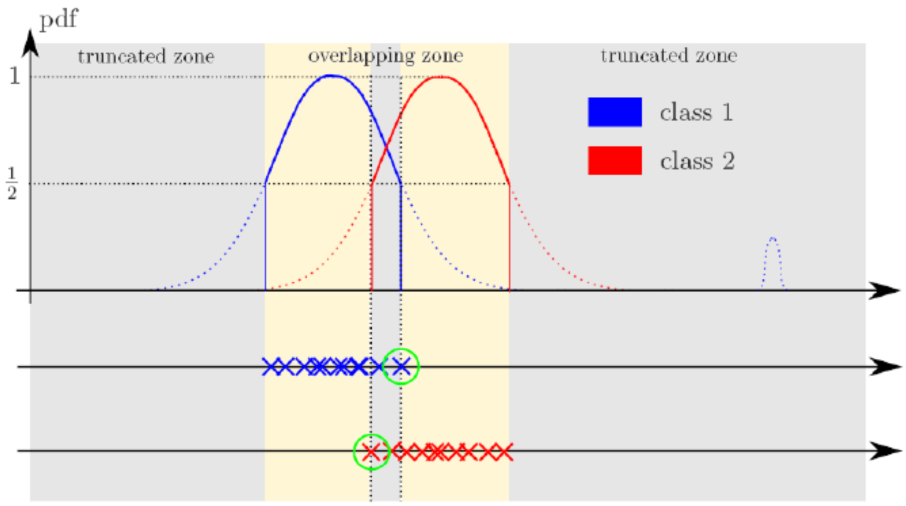


Figure 2.12: The modified DP algorithm [10].

find the best classifiable points of all three channels [Fig. 2.13]. These points serve as features for the following classification. The DP points are invariant and denote those time points of the three channels where the difference between an elicited P300 and no P300 is largest.

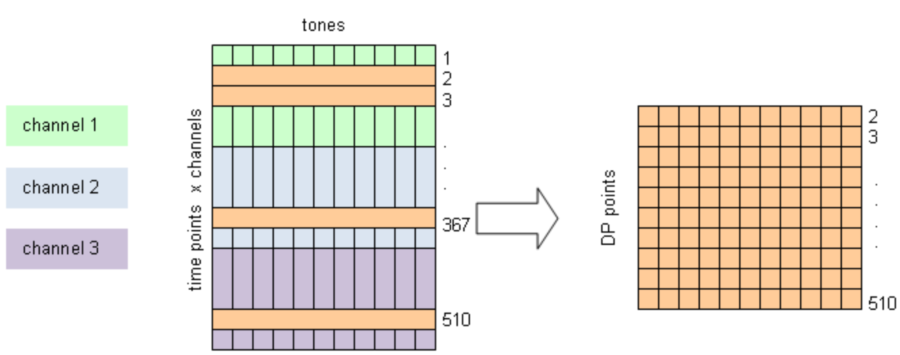


Figure 2.13: Feature selection with the DP algorithm.
The DP algorithm is applied to the whole input population and selects the best classifiable points of all channels.

Artefact Removal

If an amplitude value of 70 µV or more is found in the data segment of a stimulus, the data segment for this stimulus is discarded (for all channels).

2.5.3 Classification

Three different classification methods are investigated in this work:

1. SWLDA
2. Fisher's discriminant (LDA)
3. Multiple LDA

All three classification methods classify linear separable data.
Every tone stream is probed for a P300 potential, i.e. the Morse speller implements two classifiers, one for the low and one for the high tone stream. The detection of a P300 potential is a binary classification problem where the classes "elicited P300" and "no elicited P300" are differentiated.

Fisher's Discriminant

Note that boldface letters denote vectors ($\mathbf{x}$) or matrices ($\mathbf{X}$) and that vectors are column vectors.

Fisher's discriminant model is

$$\mathbf{y} = \mathbf{X}\mathbf{w} + b$$

where $\mathbf{y}$ contains the discriminant scores for all observations, $\mathbf{w}$ are the weights, b is the bias (threshold) and $\mathbf{X}$ is an input matrix, where every column denotes a feature (predictor variable) and every row an observation. Thus the predicted class for one input vector (observation) $\mathbf{x}$ is

$$y(\mathbf{x}) = \mathbf{x}^T\mathbf{w} + b$$

The discriminant function projects the input vector onto one dimension ($\mathbf{x}^T\mathbf{w}$), so that the class separation is maximized and the threshold assigns the input to a class. The decision boundary is 0. The weight vector that gives an optimal projection maximizes the class separation (covariance matrix $\mathbf{S}$) and minimizes the variances within the classes (means $\mathbf{m_1}$, $\mathbf{m_2}$):

$$\mathbf{w} \propto \mathbf{S^{-1}}(\mathbf{m_2} - \mathbf{m_1})$$

For a two class problem the solution of this optimization problem can be computed with the least squares method:

$$\mathbf{w} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$

$$b = \text{mean}(\mathbf{y}) - \text{mean}(\mathbf{X}) \cdot \mathbf{w}$$

$\mathbf{X}$ contains the features (columns), $\mathbf{y}$ contains the target class to every observation in $\mathbf{X}$ (rows), and mean($\mathbf{X}$) computes the mean of every feature.

Fisher's discriminant is combined with a preceding discriminant power algorithm [Section 2.5.2], which aims to find the best classifiable features in the input space.

Multiple LDA

The features for the classification are selected with a preceding DP algorithm [Section 2.5.2]. One Fisher's discriminant is implemented for every DP feature, i.e for 10 DP points 10 LDA classifiers are used. The class that achieves the most scores within all the Fisher's discriminants is the final result (decision by majority).

Stepwise Linear Discriminant Analysis (SWLDA)

This classification method is an extension of the linear discriminant analysis where only statistically significant features (predictor variables) are used to compute the weights of the discriminant function. [8, 9]

1. Firstly, the significance of all features is computed with an F-Test. Small p-values indicate a relationship between the feature and the target variable [23].

2. The features are sorted with regard to their p-values.

3. Forward analysis: Predictor variables with a p-value smaller than the entry criterion are added to the discriminant model in ascending order.

4. After the entry of a new feature to the model a backward analysis is performed: The significance of the features in the model is tested. Features with a p-value greater than the removal criterion are removed from the model.

5. This procedure is repeated until no more features satisfy the entry / removal criteria or the discriminant model includes a predefined number of features.

The entry and removal criterion for the SWLDA are set to the usual $p<0.1$ and $p>0.15$, respectively. The setting for the number of maximum features are 4, 10 and 60.

Prediction of the Morse Symbol

The prediction of a P300 in the low tone stream and no P300 for the high tone stream results in the Morse symbol dash. A vice versa prediction (no P300 in the low stream, P300 in the high stream) results in the Morse symbol dot. Otherwise the prediction is a space symbol [Tab. 2.1].

Table 2.1: Prediction of Morse symbols.

P300 found in		**Morse**
low stream	**high stream**	**symbol**
✓	x	dash
x	✓	dot
✓	✓	space
x	x	space

Thus the theoretical probability to select a dot or a dash by chance is 25 %, the chance level for the space symbol is 50 %.

2.6 Offline Analysis

The recorded data of the experiment was examined in an offline analysis. The effects of different numbers of features, filtering and referencing on the classification score are tested [Fig. 2.14].

The filtering tasks are performed on the acquired EEG signal. The bipolar channels are also computed from the entire EEG signal. After the signal processing steps are done, the data segments after the stimuli are extracted from the EEG signal. Only the data segments after deviant tones are used. This gives 8 data segments in the high tone stream and 8 data segments in the low tone stream. The average of the data segments of the high/low deviant tones is computed. The averaged data

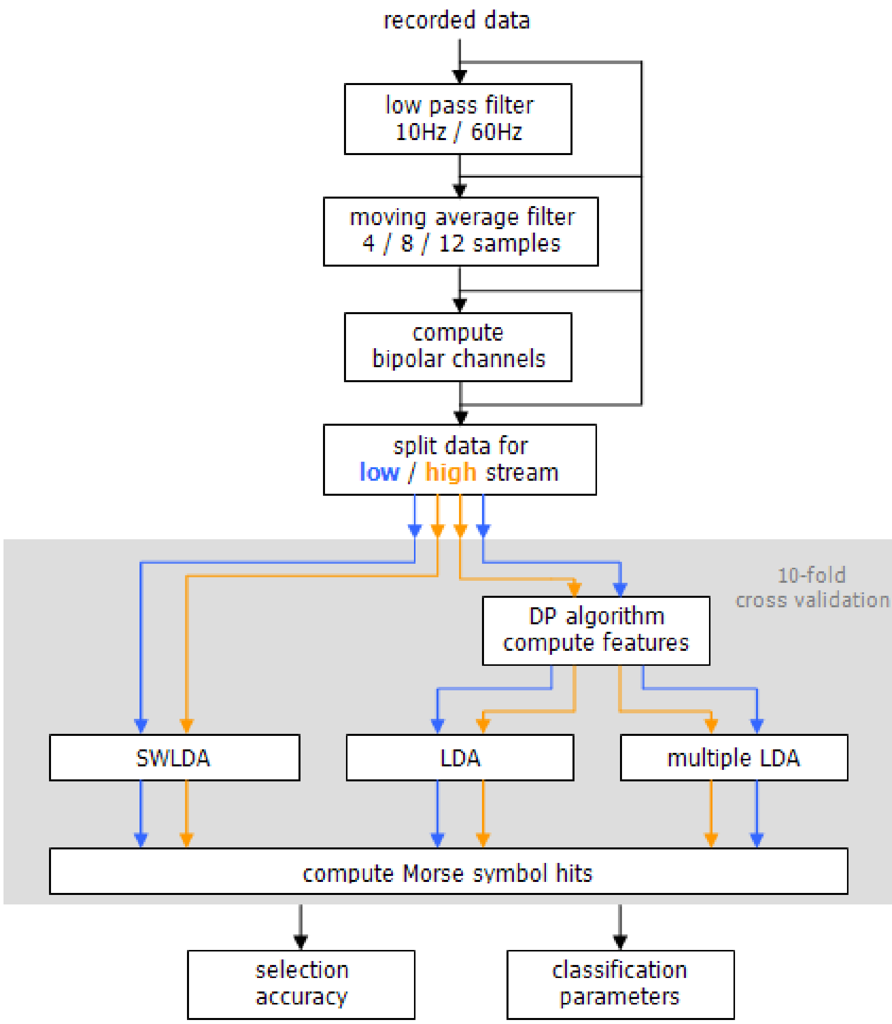


Figure 2.14: Offline analysis.

segment of the low deviants is used in the classifier for the low tone stream. The averaged data segment of the high deviants is used in the classifer for the high tone steam. The classifiers probe each tone stream for a P300 deflection and from the outcome the Morse symbol is computed.

Chapter 3

Results

Waveforms and classification scores are listed in this chapter. The average waveform of all subjects is plotted. A color coded amplitude plot for all trials with respect to the time is shown. The mean and variance of the data segments is plotted. The selected features are displayed and a correct classified trial is presented. The classification scores for the SWLDA, LDA and multiple LDA classification due to different numbers of features, classification settings, filtering and referencing are listed.

3.1 Waveform Analysis

For visual inspection, all signals were low pass filtered with 10 Hz.

Grand Average AEPs

Every subject spelled 80 Morse symbols. Thus, for every subject data of 80 trials was available. Each trial comprised 8 deviant tones in the low and the high tone stream. Only the deviant data segments are used for the classification, i.e the grand average is computed of 80 x 8 (640) data segments.

Data segments associated with deviant stimuli were averaged over all trials for each electrode location (Fz, Cz, Pz). The segments are grouped into targets and non targets for the high and the low tone stream. [Fig. 3.1] shows the AEPs elicited in the oddball paradigm of the high and low tone stream of three selected subjects. Subject 5 and 6 delivered good discriminant scores, 88.8 % and 90.4 %, subject 10 on the contrary achieved only a selection accuracy of 55 % with the SWLDA classifier. The remaining subjects can be found in the appendix [Fig. A.1].

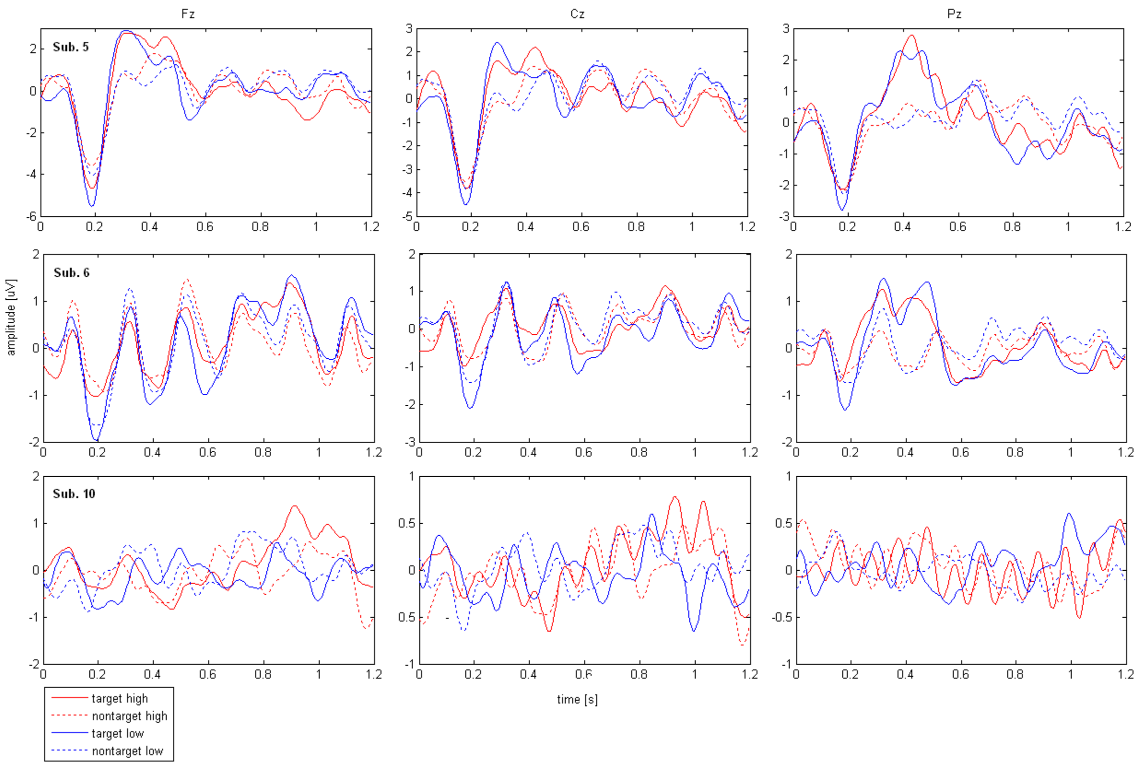


Figure 3.1: Grand average AEPs.
The grand average comprises data of 80 Morse symbol. The deviant target and non target tones in the high tone stream (red) and the low tone stream (blue) are shown. 8 deviant high and low tones are presented for one Morse symbol which are used for the classification; this gives 640 averages for every curve. The stimulus occurs from 0-40 ms. The stimulus period is 200 ms considering all tones and 400 ms for tones within the tone stream.

Variation in the Data

[Fig. 3.2] depicts the mean (solid line) and variance (dashed area) of subject 5. The mean and variance is computed of all input data – 80 trials, comprising 8 deviant tones each, for the high and low tone stream. The remaining subjects can be found in the appendix [Fig. A.2].

