Alma Mater Studiorum
Università degli studi di Bologna
Dottorato di ricerca in Bioingegneria
Ciclo XXVI
Settore Concorsuale: 09/G2
Settore Scientifico Disciplinare di afferenza: Ing-Inf/06

BRAIN-COMPUTER INTERFACES FOR AUGMENTED COMMUNICATION: ASYNCHRONOUS AND ADAPTIVE ALGORITHMS AND EVALUATION WITH END USERS

Francesca Schettini

Relatore:
Prof. ssa Serenella Salinari

Revisori:
Prof. Mauro Ursino
Prof. Reinhold Scherer

Co-relatore:
Prof. Febo Cincotti

Coordinatore Dottorato:
Prof.ssa Elisa Magosso

ESAME FINALE ANNO 2014

A Fabio e alla mia famiglia…

ABSTRACT

This PhD thesis aimed at addressing some of the issues that, at the state of the art, avoid the P300-based brain computer interface (BCI) systems to move from research laboratories to end users' home. An innovative asynchronous classifier has been defined and validated. It relies on the introduction of a set of thresholds in the classifier, and such thresholds have been assessed considering the distributions of score values relating to target, non-target stimuli and epochs of voluntary no-control. With the asynchronous classifier, a P300-based BCI system can adapt its speed to the current state of the user and can automatically suspend the control when the user diverts his attention from the stimulation interface. Since EEG signals are non-stationary and show inherent variability, in order to make long-term use of BCI possible, it is important to track changes in ongoing EEG activity and to adapt BCI model parameters accordingly. To this aim, the asynchronous classifier has been subsequently improved by introducing a self-calibration algorithm for the continuous and unsupervised recalibration of the subjective control parameters. Finally an index for the online monitoring of the EEG quality has been defined and validated in order to detect potential problems and system failures. This thesis ends with the description of a translational work involving end users (people with amyotrophic lateral sclerosis - ALS). Focusing on the concepts of the user centered design approach, the phases relating to the design, the development and the validation of an innovative assistive device have been described. The proposed assistive technology (AT) has been specifically designed to meet the needs of people with ALS during the different phases of the disease (i.e. the degree of motor abilities impairment). Indeed, the AT can be accessed with several input devices either conventional (mouse, keyboard, touch screen) or alterative (switches, headtracker) up to an embedded P300-based BCI.

1 INTRODUCTION

Abilities such as read minds, communicate without speaking, move objects with the power of thought, have always fascinated people and are often reported in science fictions and novels. Brain Computer Interface (BCI) systems can provide some of these features by detecting and processing specific cerebral potentials in order to convert them into control signals for an external device (computers, prosthesis, domotic appliances, etc.). Since they do not rely on muscles and peripheral nerves, the target users for BCI systems are mainly people with severe motor disability, such as people with amyotrophic lateral sclerosis (ALS), muscular sclerosis, and outcomes of accidents. Recently new applications have been also suggested for healthy people such as games or monitoring of psychophysiological factors (stress, attention level, fatigue) during demanding and critical tasks for specific professional figures (e.g. surgeons, pilots, etc.). However the state of the art of the BCI systems is considerably far from what has been imagined in science fictions. In fact, until now, BCIs are laboratory prototypes, and despite the great progresses of the last 25 years, they still suffer from problems of speed, reliability and overall usability. Furthermore the available systems still need complex procedures for configuration and calibration requiring a specialized technical staff for their set up operations

Among non-invasive EEG-based BCIs, the P300 event related potential (ERP) has been identified as one of the most effective control feature for applications concerning communication and environmental control. In fact, P300-based BCIs allow users to select an item (a character, a word, an icon) between several possible choices in a relative short time and without user training. Since the P300 potential is a cognitive potential that can be influenced by several factors, such as subject age or sex, environmental conditions, stimulation modalities, a short calibration period (a 10/20 minutes) is required in order to identify the specific user’s control parameters. Moreover the effectiveness of the control parameters can be influenced by the variations of the morphology of the P300 potential over time.

The main objective of this PhD thesis has been to identify some of the weak points of the current P300-based BCIs that prevent these systems to move from labs to the users' home and suggest solutions to improve their overall usability.