[Fig. 3.3] shows a different representation: the averaged data segments (0–1200 ms) of all trials are plotted, the color depicts the amplitude value with respect to the time. The target and non target waveforms of the locations Fz, Cz, Pz for subject

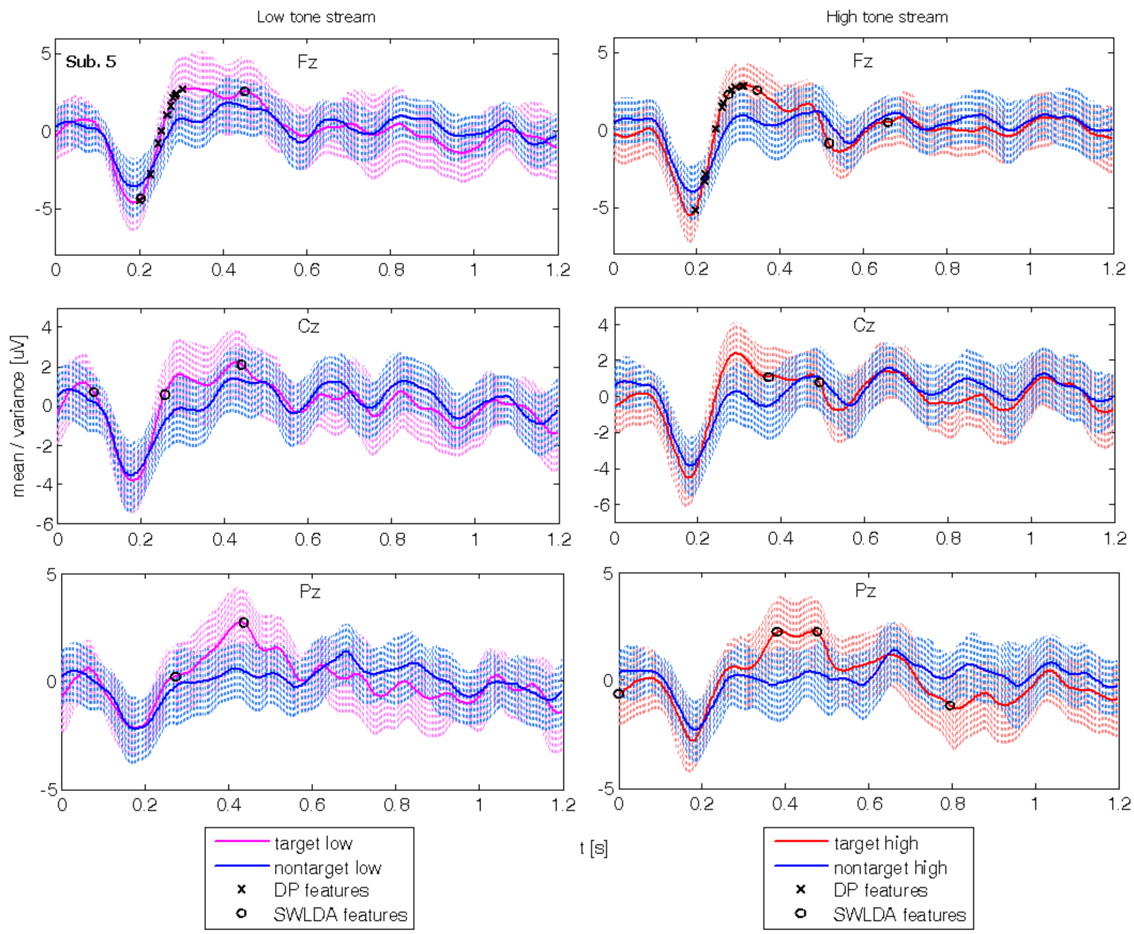


Figure 3.2: Mean, variance and features - Subject 5.
The plot shows the mean (solid line) and the variance (dashed area) of target and non target responses in the low (left column) and high (right column) tone stream. Moreover the features determined by the discriminant power algorithm (10 DP points) and the significant features of the SWLDA classification (10 maximum features) are plotted. The features are computed with 10-fold cross validation whereas the mean and the variance are computed of the overall data (80 Morse symbols x 8 deviant tones per tone stream = 640 averages).

5 are depicted. The figures for the remaining subjects can be found in the appendix [Fig. A.3]. All subjects use the same amplitude range to keep the plots consistent.

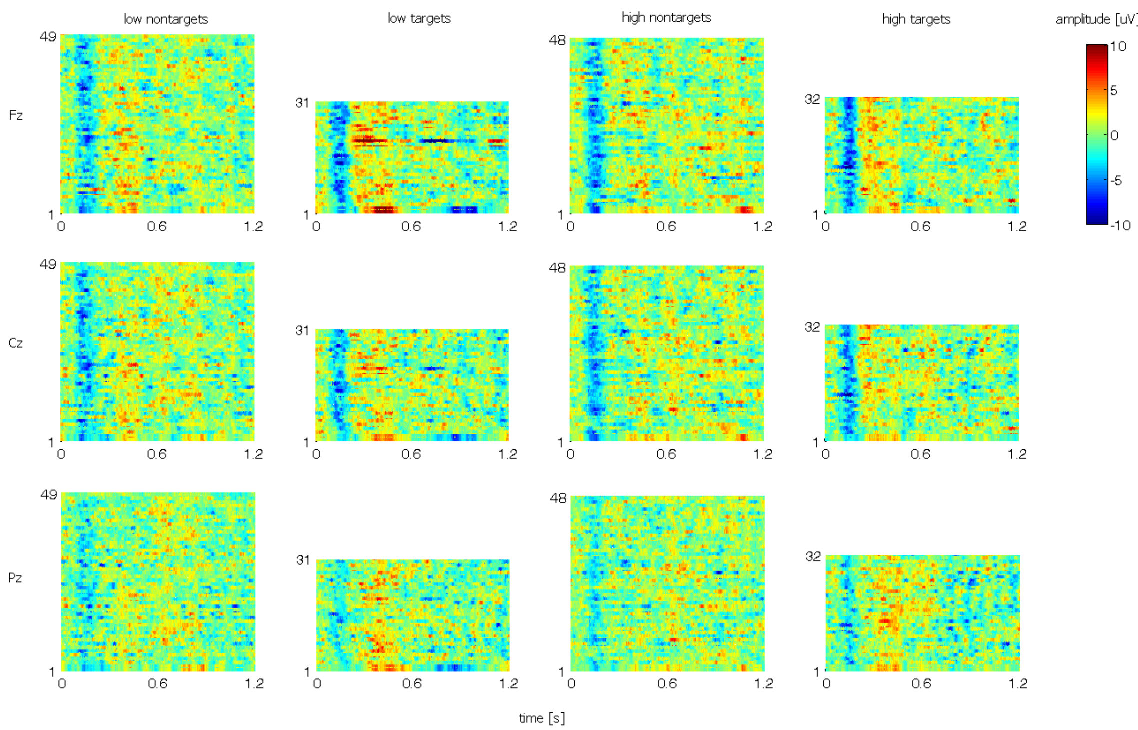


Figure 3.3: AEPs with respect to the trial and time - Subject 5.
The average amplitude for every trial (8 averages) is depicted color coded. The trials are plotted vertically against a 1200 ms data segment (horizontal). The AEPs are shown for the low and the high tone stream and the locations Fz, Cz and Pz.

Features of the SWLDA Classifier

[Fig. 3.2] presents an example of the significant features in the SWLDA classification. The data was low pass filtered with 10 Hz, and 0–800 ms data segments were used for the classification. The number of maximum features for the SWLDA was set to 10. The figure depicts the mean (solid line) and variance (dashed area) of all 80 trials, comprising 8 deviant tones each, for the high and low tone stream. The 10 features, which the SWLDA classifier computed to be most significant, are marked in the plots with circles. Subject 5 obtained a selection accuracy of 88.8 % with 10-fold cross validation. The significant features of the remaining subjects can be found in the appendix [Fig. A.2].

Features of the LDA Classifier

[Fig. 3.2] also presents an example of the feature selection with the discriminant power algorithm. Ten DP points in a data segments of 200–1200 ms were extracted. The features of the offline analysis are marked with crosses in the plots. This setting yielded a selection accuracy of 74.4 % for subject 5 with Fisher's discriminant classifier and 10-fold cross validation. The figures for the remaining subjects can be found in the appendix [Fig. A.2].

Examples of Correctly Classified Trials

[Fig. 3.4] and [Fig. 3.5] present arbitrary trials of subject 6, where the classification of high and low targets and non targets was correct. The oddball paradigm for one trial presents 8 deviant tones for each tone stream, i.e. 8 data segments are averaged in a trial.

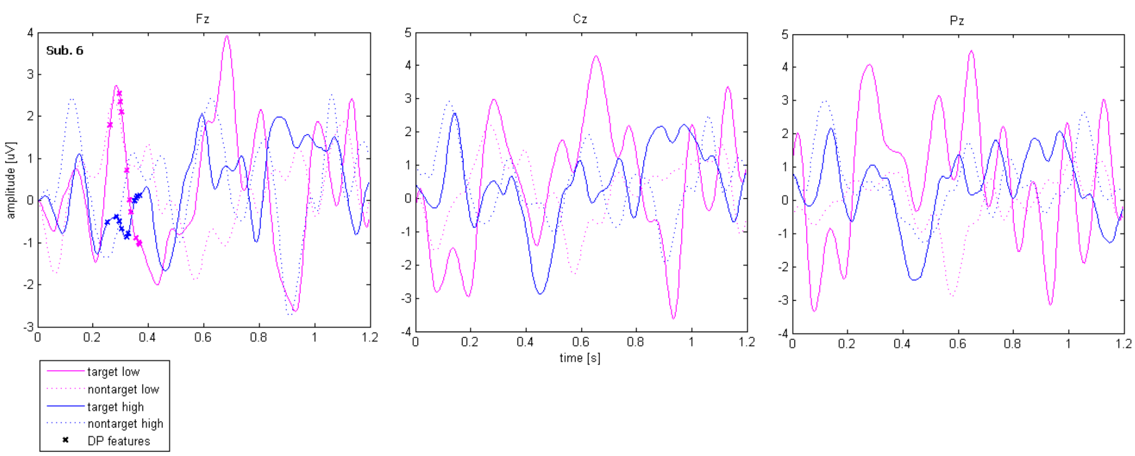


Figure 3.4: Correct SWLDA classification - Subject 6.
This plot depicts the correct classification of a high and low target and non target of an arbitrary trial. An SWLDA classifier with 10 maximum features was used. 8 deviant high and low tones were presented in every trial. Thus, 8 data segments are averaged in every trial and probed for a P300 deflection.

The feature extraction contains chance components, i.e. the selected features are not the same for the different figures, since they were plotted in different runs.

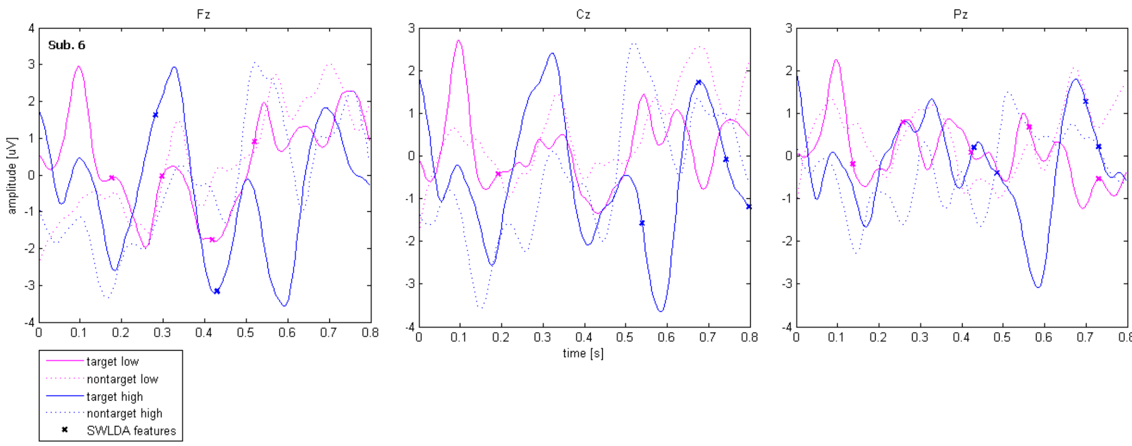


Figure 3.5: Correct LDA classification - Subject 6.
This plot shows the correct classification of a high and low target and non target of an arbitrary trial. An LDA classifier with preceding discriminant power algorithm (10 features) was used. 8 deviant high and low tones were presented in every trial. Thus, 8 data segments are averaged in every trial and probed for a P300 deflection.

3.2 Classification Performance

In the following tabels the best overall selection accuracy for every subject (of all settings, features selection and classification methods) is marked in boldface numbers.

The results of the different settings (number of features, data segments, filtering, and referencing) were examined with a T-test. The significance level was set to 0.01 and a Bonferrroni correction used.

Stepwise Linear Discriminant Analysis

The SWLDA was performed using different settings; 4, 10 and 60 maximum features were tested [Tab. 3.1]; no significant difference was found between the settings. The entry and removal criterion of the SWLDA classifier were kept to $p<0.10$ and $p>0.15$ for all classifications. Furthermore, the optimal data segment length after the stimulus was examined [Tab. 3.2], no significant difference was found.
Moreover, the effects of digital filtering on the selection accuracy were investigated [Tab. 3.3]. The signal is analog band pass filtered between 0.5 and 100 Hz in a preprocessing step. An additional low pass filter of 10 Hz and 60 Hz was examined. Furthermore, a moving average filter of 4, 8 and 12 samples was tested, which

Table 3.1: Accuracy and maximum SWLDA features.

max. features	**selection accuracy**										
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**mean**
4	69.4	70	69.4	72.5	**88.8**	**90.6**	81.3	61.3	76.3	50	72.9
10	71.9	76.9	63.7	69.4	81.9	83.1	79.4	63.1	81.3	45.6	71.6
60	67.5	74.4	71.9	65.6	78.8	83.1	81.3	65	82.5	45.6	71.6

Monopolar reference, 0–800 ms data segment, low pass filter 10 Hz, artefact removal.

Table 3.2: Accuracy and data segment length.

length [ms]	**selection accuracy**										
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**mean**
0–800	71.9	76.9	63.7	69.4	81.9	83.1	79.4	63.1	81.3	45.6	71.6
200–800	72.5	77.5	69.4	65.6	85	87.4	**84.4**	59.4	85	40	72.6
0–1200	66.3	71.3	63.7	68.8	80.6	82.5	72.5	64.4	80.6	46.9	69.8
200–1200	60.6	75.6	70	75.6	78.8	86.9	79.4	66.9	75.6	53.8	72.3

Monopolar reference, low pass filter 10 Hz, max. features 10, artefact removal.

reduces the input frequency of 256 Hz to 64 Hz, 32 Hz, and 21.3 Hz, respectively. A significant difference was found between no filter–MA 21.3 Hz, LP 60 Hz–MA 32 Hz, LP 60 Hz–MA 21.3 Hz, MA 64 Hz–MA 32 Hz, and MA 64 Hz–MA 21.3 Hz. It seems that for finding the best user adjustment the different filter settings have to be tested.

Table 3.3: Accuracy and low pass filtering (LP) / moving average filtering (MA).

filter	**selection accuracy**										
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**mean**
no	70	71.9	63.7	61.9	74.4	78.1	68.8	54.4	**86.3**	53.8	68.3
LP 10 Hz	72.5	77.5	69.4	65.6	85	87.4	84.4	59.4	85	40	72.6
LP 60 Hz	66.3	**78.8**	58.8	65.6	77.5	84.4	78.1	63.7	83.1	52.5	70.9
MA 64 Hz	68.8	72.5	54.4	64.4	75	82.5	73.1	52.5	81.3	52.5	67.7
MA 32 Hz	**73.1**	76.3	65.6	70	80.6	88.1	80.6	56.9	83.1	55	72.9
MA 21.3 Hz	67.5	68.1	63.7	68.8	86.6	89.4	73.8	**70**	76.9	43.8	70.8

Monopolar reference, 200–800 ms data segment, max. features 10, artefact removal.

The EEG was recorded monopolarly. For comparison three bipolar channels were computed out of the monopolar channels. A significant difference was found; two out of the 10 subjects (S3, S4) achieved their best selection accuracies with bipolar referencing [Tab. 3.4].

Table 3.4: Accuracy and EEG referencing.

reference	selection accuracy										
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**mean**
monopolar	72.5	77.5	69.4	65.6	85	87.4	84.4	59.4	85	40	72.6
bipolar	71.9	78.1	**73.1**	**78.8**	84.4	81.3	78.8	65	75	49.4	73.6

200–800 ms data segment, low pass filter 10 Hz, max. features 10, artefact removal.

Fisher's discriminant (LDA)

Since the linear discriminant analysis performs no intern significance test of the features, it was coupled with a preceding discriminant power algorithm. Four, 10 and 20 features (DP points) for the following classification were tested [Tab. 3.5]; the optimal number of features seemed to be individual.

Table 3.5: Accuracy and number of discriminant power points.

DP features	selection accuracy										
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**mean**
4	58.8	46.3	53.1	50	60	56.3	48.1	50.6	59.4	49.4	53.2
10	55.6	43.1	54.4	50	56.3	53.1	50	41.3	60	51.9	51.6
20	51.9	46.9	51.2	50.6	51.2	59.4	49.4	46.3	63.7	55.6	52.6

Monopolar reference, 0–800 ms data segment, artefact removal.

Comparing the different data segment lengths yielded no significant difference [Tab. 3.6]. For all further examinations 4 discriminant power features and a 200–800 ms data segment were chosen as the basic settings.

Table 3.6: Accuracy and data segment length.

length [ms]	selection accuracy										
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**mean**
0–800	58.8	46.3	53.1	50	60	56.3	48.1	50.6	59.4	49.4	53.2
0–1200	62.5	43.8	42.5	53.8	64.4	56.3	54.4	58.1	63.1	48.1	54.7
200–800	59.4	55.6	59.4	60.6	70	66.3	57.5	53.1	66.3	56.3	60.4
200–1200	60.6	50	61.3	66.9	74.4	63.1	54.4	53.8	65	56.9	60.6

Monopolar reference, 4 DP features, artefact removal.

The compared filter settings, which included no filter, a low pass filter of 10 Hz and 60 Hz, and moving average filters of 4, 8 and 12 samples. The moving average filter of 4, 8 and 12 samples reduced the signal frequency to 64 Hz, 32 Hz and 21.3 Hz, respectively. No significant difference was noted. The best results are individual [Tab. 3.7].

Table 3.7: Accuracy and moving average filtering (MA).

filter	**selection accuracy**										
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**mean**
no	61.3	51.2	58.1	50.6	71.3	58.1	56.3	56.9	67.5	55.6	58.7
10Hz	55.6	55.6	53.1	45	66.9	59.4	52.5	52.5	65	53.8	51.2
60Hz	51.2	58.8	53.8	48.1	62.5	61.3	53.1	60.6	56.9	56.9	56.3
MA 64 Hz	60.6	51.9	53.1	51.2	76.3	59.4	60	60	57.5	60	59
MA 32 Hz	55.6	53.8	53.8	60.6	75.6	53.1	61.9	60.6	65	58.8	59.9
MA 21.3 Hz	60	55	50.6	62.5	58.1	53.8	53.1	53.8	63.7	45	55.6

Monopolar reference, 200–800 ms data segment, 4 DP features.

The bipolar channels that were computed from the monopolar channels improved the selection accuracy of 2 subjects (S2, S6). For all other subjects it did not gain better results [Tab. 3.8].

Table 3.8: Accuracy and referencing.

reference	**selection accuracy**										
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**mean**
monopolar	59.4	55.6	59.4	60.6	70	66.3	57.5	53.1	66.3	56.3	60.4
bipolar	55.6	61.3	44.4	53.8	46.3	73.8	57.5	41.9	53.8	**61.9**	55

4 DP points, 200–800 ms data segment, artefact removal.

Multiple LDA Classification

The multiple LDA classifier implements one Fisher's discriminant for every feature. The features are selected with a preceding discriminant power algorithm. A comparison of the different numbers of features showed that the best configuration is individual [Tab. 3.9]. Investigated data segment lengths unfolded no significant difference [Tab. 3.10].

Table 3.9: Accuracy and number of discriminant power features.

DP features	**selection accuracy**										
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**mean**
4	51.2	51.3	56.3	60.6	70	57.5	60.6	54.4	57.5	50	57.1
10	61.9	48.8	58.1	54.4	75	60.6	63.1	54.4	66.3	47.5	59
20	61.9	60	50	53.8	76.9	59.4	59.4	50.6	68.1	49.4	58.9

Monopolar reference, 200–1200 ms data segment, artefact removal.

Table 3.10: Accuracy and data segment length.

length [ms]	**selection accuracy**										
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**mean**
0–800	53.8	59.4	56.9	54.4	70.6	58.8	62.5	45.6	58.8	49.4	57
200–800	60	58.1	57.5	49.4	73.8	56.3	55	53.1	63.1	53.8	58
0–1200	54.4	46.9	43.8	51.9	56.3	57.5	50	41.9	62.5	47.5	51.3
200–1200	61.9	48.8	58.1	54.4	75	60.6	63.1	54.4	66.3	47.5	59

Monopolar reference, 10 DP features.

The effects of different filter settings were examined [Tab. 3.11]. A significance difference was found between LP 10 Hz–MA 32 Hz, LP 60 Hz–MA 32 Hz, MA 64 Hz–MA 32 Hz, and MA 32 Hz–MA 21.3 Hz. No setting outperformed the others tremendously.

Table 3.11: Accuracy and moving average filtering (MA).

filter	**selection accuracy**										
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**mean**
no	61.9	48.8	58.1	54.4	75	60.6	63.1	54.4	66.3	47.5	59
10 Hz	54.4	55.6	52.5	54.4	71.3	62.5	64.4	61.9	63.1	56.9	59.7
60 Hz	61.3	57.5	55.6	61.3	77.5	61.3	65	55	68.1	53.8	61.6
MA 64 Hz	65.6	58.1	52.5	46.9	71.3	50.6	56.3	44.4	65.6	52.5	56.4
MA 32 Hz	59.4	58.1	51.9	52.5	71.3	59.4	54.4	56.9	61.3	48.8	57.4
MA 21.3 Hz	62.5	59.4	45	58.1	63.1	51.2	74.4	53.8	63.1	43.8	57.4

Monopolar reference, 200–1200 ms data segment, 10 DP features.

Furthermore, the results of the classification with bipolar channels were investigated [Tab. 3.12]. Generally, the monopolar reference seemed to performe better; 2 subjects (S2, S6) reached a better selection accuracy with bipolar channels.

Table 3.12: Accuracy and referencing.

reference	**selection accuracy**										
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**mean**
monopolar	61.9	48.8	58.1	54.4	75	60.6	63.1	54.4	66.3	47.5	59
bipolar	48.8	63.7	49.4	43.1	56.9	72.5	58.8	46.3	55.6	58.1	55.3

200–1200 ms data segment, 10 DP features.

Best Classification

[Tab. 3.13] summarizes the best selection accuracies and settings of all subjects. The mean selection accuracy of all 10 subjects is 78.6 %. Nine users reached a selection accuracy of 70 % or more, subject 10 reached only 61.9 %. The 9 participants with a selection accuracy of at least 70 % achieved this performance with the SWLDA classifier. Subject 10 reached 61.9 % with Fisher's discriminant.

Table 3.13: Best accuracies and settings for all subjects.

S1	73.1	SWLDA monopolar reference, 200–800 ms data segment, max. features 10, moving average filter 32 Hz, artefact removal
S2	78.8	SWLDA monopolar reference, 200–800 ms data segment, max. features 10, low pass filter 60 Hz, artefact removal
S3	73.1	SWLDA bipolar reference, 200–800 ms data segment, max. features 10, low pass filter 10 Hz, artefact removal
S4	78.8	SWLDA bipolar reference, 200-800 ms data segment, max. features 10, low pass filter 10 Hz, artefact removal
S5	88.8	SWLDA monopolar reference, 0–800 ms data segment, max. features 4, low pass filter 10 Hz, artefact removal
S6	90.6	SWLDA monopolar reference, 0–800 ms data segment, max. features 4, low pass filter 10 Hz, artefact removal
S7	84.4	SWLDA monopolar reference, 200–800 ms data segment, max. features 10, low pass filter 10 Hz, artefact removal
S8	70	SWLDA monopolar reference, 200–800 ms data segment, max. features 10, moving average filter 21.3 Hz, artefact removal
S9	86.3	SWLDA monopolar reference, 200–800 ms data segment, max. features 10, artefact removal
S10	61.9	LDA bipolar reference, 200–800 ms data segment, 4 DP features, artefact removal

S1 to S10 denote subjects 1 to 10, the selection accuracy is denoted in %.

3.3 Theoretical Spelling Rate

A theoretical spelling rate is derived since no online tests were conducted to measure the real spelling rate.

The letters of the alphabet are built of 1 to 4 Morse symbols:
2 letters (E, T) with 1 Morse symbol
4 letters (A, I, M, N) with 2 Morse symbols
8 letters (D, G, K, O, R, S, U, W) with 3 Morse symbols
12 letters (B, C, F, H, J, L, P, Q, V, X, Y, Z) with 4 Morse symbols

A Morse symbol is selected in one trial, which lasts 33s. In addition to the mere Morse symbols for a letter, a space symbol is needed to end the Morse word. This raises the number of trials needed by one. A random pause of 1.5s to 2.5s lies between two trials; the average of 2s is used in the calculations. This leads to the following times to code a character:

$$t_{1S} = 2 \cdot t_{\text{trial}} + 1 \cdot t_{\text{pause}} = 2 \cdot 33s + 1 \cdot 2s = 68s$$

$$t_{2S} = 3 \cdot t_{\text{trial}} + 2 \cdot t_{\text{pause}} = 3 \cdot 33s + 2 \cdot 2s = 103s$$

$$t_{3S} = 4 \cdot t_{\text{trial}} + 3 \cdot t_{\text{pause}} = 4 \cdot 33s + 3 \cdot 2s = 138s$$

$$t_{4S} = 5 \cdot t_{\text{trial}} + 4 \cdot t_{\text{pause}} = 5 \cdot 33s + 4 \cdot 2s = 173s$$

where nS depicts a letter with n Morse symbols. The average time to select a letter (l) with respect to the letter frequency (p) in the German language [45] is:

$$t = p(l_{1S}) \cdot t_{1S} + p(l_{2S}) \cdot t_{2S} + p(l_{3S}) \cdot t_{3S} + p(l_{4S}) \cdot t_{4S} =$$

$$= 0.2355 \cdot t_{1S} + 0.2637 \cdot t_{2S} + 0.3232 \cdot t_{3S} + 0.1776 \cdot t_{2S} =$$

$$= 118.5s = 1\text{min } 59s \approx 2\text{min}$$

The best and worst case times for spelling a letter are:

$$t_{\text{best}} = 68s = 1\text{min } 8s$$

$$t_{\text{worst}} = 173s = 2\text{min } 58s$$

Chapter 4

Discussion

In this work, a BCI system based on auditory stimuli and the Morse code was proposed. The speller was demonstrated to work. In an offline study with 10 subjects, three classification methods and several different settings were investigated: (1) an SWLDA classification (2) a Fisher discriminant with preceding discriminant power algorithm, and (3) multiple LDA classifiers for every feature of the preceding discriminant power algorithm. The SWLDA classification was found to perform best. None of the different number of features, data segments, filtering settings, and referencing outperformed the other settings substantially.
Furdea et al. already showed that an ERP based auditory speller with an oddball paradigm worked for healthy participants [12]. Kanoh et al. demonstrated that a binary decision can be communicated with selective attention to one of two segregated tone streams [16]. This coincides with the findings here that a Morse speller based on auditory stream segregation and P300 detection works. A monopolar referenced channel set was used because in a study by Krusienski et. al no difference between monopolar reference (ear) and common average reference was found [20]. The data segment was set to at least 800 ms after the cue onset, since Furdea et al. found the average peak latencies for AEPs to be 575 ms [12].

Krusienski et al. compared classification methods for P300 potentials: Pearson's correlation method, Fisher's linear discriminant, SWLDA, and support vector machines. All algorithms were capable of classifying the data, but Fisher's linear discriminant and SWLDA provided the best results [19]. Thus, an SWLDA classifier, a single LDA and multiple LDAs were tested. The SWLDA classification has the additional advantage that it eliminates statistically not significant features for a large, unknown feature space [19]. Therefore, the LDA classifiers were combined with a preceding discriminant power algorithm, which aims to find the best classifiable features in the input space. The entry and removal criterion for the SWLDA were set to the usual $p<0.1$ and $p>0.15$, respectively. The setting for the number of maximum features are ambiguous; Krusienski et al. found the classification to be best for more than 45 maximum features [20], whereas Donchin et al. used 10 [6]. Thus 4, 10 and 60 maximum features for the SWLDA were tested.

The data segments used for the classification were only those after high deviant and low deviant tones. The first idea was to use the data segments of all tones (high, high deviant, low, low deviant) because the variance for the non target data is smaller, since there are more data segments to average. However, tests showed that the classification performance was better if only data from the deviant tones was used for the classifier.

The feasibility of the speller was shown in the experiment, the subjects had no problems to use the speller. The participants had no difficulties to distinguish the low and the high tone stream after a short instruction. However, most participants reported initial problems with hearing all deviant tones, in particular with the high ones. In the course of the experiment, all users reported to be able to distinguish the four tones better, usually after 20 to 30 spelled symbols. This matches the findings of a study by Klobassa et al., where the participants reported than the auditory paradigm required more attention than a visual one, and a training effect for the auditory P300 based paradigm was found [17]. The findings suggest that the Morse speller is easy to use after some training. In a real setting, an additional demand is that the Morse coding of letters is available. Thus, the Morse code needs either to be presented for support, or it needs to be learned by heart.

Waveform Analysis

A distinct waveform for the deviant stimuli was found in all subjects. Kanoh et al. already proved that deviant auditory stimuli in a BCI based on auditory stream segregation elicit a mismatch negativity (MMN) and a P300 deflection [16]. The MMN occurs typically 150–250 ms after the deviant stimulus, the P300 component

300 ms after the onset of a deviant event. The AEP response shows variability in the user performance. In some waveforms, a rhythmic pattern can be observed, e.g. in subject 6, while others lack it at all (subject 7). Klobassa et al. found the best P300 response to be usually at the location Pz [17]. This fact can be observed in subject 3 and subject 7. However, Schreuder et al. found the trend that the maximum difference of targets and non targets shifts to more frontal areas for faster stimulus rates [37]. This situation can be observed in subject 8.
The amplitude values were typically in the range of 3 to −3 µV, although the amplitude of subject 5 ranged from 3 to −6 µV and the amplitude of subject 10 ranged only within 1.5 to −1.5 µV.

Variation in the Data

A great variation in the data was found. One reason might be that the chosen time for the presentation of the auditory stimuli (30 s) was too long. It is hardly possible to avoid movement artefacts (e.g. eye blinks) for this period. Moreover, the break (1.5 s–2.5 s) between two trials might have been to short and did not give the participants enough time to relax. Another assumption is that the subjects did not perceive all deviant tones in the beginning of the test and therewith corrupted parts of the data. Furthermore, the participants reported the experiment to be demanding, all showed signs of tiredness after about half an hour of spelling. Thus, it can be assumed that due to the lack of concentration, the execution of the task was difficult.

Classification Performance

The SWLDA classification achieved the best results, followed by the multiple LDA method and Fisher's discriminant. Nine participants reached a selection accuracy greater than 70 %, one over 60 %. The theoretical chance level for the Morse speller is 25 % for a dot or dash and 50 % for a space symbol [Section 2.5.3]. The actual chance level for this experiment is 30.3 %.
Selection accuracies over 80 % were only achieved with the SWLDA classifier (4 subjects). Subject 1 to 9 reached their best results (over 70 %) with the SWLDA classifier, wherof for subject 3 and 4 bipolar channels were used. Subject 10's best accuracy was 61.9 % with Fisher's discriminant, the accuracy for the SWLDA classification was 53.8 %.

Interestingly, the selected features of the discriminant power algorithm were all found between 200 ms and 400 ms after the stimulus onset on channel Fz. The SWLDA on the contrary spreaded the significant features over the data segments and the channels.

Theoretical Spelling Rate

The average time for spelling a letter (in a German text) is 118.5 s or approximately 2 minutes. At first glance, this does not seem to be very fast. But a study found, when investigating a BCI for answering yes/no questions, that for involved people 1 to 5 minutes is no unreasonable time to wait for an answer [39]. So a spelling speed of 2 minutes per character might be sufficient.

Proposed Improvements

The paradigm was shown to work, but the variance in the data of the participants was found to be large. This set an upper limit to the classification score. To improve the performance of the speller, the following steps might be fruitful:

1. The Morse speller was adopted by all users in the experiments. The only drawback was the high deviant tone that was quite hard to hear in the beginning and therefore unsettled the participants. Thus, it is strongly recommended to make the high deviant tone "better audible". One suggestion is to enlarge the pitch difference of the tones. Kanoh et. al used in a BCI based on two segregated tone streams 1000 Hz (700 Hz) and 2000 Hz (2600 Hz) for the normal (deviant) tones [16]. In contrast, the Morse speller uses 880 Hz (784 Hz) and 1760 Hz (1568 Hz) for the normal (deviant) tones in the low and the high tone stream.

2. Sellers et al. found the best spelling results in a visual matrix speller with a 175 ms inter stimulus time in the oddball paradigm [40]. For now, the interstimulus time between a low and a high tone in the Morse speller is 160 ms with a tone duration of 40 ms. Thus, the interstimulus time in the high and in the low tone stream itself is 360 ms. A shortening of the interstimulus time might prove worthwhile and as well improve the bit rate.

3. Sellers et al. investigated matrix sizes and classification accuracies in a visual matrix speller and found that the 3x3 matrix yielded the best results [40]. A 3x3 matrix corresponds to an oddball paradigm with one target in 9 non targets (roughly 10 % targets). Based on this study, 10 % deviant stimuli were used in the Morse speller. An increase of deviant cues might bring an improvement, a rise to 20 % is suggested.