The first section provides the basic concepts of neurophysiology useful to understanding the operation modalities and problems of P300-based BCI systems. The mechanisms and the key concepts regarding event-related potentials will be described with a particular attention to the P300 ERP. Brain computer interfaces will be defined providing a brief description of the main applications available and focusing on the translation of these systems to the assistive technologies field.

The second section addresses the problem of reducing the existing gap between BCI systems and conventional assistive devices. A new classification algorithm will be described and validated both with healthy subjects and potential end users. The proposed algorithm aims at increasing the usability of the P300-based BCI systems in non-experimental condition by providing two important features: *i)* dynamic-stopping, i.e. the system can automatically adapt the speed of selection depending on the current state of the user; *ii)* voluntary no-control period detection, i.e. the system is able to understand when the user diverts his attention from the stimulation interface and to automatically suspend the selections on the interface.

The third section addresses the problem of the system calibration over time. First, the benefits of a continuous recalibration of the system will be pointed out with respect to a single initial calibration, and then a self-calibration algorithm will be defined and validated. Such algorithm was introduced into the asynchronous classifier and allows to label data collected in unsupervised manner and to continuously update the control parameters of the system.

The forth section investigates the effects of the artifacts rejection on the performance of BCI systems. Different types of artifacts have been taken into account, and an index for the online assessment of the quality of the EEG signal has been defined and validated.

The last section regards a practical and innovative application. In fact it describes an assistive device designed to be accessed by means of several input devices, including a P300-based BCI. This was done in order to copy with to the residual motor abilities of the users with amyotrophic lateral sclerosis (ALS) during the different stages of the disease. The preliminary evaluation of this device was conducted in terms of overall usability of the system involving subjects with ALS.

2 PRELIMINARY CONCEPTS

2.1 Neurophysiological basis: the nervous system

The nervous system is responsible for the detection and elaboration of both external and internal stimuli to the body. Complex mental functions, such as memory, learning and emotions are associated with it. It can be divided in two main sections: the somatic and autonomic nervous systems. The somatic peripheral nervous system manages information from receptors for pain, temperature, and mechanical stimuli in the skin, muscles, and joints to the central nervous system, and the motor neurons, which return impulses from the central nervous system to these same areas of the body. The autonomic nervous system is concerned with the involuntary regulation of smooth muscle, cardiac muscle, and glands. The central nervous system (CNS) consists of the encephalon, enclosed in the cranium, and the spinal cord which is contained in the spinal canal. The CNS is responsible for the integration, analysis and coordination of sensory data, but as well of motor commands. It is also the site of very important functions such as intelligence, memory, learning and emotion. In contrast to the peripheral nervous system, the CNS is not only able to collect and transmit information, but is also able to integrate them. The peripheral nervous system consists of the whole nervous tissue outside the CNS and performs the function of transmission, through bundles of conduction of afferent signals from a peripheral unit (organs) or outgoing (efferent) towards a peripheral unit.

The brain is a large soft mass of nervous tissue and has three major parts: cerebrum, diencephalon, and brain stem and cerebellum. The cerebrum is divided into two hemispheres and is the largest and most observable portion of the brain. In fact it consists of many convoluted ridges (gyri), narrow grooves (sulci), and deep fissures. The cerebral cortex consist of the outer layer of the cerebrum and it is composed of gray matter (neurons with unmyelinated axons) that is 2-4 mm thick and contains over 50 billion neurons and 250 billion glial cells called neuroglia. The thicker inner layer is the white matter that consists of interconnecting groups of myelinated axons that project from the cortex to other cortical areas or from the thalamus (part of the diencephalon) to the cortex. The connection between the two cerebral hemispheres is called the corpus callosum (Figure 2.1).