4. An increase in the number of stimulus sequences leads to an increase in the classification accuracy [39]. Thus the stimulus sequence (30s) might be shortened and more than one stimulus sequence presented for a Morse symbol. A shorter tone presentation facilitates the reduction of motional artefacts and a repeated stimulus sequence increases the classification accuracy.

5. The pause between two trials, which was set to 1.5s–2.5s, could be extended to give the participant more time to relax between the spelling of two Morse symbols and therewith soften the tension of the task.

6. For the visual speller matrix, the best channel set was found to be the classical locations (Fz, Cz, Pz) plus the posterior locations (PO7, PO8, Oz) [20]. Therefore, in an auditory paradigm a channel set including the classical locations and primary auditory regions could improve the system.

7. A high pass filter with a high cut-off frequency distorts the signal and reduces the amplitudes of lower frequency components such as the P300 and N400 deflections. Moreover, notch filters and online filtering also distort the ERP waveform [26]. Thus, reducing the high pass cut-off frequency to 0.01 Hz - 0.1 Hz, omitting the notch filter and keeping the online filtering to a minimum might be worth a try.

The future steps in the development of the Morse speller include the improvement of the classification score to increase the robustness of the speller, and online experiments.

Bibliography

[1] A. Bidet-Cauletand and O. Bertrand. Neurophysiological mechanisms involved in auditory perceptual organization. *Frontiers in Neuroscience*, 3, September 2009.

[2] N. Birbaumer, T. Elbert, A. Canavan, and B. Rockstroh. Slow potentials of the cerebral cortex and behavior. *Physiological Reviews*, 70:1–41, 1990.

[3] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor. A spelling device for the paralysed. *Nature*, 398, March 1999.

[4] B. Blankertz, G. Dornhege, M. Krauledat, M. Schröder, J. Williamson, R. Murray-Smith, and K. Müller. The Berlin brain-computer interface presents the novel mental typwriter Hex-O-Spell. In *Proceedings of the 3rd International Brain-Computer Interface Workshop and Training Course 2006*, 2006.

[5] M. G. Coles and M. D. Rugg. *Electrophysiology of mind: event-related brain potentials and cognition*. Oxford University Press, 1996.

[6] E. Donchin, K. M. Spencer, and R. Wijesinghe. The mental prosthesis: assessing the speed of a P300-based brain-computer interface. *IEEE Transactions on Rehabilitation Engineering*, 18, June 2000.

[7] G. Dornhege, J. del R. Millan, T. Hinterberger, D. J. McFarland, and K. Müller. *Toward brain-computer interfacing*. MIT Press, 2007.

[8] N. Draper and H. Smith. *Applied regression analysis*. Wiley-Interscience, 1998.

[9] L. A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, 70:510–523, 1988.

[10] P. Ferrez. Error-related EEG potentials in brain-computer interfaces. Master's thesis, École Polytechnique Fédérale de Lausanne, 2007.

[11] G. Friehs, R. Penn, M. Park, M. Goldman, V. Zerris, L. Hochberg, D. Chen, J. Mukand, and J. Donoghue. Initial surgical experience with an intracortical microelectrode array for brain-computer interface applications: 881. *Neurosurgery*, 59:481, 2006.

[12] A. Furdea, S. Halder, D. J. Krusienski, D. Boss, F. Nijboer, N. Birbaumer, and A. Kübler. An auditory oddball (P300) spelling system for brain-computer interfaces. *Psychophysiology*, 46:617–625, 2009.

[13] M. Hämäläinen, R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa. Magnetoencephalography - theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, 65(2):413–497, 1993.

[14] L. Hochberg, M. Serruya, G. Friehs, J. Mukand, M. Saleh, A. Caplan, A. Branner, D. Chen, R. Penn, and J. Donoghue. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature*, 442:164–171, 2006.

[15] H. Jasper. The ten-twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology*, 10:371–375, 1958.

[16] S. Kanoh, K. Miyamoto, and T. Yoshinobu. A brain-computer interface (BCI) system based on auditory stream segregation. In *Engineering in Medicine and Biology Society, 30th Annual International Conference of the IEEE*, pages 642–645, August 2008.

[17] D. S. Klobassa, T. M. Vaughan, P. Brunner, N. E. Schwartz, J. R. Wolpaw, C. Neuper, and E. W. Sellers. Toward a high-throughput auditory P300-based brain-computer interface. *Clinical Neurophysiology*, 120:1252–1261, 2009.

[18] J. Kronegg, S. Voloshynovskiy, and T. Pun. Analysis of bit-rate definitions for Brain-Computer Interfaces. *Electroencephalography and Clinical Neurophysiology*, 70:510–523, 1988.

[19] D. Krusienski, E. W. Sellers, F. Cabestaing, S. Bayoudh, D. Farland, T. Vaughan, and J. Wolpaw. A comparison of classification techniques for the P300 Speller. *Journal of Neural Engineering*, 3:299–305, 2006.

[20] D. Krusienski, E. W. Sellers, D. Farland, T. Vaughan, and J. Wolpaw. Toward enhanced P300 speller performance. *Journal of Neuroscience Methods*, 167:15–21, 2007.

[21] A. Kübler, A. Furdea, S. Halder, E. M. Hammer, F. Nijboer, and B. Kotchoubey. A brain-computer interface controlled auditory event-related potential (P300) spelling system for locked-in patients. *Annals of the New York Academy of Sciences, Disorder of Consciousness*, 1157:90–100, 2009.

[22] A. Kübler, N. Neumann, J. Kaiser, B. Kotchoubey, T. Hinterverger, and N. P. Birbaumer. Brain-computer communication: self-regulation of slow cortical potentials for verbal communication. *Archives of Physical Medicine and Rehabilitation*, 82:1533–1539, 2001.

[23] D. T. Larose. *Data mining methods and models*. John Wiley and Sons, 2006.

[24] E. C. Leuthardt, G. Schalk, J. R. Wolpaw, J. G. Ojemann, and D. W. Moran. A brain-computer interface using electrocorticographic signals in humans. *Journal of Neural Engineering*, 1:63–74, 2004.

[25] S. P. Levine, J. E. Huggins, S. L. BeMent, R. K. Kushwaha, L. A. Schuh, M. M. Rohde E. A. Passaro, D. A. Ross, K. V. Elisevich, and B. J. Smith. A direct brain interface based on event-related potentials. *Rehabilitation Engineering*, 8:180–185, 2000.

[26] S. J. Luck. *An introduction to the event-related potential technique*. The MIT Press, 2005.

[27] M. Middendorf, G. McMillan, G. Calhoun, and K. S. Jones. Brain-computer interfaces based on the steady-state visual-evoked response. *IEEE Transactions on Rehabilitation Engineering*, 8:211–214, 2000.

[28] G. Müller-Putz. Grundlagen der Hirn-Computer Kommunikation. Lecture notes, Graz University of Technology, 2008.

[29] G. Müller-Putz, R. Scherer, C. Neuper, and G. Pfurtscheller. Steady-state somatosensory evoked potentials: suitable brain signals for brain-computer interfaces? *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14:30–37, 2006.

[30] F. Nijboer, A. Furdea, I. Gunst, J. Mellinger, D. J. McFarland, N. Birbaumer, and A. Kübler. An auditory brain-computer interface (BCI). *Journal of Neuroscience Methods*, 167:43–50, 2008.

[31] F. Nijboer, E. W. Sellers, J. Mellinger, M. A. Jordan, T. Matuy, A. Furdea, S. Halder, U. Motchy, D. J. Krusienski, T. M. Vaughan, J. R. Wolpaw, N. Birbaumer, and A. Kübler. A P300-based brain-computer interface for people with amyotrophic lateral sclerosis. *Clinical Neurophysiology*, 119, March 2008.

[32] G. Pfurtscheller. Trends in der Neurorehabilitation. Lecture notes, Graz University of Technology, 2008.

[33] G. Pfurtscheller and F. H. Lopes da Silva. Event-related EEG/MEG synchronisation and desynchronisation: basic principles. *Clinical Neurophysiology*, 110:1842–1857, 1999.

[34] G. Pfurtscheller and F. H. Lopes da Silva. 15 years of BCI research at Graz University of Technology: current projects. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14:205–210, 2006.

[35] D. Regan. *Human brain electrophysiology: Evoked potentials and evoked magnetic fields in science and medicine*. Elsevier, 1989.

[36] S. Rezaei, K. Tavakolian, A. M. Nasrabadi, and S. K. Setarehdan. Different classification techniques considering brain computer interface applications. *Journal of Neural Engineering*, 3(2):139, 2006.

[37] E. M. Schreuder, M. Tangermann, and B. Blankertz. Initial results of a high-speed spatial auditory BCI. *International Journal of Bioelectromagnetism (IJBEM)*, 2009.

[38] S.Coyle, T. Ward, C. Markham, and G. McDarby. On the suitability of near-infrared (NIR) systems for next-generation braincomputer interfaces. *Physiological Measurement*, 25(4), 2004.

[39] E. W. Sellers and E. Donchin. A P300-based brein-computer interface: Inital test by ALS patients. *Clinical Neurophysiology*, 117:538–548, 2005.

[40] E. W. Sellers, D. Krusienski, D. Farland, T. Vaughan, and J. Wolpaw. A P300 event-related brain-computer interface (BCI): The effects of matrix size and interstimulus interval. *Biological Pschology*, 73:242–252, 2006.

[41] E. W. Sellers, A. Kübler, and E. Donchin. Brain-computer interface research at the Univeristy of South Florida Cognitive Psychophysiology Laboratory: The P300 speller. *IEEE Transactions on Rehabilitation Engineering*, 18, June 2006.

[42] L. Uldry, P. W. Ferrez, and J. del R. Millán. Feature selection methods on distributed linear inverse solutions for a non-invasive brain-machine interface. Master's thesis, École Polytechnique Fédérale de Lausanne, 2007.

[43] J. J. Vidal. Toward direct brain-computer communication. In *Annual Review of Biophysics and Bioengineering,*, volume 2, pages 157–180, 1973.

[44] N. Weiskopf, K. Mathiak, S. W. Bock, F. Scharnowski, R. Veit, W. Grodd, R. Goebel, and N. Birbaumer. Principles of a Brain-Computer Interface (BCI) based on real-time functional magnetic resonance imaging (fMRI). *IEEE Transactions on Biomedical Engineering*, 51(6), 2004.

[45] Wikipedia the free encyclopedia. http://en.wikipedia.org, 2010.

[46] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6):767–791, 2002.

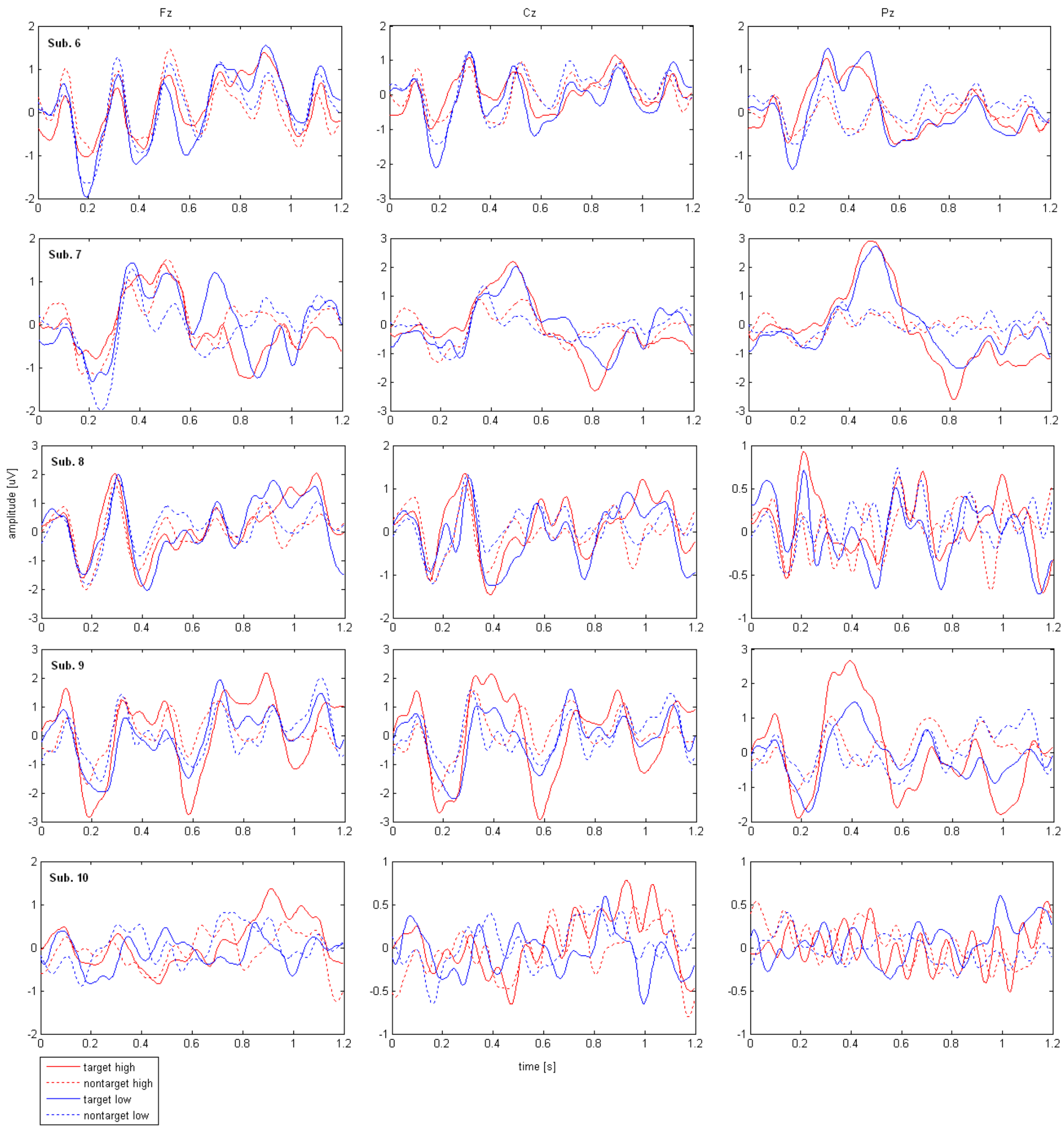
[47] J. R. Wolpaw, H. Ramoser, D. J. McFarland, and G. Pfurtscheller. EEG-based communication: improved accuracy by response verification. *IEEE Transactions on Rehabilitation Engineering*, 6(3):326–333, 1998.

Appendix A

Additional figures

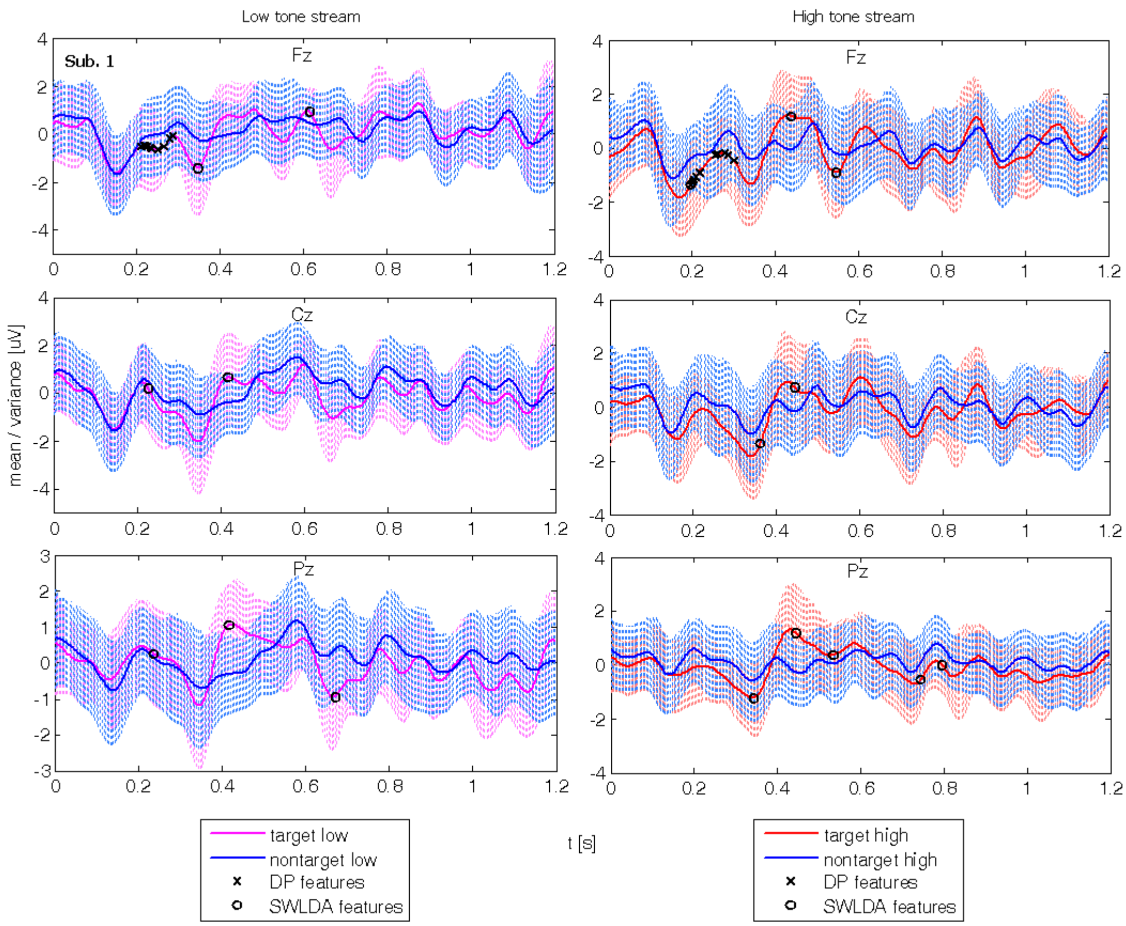
(a) Subject 1 - 5

Figure A.1: Grand average AEPs.
The grand average comprises data of 80 Morse symbol. The deviant target and non target tones in the high tone stream (red) and the low tone stream (blue) are shown. 8 deviant high and low tones are presented for one Morse symbol which are used for the classification; this gives 640 averages for every curve. The stimulus occurs from 0–40ms. The stimulus period is 200ms considering all tones and 400ms for tones within the tone stream.



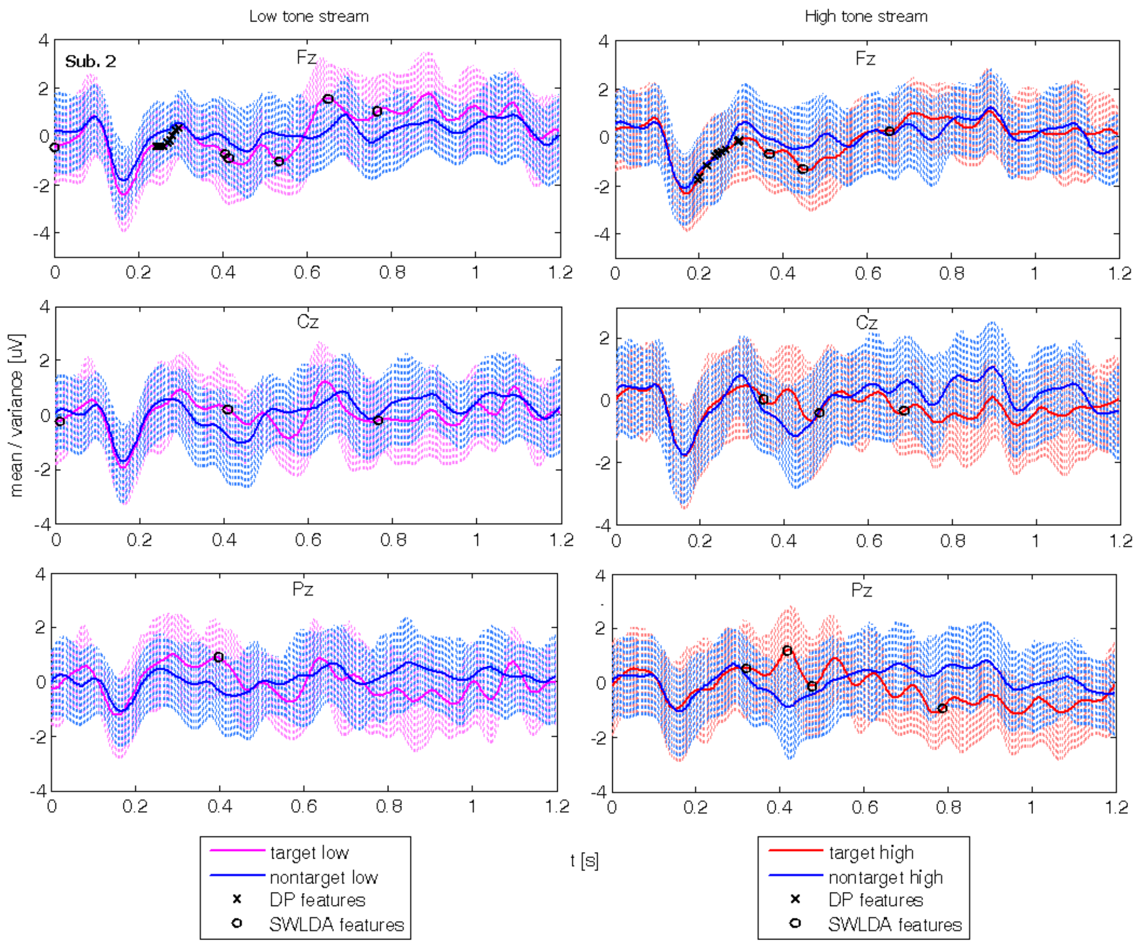
(b) Subject 6 - 10

Figure A.1: Grand average AEPs.
The grand average comprises data of 80 Morse symbol. The deviant target and non target tones in the high tone stream (red) and the low tone stream (blue) are shown. 8 deviant high and low tones are presented for one Morse symbol which are used for the classification; this gives 640 averages for every curve. The stimulus occurs from 0–40ms. The stimulus period is 200ms considering all tones and 400ms for tones within the tone stream.



(a) Subject 1

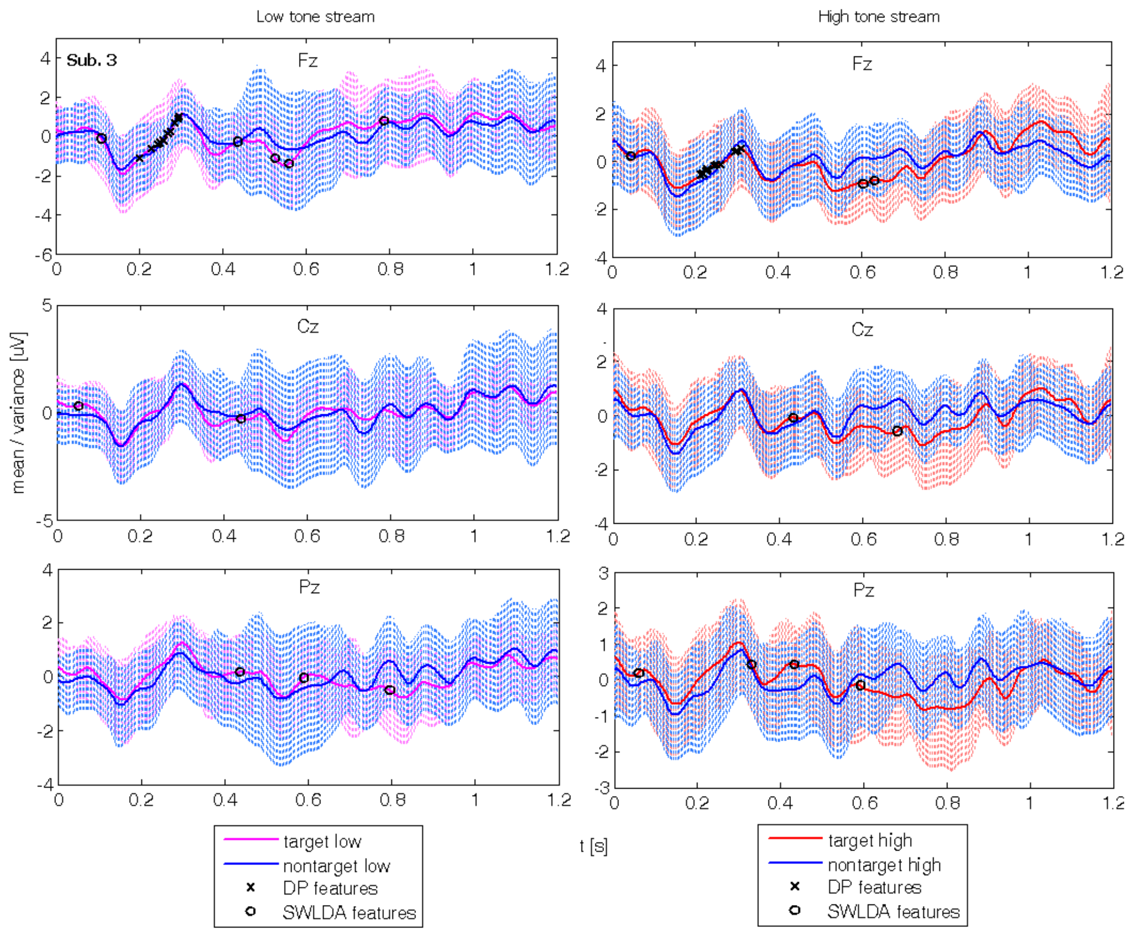
Figure A.2: Mean, variance and features.
The plot shows the mean (solid line) and the variance (dashed area) of target and non target responses in the low (left column) and high (right column) tone stream. Moreover, the features determined by the discriminant power algorithm (10 DP points) and the significant features of the SWLDA classification (10 maximum features) are plotted. The features are computed with 10-fold cross validation whereas the mean and the variance are computed of the overall data (80 Morse symbols x 8 deviant tones per tone stream = 640 averages).



(b) Subject 2

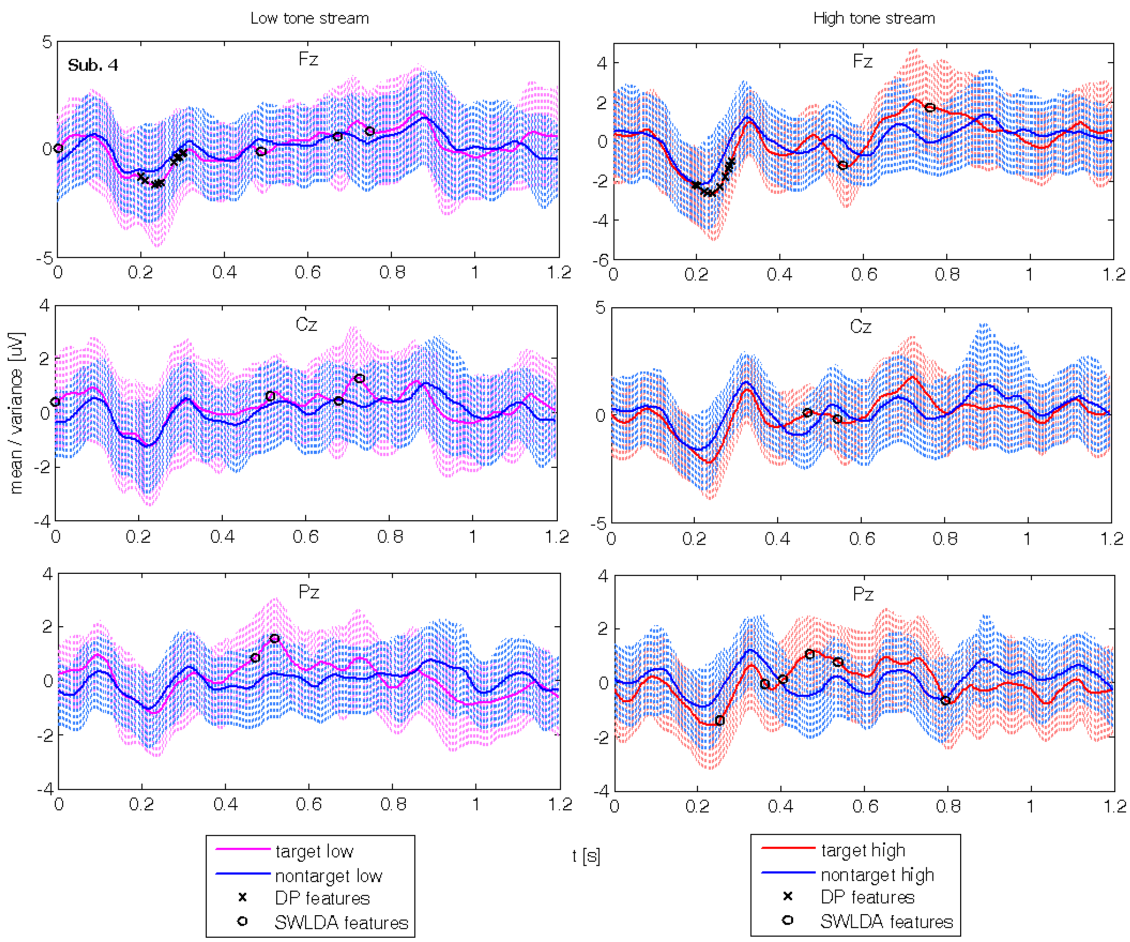
Figure A.2: Mean, variance and features.

The plot shows the mean (solid line) and the variance (dashed area) of target and non target responses in the low (left column) and high (right column) tone stream. Moreover, the features determined by the discriminant power algorithm (10 DP points) and the significant features of the SWLDA classification (10 maximum features) are plotted. The features are computed with 10-fold cross validation whereas the mean and the variance are computed of the overall data (80 Morse symbols x 8 deviant tones per tone stream = 640 averages).



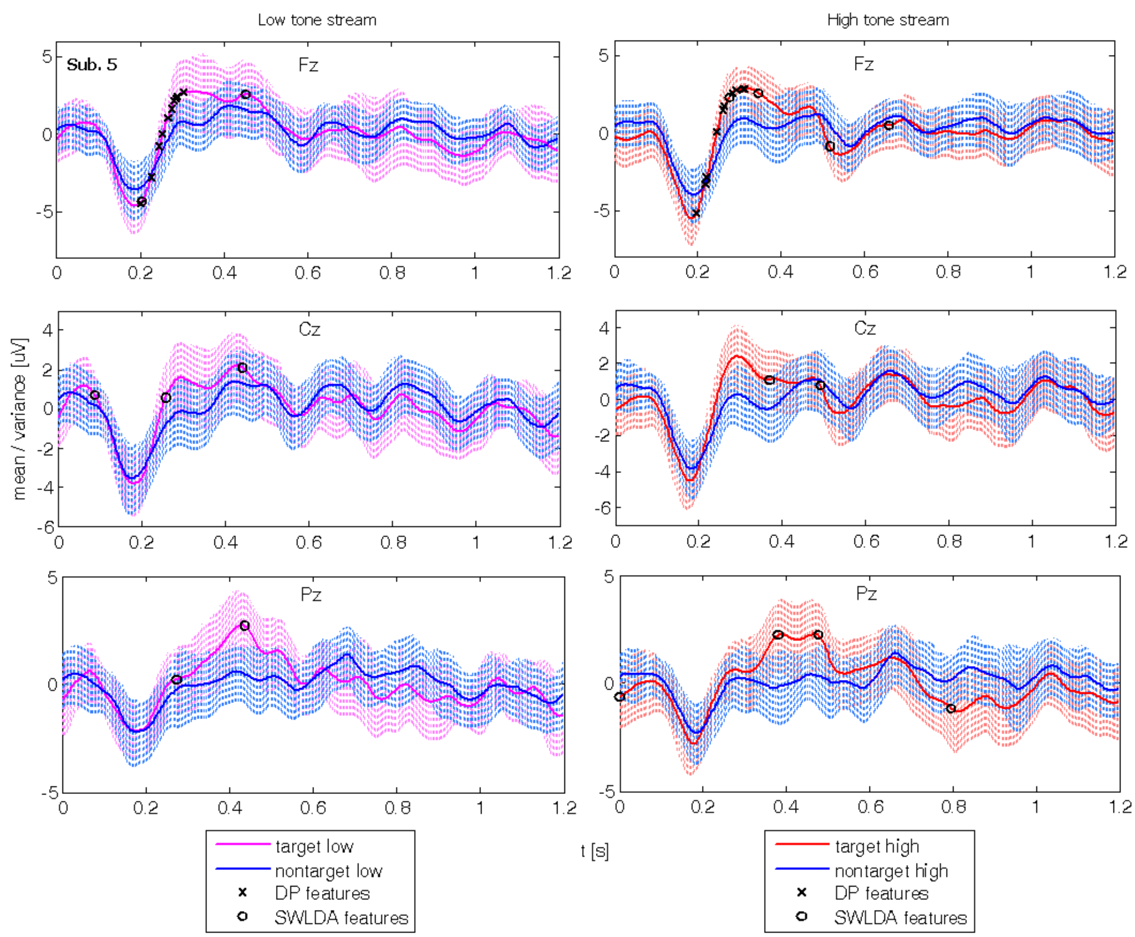
(c) Subject 3

Figure A.2: Mean, variance and features.
The plot shows the mean (solid line) and the variance (dashed area) of target and non target responses in the low (left column) and high (right column) tone stream. Moreover, the features determined by the discriminant power algorithm (10 DP points) and the significant features of the SWLDA classification (10 maximum features) are plotted. The features are computed with 10-fold cross validation whereas the mean and the variance are computed of the overall data (80 Morse symbols x 8 deviant tones per tone stream = 640 averages).



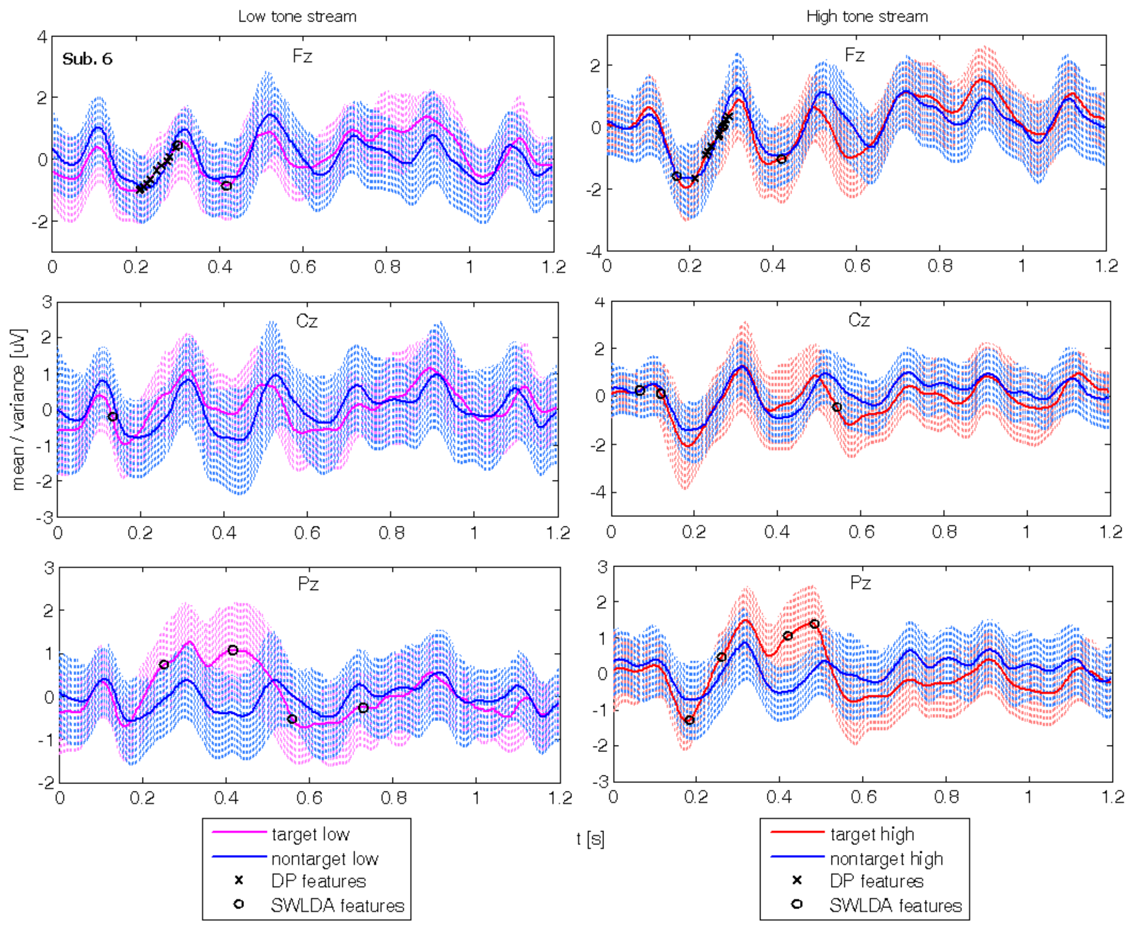
(d) Subject 4

Figure A.2: Mean, variance and features.
The plot shows the mean (solid line) and the variance (dashed area) of target and non target responses in the low (left column) and high (right column) tone stream. Moreover, the features determined by the discriminant power algorithm (10 DP points) and the significant features of the SWLDA classification (10 maximum features) are plotted. The features are computed with 10-fold cross validation whereas the mean and the variance are computed of the overall data (80 Morse symbols x 8 deviant tones per tone stream = 640 averages).



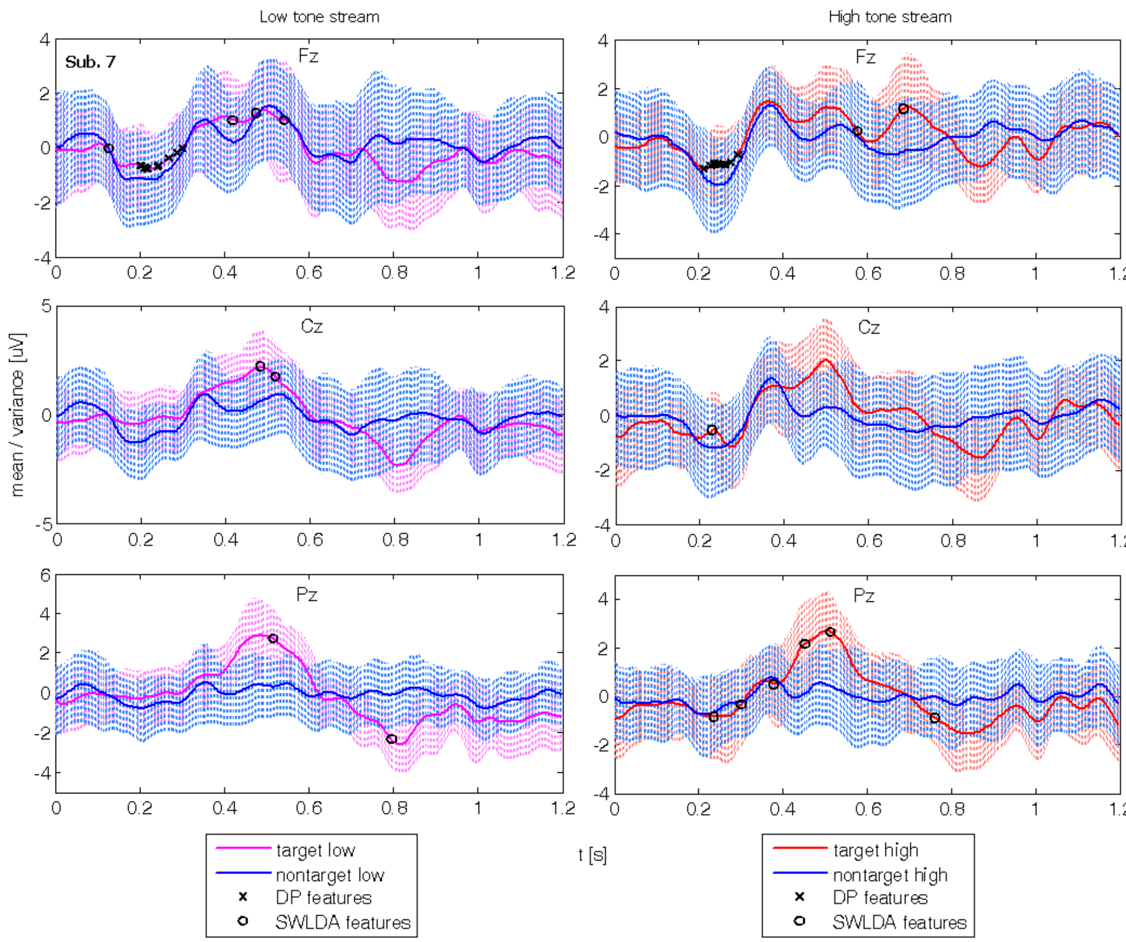
(e) Subject 5

Figure A.2: Mean, variance and features.
The plot shows the mean (solid line) and the variance (dashed area) of target and non target responses in the low (left column) and high (right column) tone stream. Moreover, the features determined by the discriminant power algorithm (10 DP points) and the significant features of the SWLDA classification (10 maximum features) are plotted. The features are computed with 10-fold cross validation whereas the mean and the variance are computed of the overall data (80 Morse symbols x 8 deviant tones per tone stream = 640 averages).



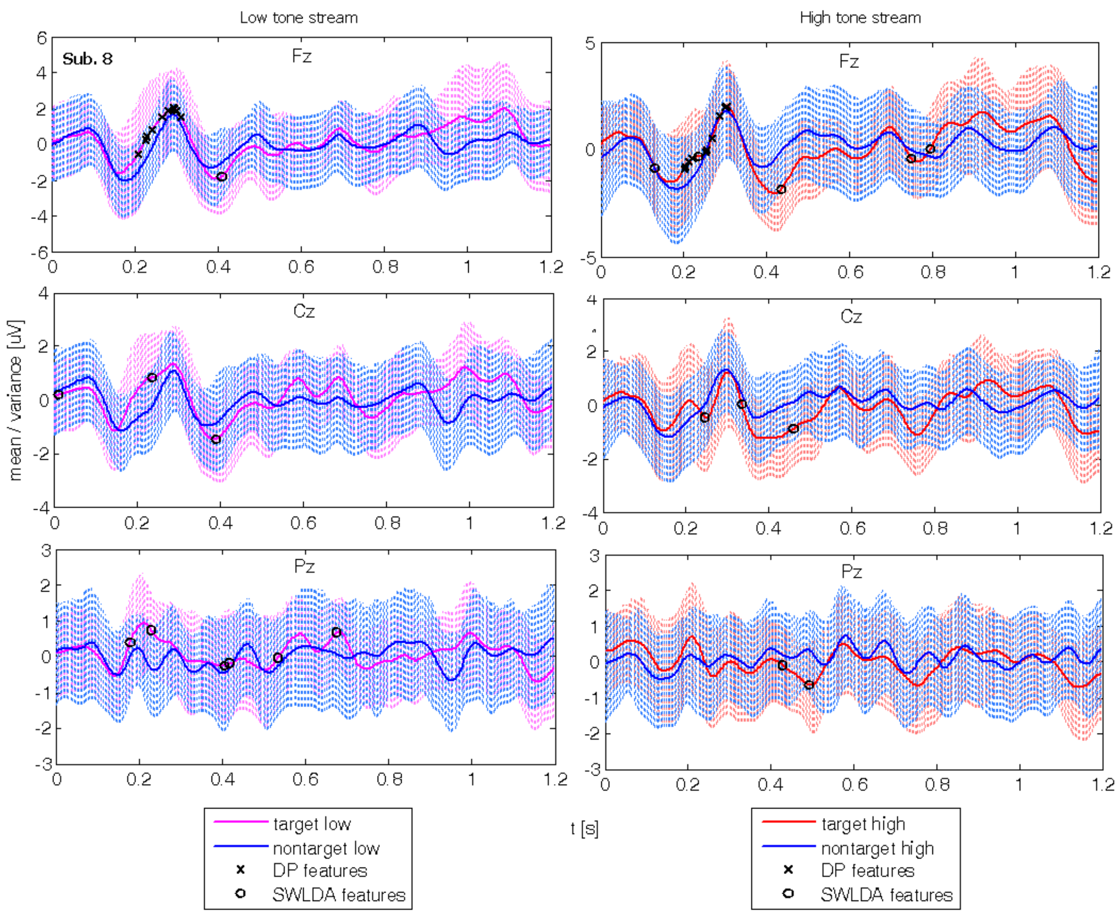
(f) Subject 6

Figure A.2: Mean, variance and features.
The plot shows the mean (solid line) and the variance (dashed area) of target and non target responses in the low (left column) and high (right column) tone stream. Moreover, the features determined by the discriminant power algorithm (10 DP points) and the significant features of the SWLDA classification (10 maximum features) are plotted. The features are computed with 10-fold cross validation whereas the mean and the variance are computed of the overall data (80 Morse symbols x 8 deviant tones per tone stream = 640 averages).



(g) Subject 7

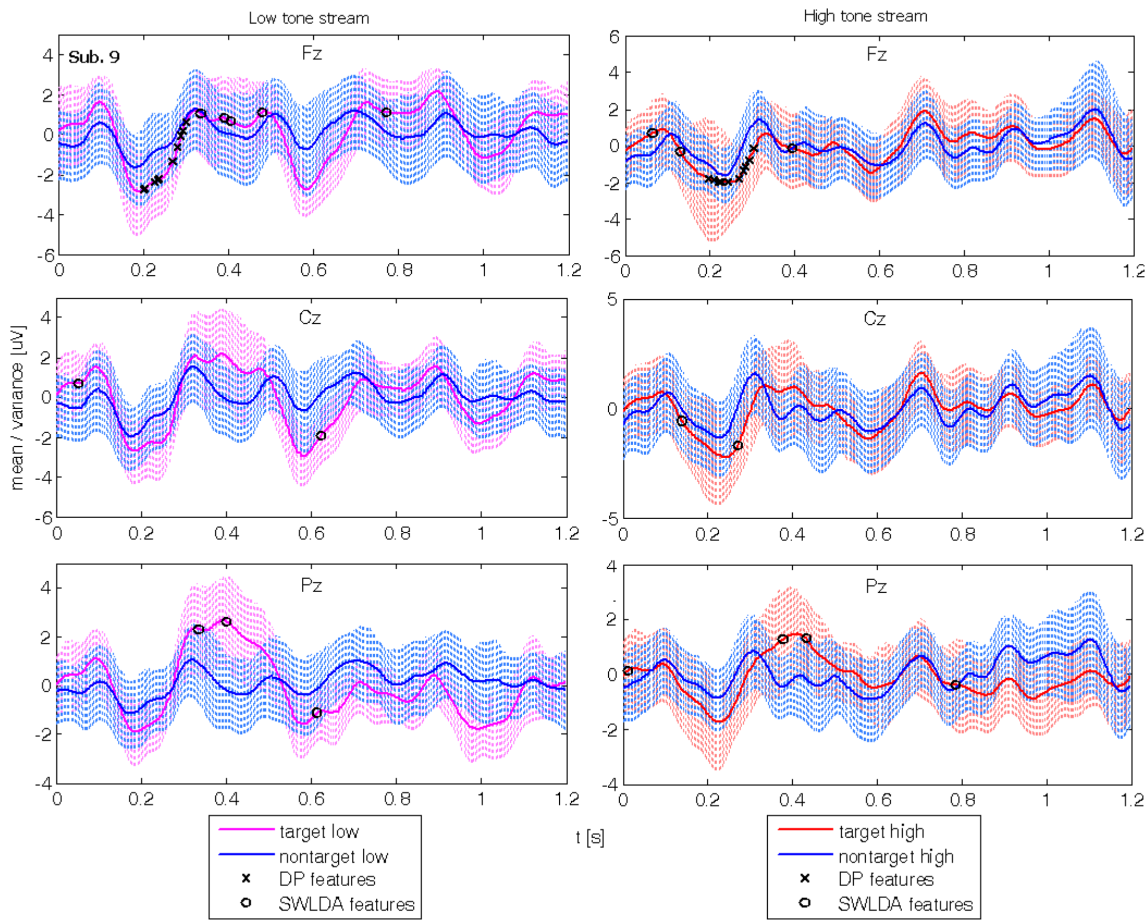
Figure A.2: Mean, variance and features.
The plot shows the mean (solid line) and the variance (dashed area) of target and non target responses in the low (left column) and high (right column) tone stream. Moreover, the features determined by the discriminant power algorithm (10 DP points) and the significant features of the SWLDA classification (10 maximum features) are plotted. The features are computed with 10-fold cross validation whereas the mean and the variance are computed of the overall data (80 Morse symbols x 8 deviant tones per tone stream = 640 averages).



(h) Subject 8

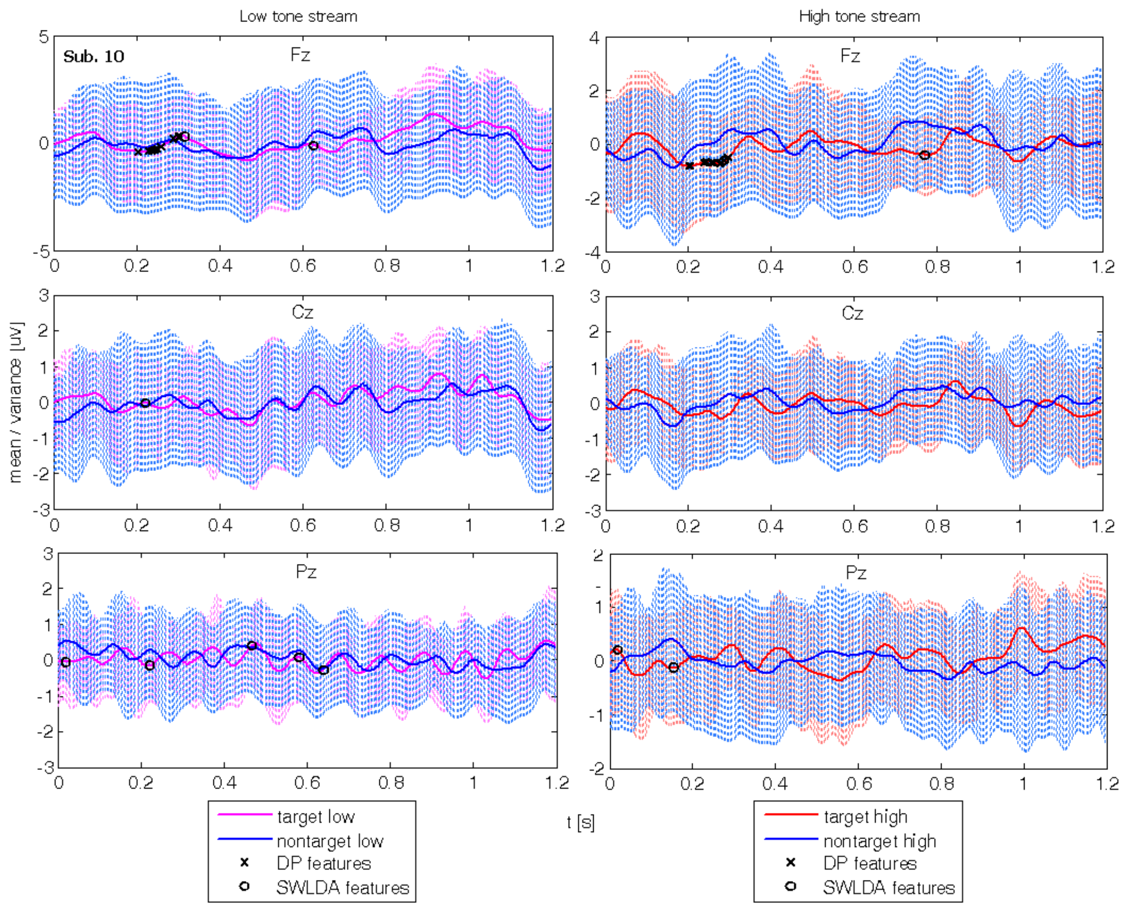
Figure A.2: Mean, variance and features.

The plot shows the mean (solid line) and the variance (dashed area) of target and non target responses in the low (left column) and high (right column) tone stream. Moreover, the features determined by the discriminant power algorithm (10 DP points) and the significant features of the SWLDA classification (10 maximum features) are plotted. The features are computed with 10-fold cross validation whereas the mean and the variance are computed of the overall data (80 Morse symbols x 8 deviant tones per tone stream = 640 averages).



(i) Subject 9

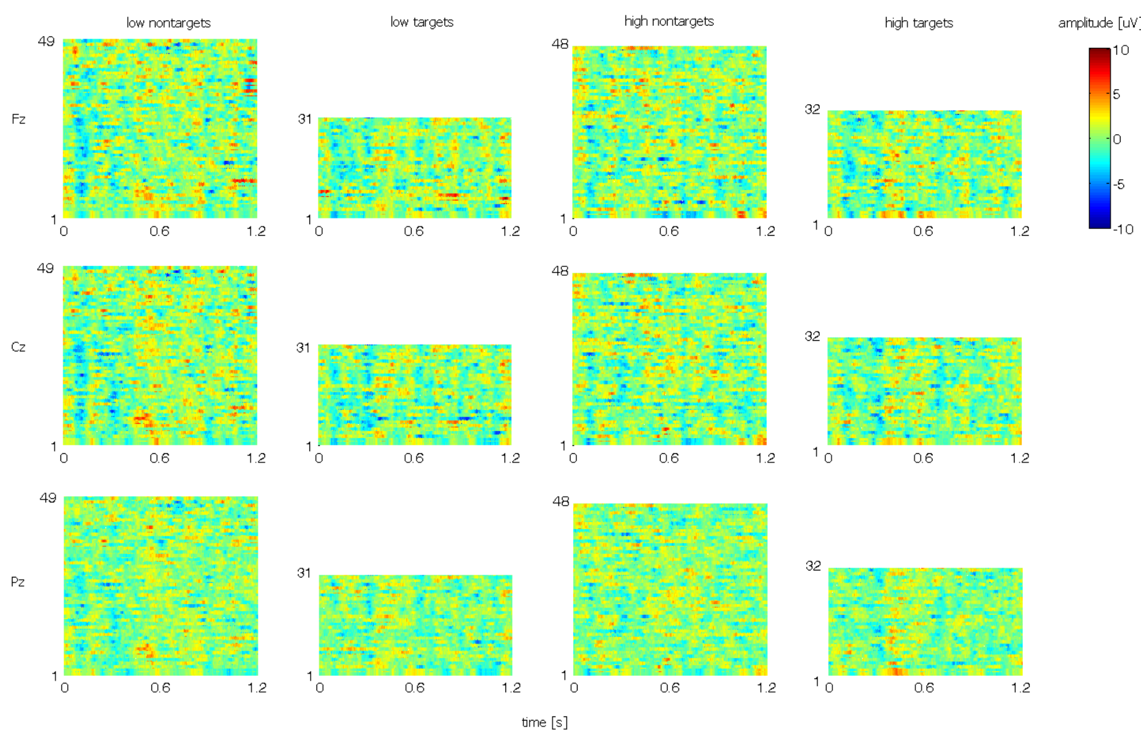
Figure A.2: Mean, variance and features.
The plot shows the mean (solid line) and the variance (dashed area) of target and non target responses in the low (left column) and high (right column) tone stream. Moreover, the features determined by the discriminant power algorithm (10 DP points) and the significant features of the SWLDA classification (10 maximum features) are plotted. The features are computed with 10-fold cross validation whereas the mean and the variance are computed of the overall data (80 Morse symbols x 8 deviant tones per tone stream = 640 averages).



(j) Subject 10

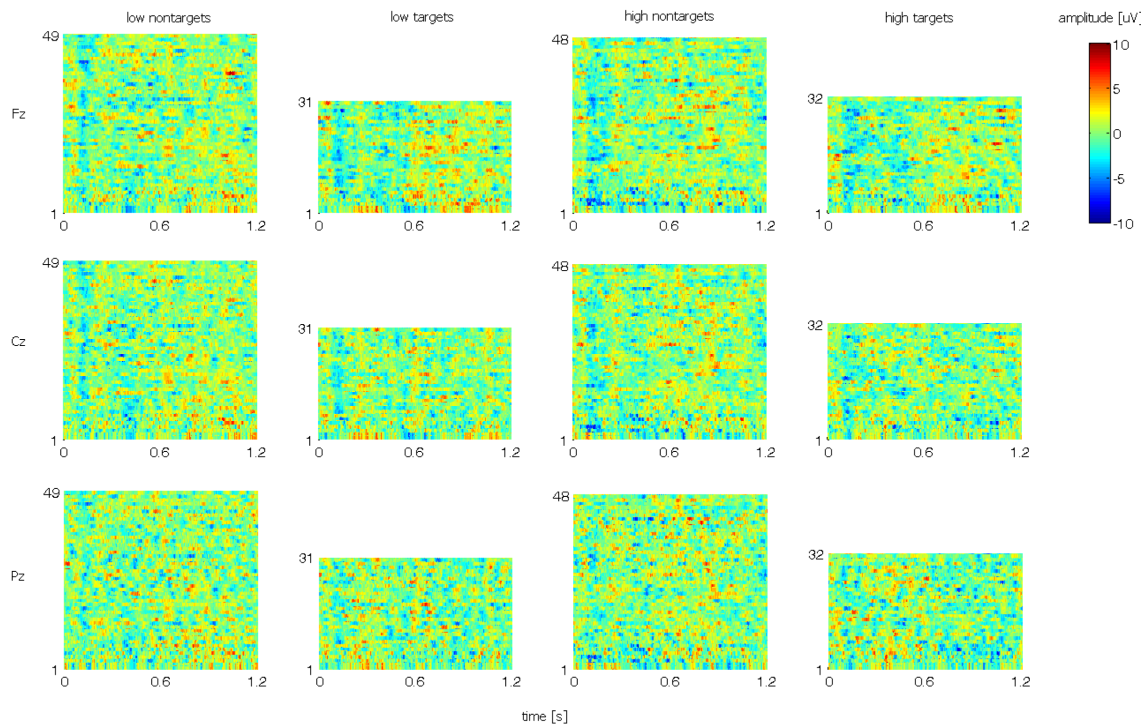
Figure A.2: Mean, variance and features.

The plot shows the mean (solid line) and the variance (dashed area) of target and non target responses in the low (left column) and high (right column) tone stream. Moreover, the features determined by the discriminant power algorithm (10 DP points) and the significant features of the SWLDA classification (10 maximum features) are plotted. The features are computed with 10-fold cross validation whereas the mean and the variance are computed of the overall data (80 Morse symbols x 8 deviant tones per tone stream = 640 averages).



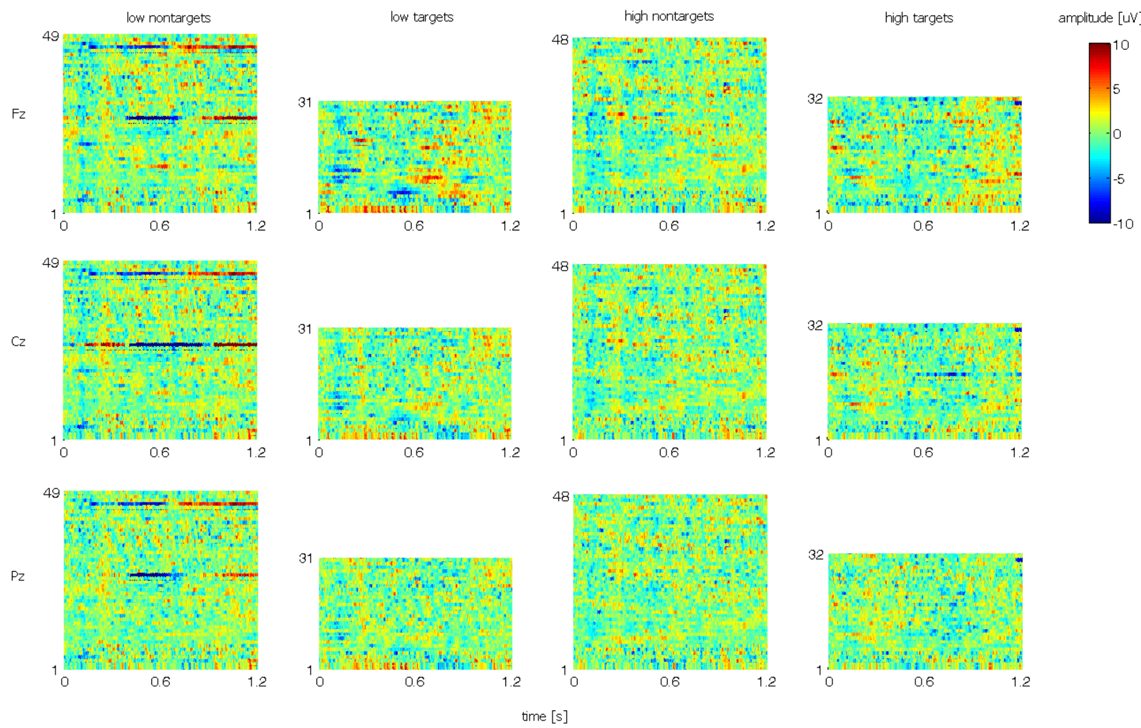
(a) Subject 1

Figure A.3: AEPs with respect to the trial and time.
The average amplitude for every trial (8 averages) is depicted color coded. The trials are plotted vertically against a 1200ms data segment (horizontal). The AEPs are shown for the low and the high tone stream and the locations Fz, Cz and Pz.



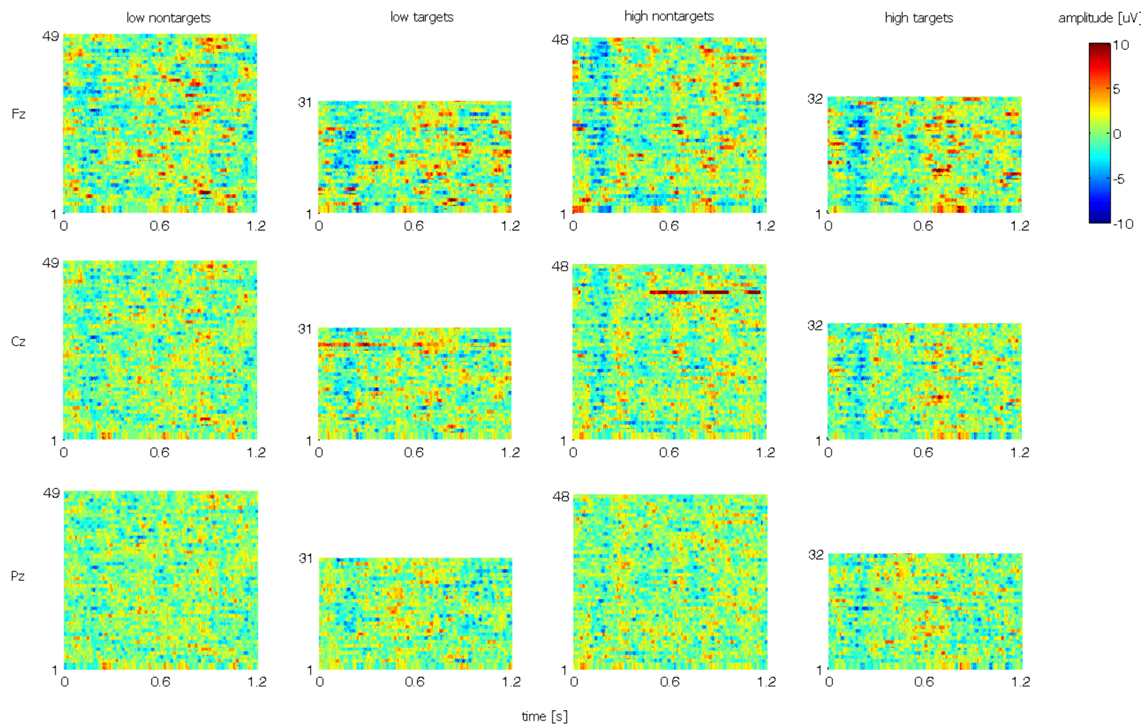
(b) Subject 2

Figure A.3: AEPs with respect to the trial and time.
The average amplitude for every trial (8 averages) is depicted color coded. The trials are plotted vertically against a 1200ms data segment (horizontal). The AEPs are shown for the low and the high tone stream and the locations Fz, Cz and Pz.



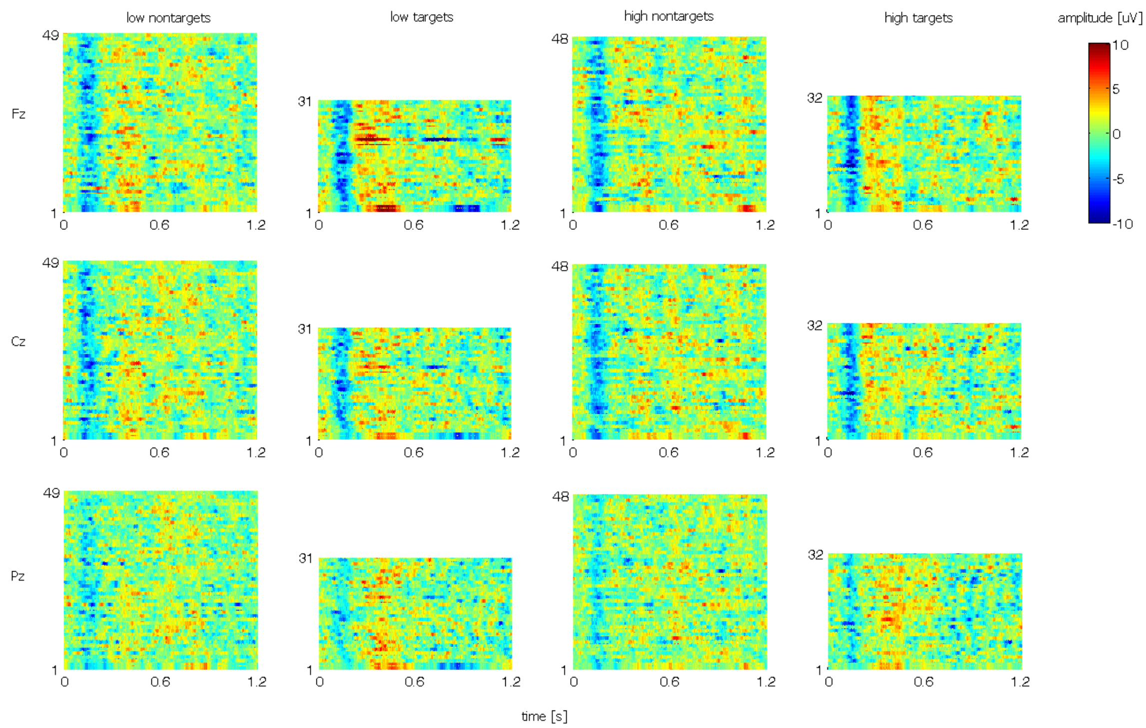
(c) Subject 3

Figure A.3: AEPs with respect to the trial and time.
The average amplitude for every trial (8 averages) is depicted color coded. The trials are plotted vertically against a 1200ms data segment (horizontal). The AEPs are shown for the low and the high tone stream and the locations Fz, Cz and Pz.



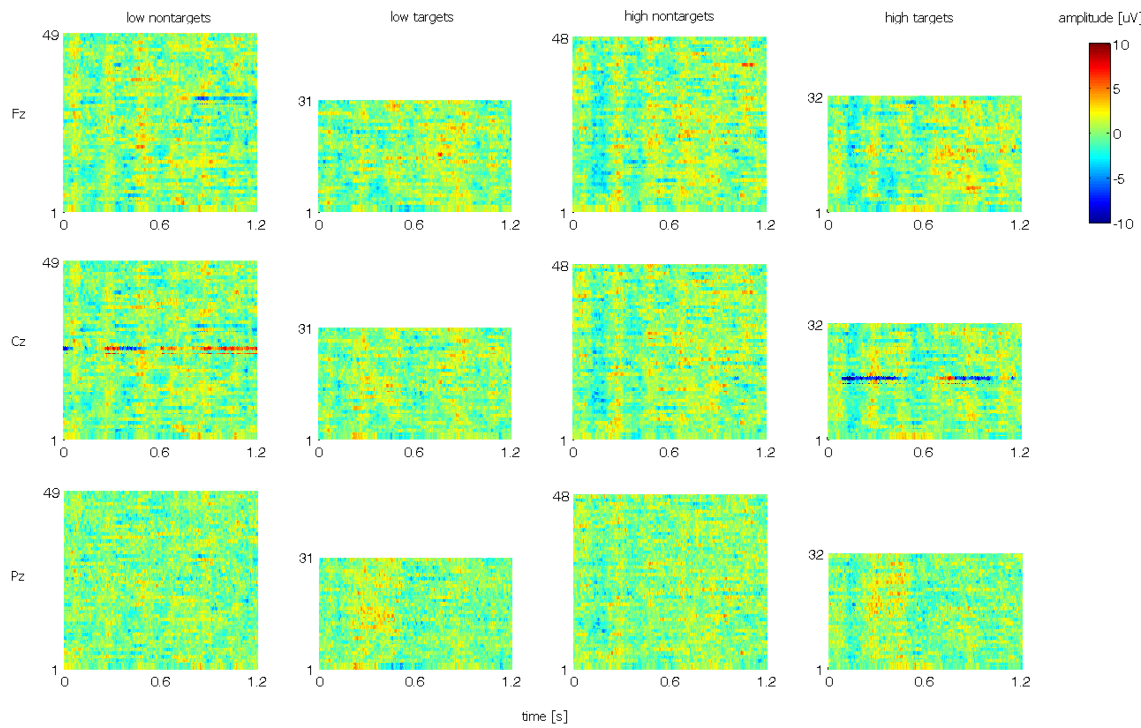
(d) Subject 4

Figure A.3: AEPs with respect to the trial and time.
The average amplitude for every trial (8 averages) is depicted color coded. The trials are plotted vertically against a 1200ms data segment (horizontal). The AEPs are shown for the low and the high tone stream and the locations Fz, Cz and Pz.



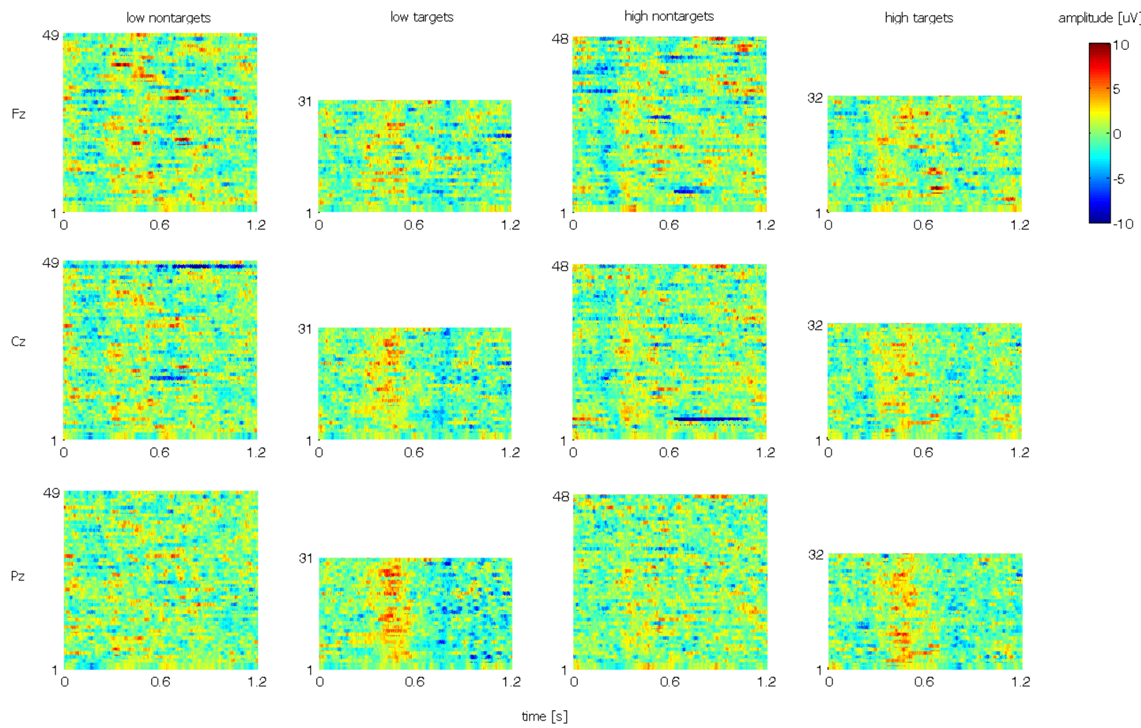
(e) Subject 5

Figure A.3: AEPs with respect to the trial and time.
The average amplitude for every trial (8 averages) is depicted color coded. The trials are plotted vertically against a 1200ms data segment (horizontal). The AEPs are shown for the low and the high tone stream and the locations Fz, Cz and Pz.



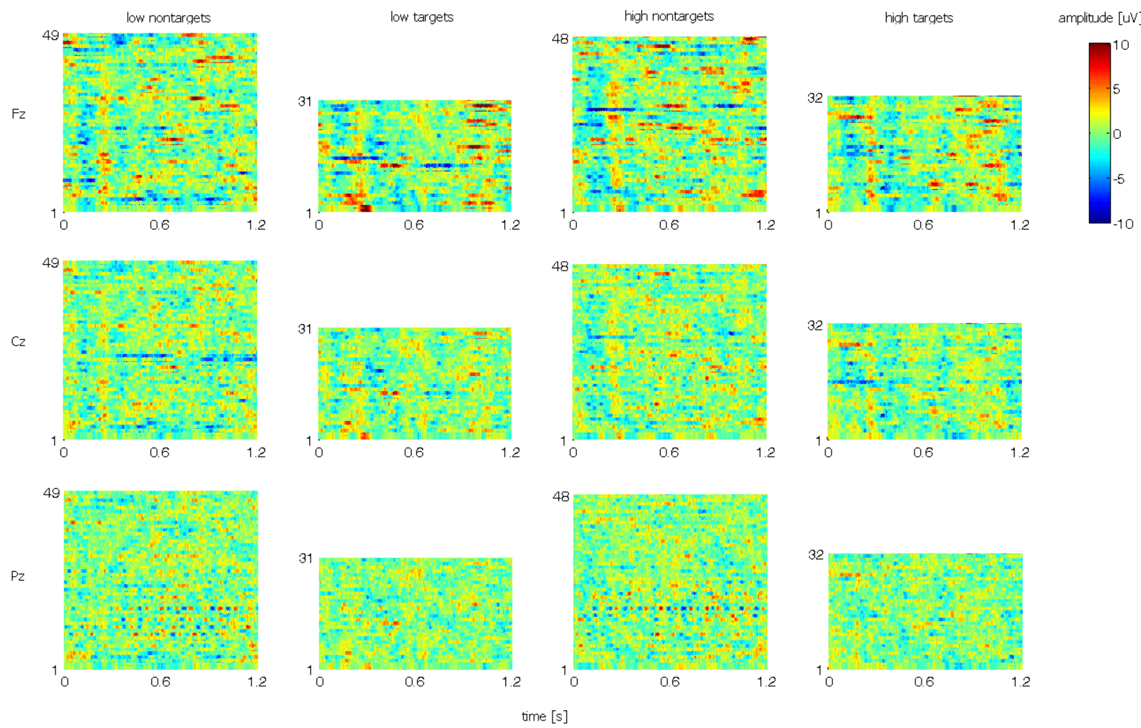
(f) Subject 6

Figure A.3: AEPs with respect to the trial and time.
The average amplitude for every trial (8 averages) is depicted color coded. The trials are plotted vertically against a 1200ms data segment (horizontal). The AEPs are shown for the low and the high tone stream and the locations Fz, Cz and Pz.



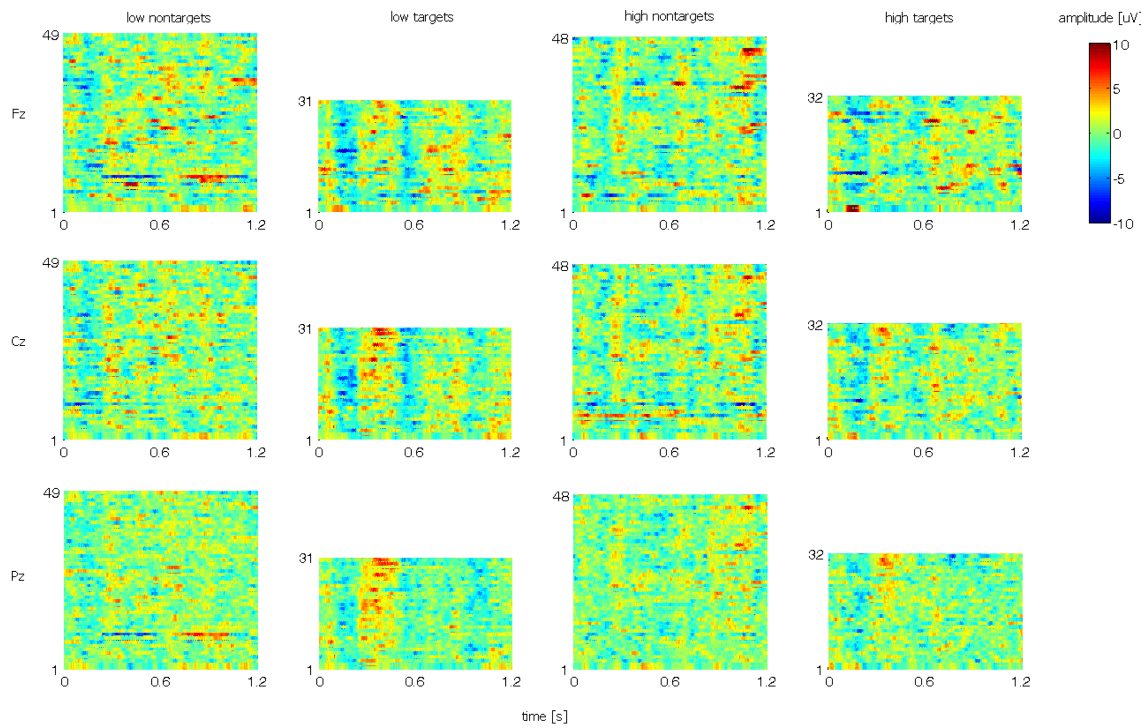
(g) Subject 7

Figure A.3: AEPs with respect to the trial and time.
The average amplitude for every trial (8 averages) is depicted color coded. The trials are plotted vertically against a 1200ms data segment (horizontal). The AEPs are shown for the low and the high tone stream and the locations Fz, Cz and Pz.



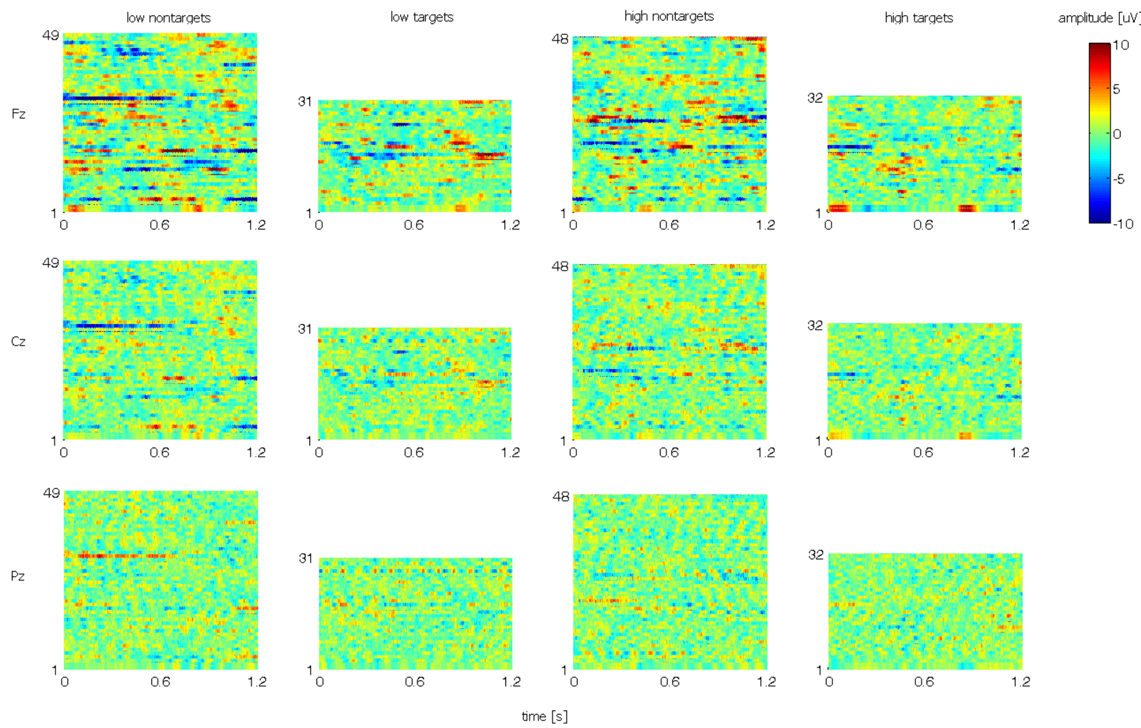
(h) Subject 8

Figure A.3: AEPs with respect to the trial and time.
The average amplitude for every trial (8 averages) is depicted color coded. The trials are plotted vertically against a 1200ms data segment (horizontal). The AEPs are shown for the low and the high tone stream and the locations Fz, Cz and Pz.



(i) Subject 9

Figure A.3: AEPs with respect to the trial and time.
The average amplitude for every trial (8 averages) is depicted color coded. The trials are plotted vertically against a 1200ms data segment (horizontal). The AEPs are shown for the low and the high tone stream and the locations Fz, Cz and Pz.



(j) Subject 10

Figure A.3: AEPs with respect to the trial and time.
The average amplitude for every trial (8 averages) is depicted color coded. The trials are plotted vertically against a 1200ms data segment (horizontal). The AEPs are shown for the low and the high tone stream and the locations Fz, Cz and Pz.