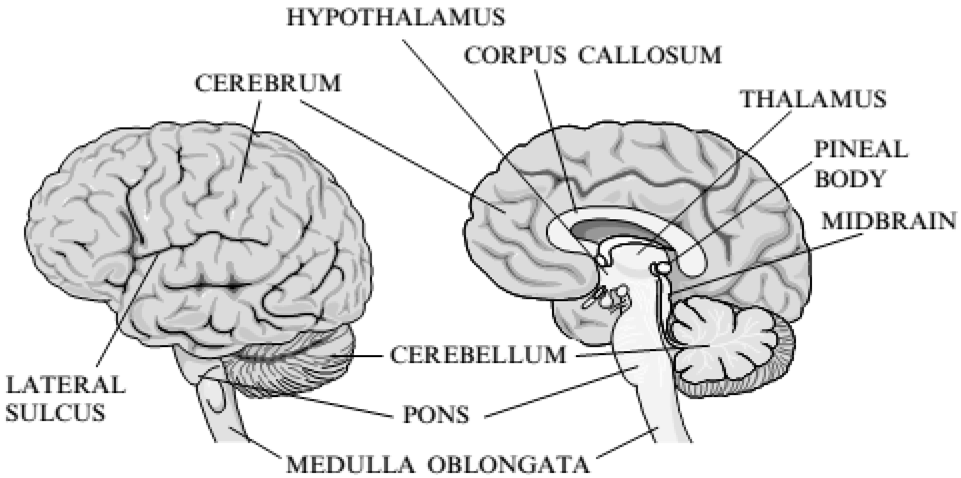

Figure 2.1: a) The exterior surface of the brain. b) A midsagittal section through the brain

The left side of the cortex controls motor and sensory functions from the right side of the body, whereas the right side controls the left side of the body. Association areas that interpret incoming data or coordinate a motor response are connected to the sensory and motor regions of the cortex. Fissures divide each cerebral hemisphere into a series of lobes that have different functions. The functions of the frontal lobes include initiating voluntary movement of the skeletal muscles, analyzing sensory experiences, providing responses relating to personality, and mediating responses related to memory, emotions, reasoning and judgment, planning, and speaking. The parietal lobes respond to stimuli from cutaneous (skin) and muscle receptors throughout the body. The temporal lobes interpret some sensory experiences, store memories of auditory and visual experiences, and contain auditory centers that receive sensory neurons from the cochlea of the ear. The occipital lobes integrate eye movements by directing and focusing the eye and are responsible for correlating visual images with previous visual experiences and other sensory stimuli. The insula is a deep portion of the cerebrum that lies under the parietal, frontal, and temporal lobes. Little is known about its function, but it seems to be associated with gastrointestinal and other visceral activities. The diencephalon is the deep part of the brain that connects the midbrain of the brain stem with the cerebral hemispheres. Its main parts are the thalamus, hypothalamus, and epithalamus. The thalamus is involved with sensory and motor systems, general neural background activity, and the expression of emotion and uniquely human behaviors. Due to its two-way communication with areas of the cortex, it is linked with thought, creativity, interpretation and understanding of spoken and written words, and identification of objects sensed by touch. The hypothalamus is involved with integration within the autonomic nervous system, temperature regulation, water and electrolyte balance, sleep–wake patterns, food intake, behavioral responses associated with emotion, endocrine control, and sexual responses. The epithalamus contains the pineal body that is thought to have a neuroendocrine function. The brain stem connects the brain with the spinal cord and automatically controls vital functions such as breathing. Its principal regions include the midbrain, pons, and medulla oblongota. The midbrain connects the pons and cerebellum with the cerebrum and is located at the upper end of the brain stem. It is involved with visual reflexes, the movement of eyes, focusing of the lenses, and the dilation of the pupils. The pons is a rounded bulge between the midbrain and medulla oblongata which functions with the medulla oblongata to control respiratory functions, acts as a relay station from the medulla oblongata to higher structures in the brain, and is the site of emergence of cranial nerve V. The medulla oblongata is the lowermost portion of the brain stem and connects the pons to the spinal cord. It contains vital centers that regulate heart rate, respiratory rate, constriction and dilation of blood vessels, blood pressure, swallowing, vomiting, sneezing, and coughing. The cerebellum is located behind the pons and is the second largest part of the brain. It processes sensory information that is used by the motor systems and is involved with coordinating skeletal muscle contractions and impulses for voluntary muscular movement that originate in the cerebral cortex. The cerebellum is a processing center that is involved with coordination of balance, body positions, and the precision and timing of movements(Blanchard, 2005).

2.2 Electroencephalography

Bioelectromagnetic signals generated by the neurons of the cortex can be detected through two different techniques: electroencephalography (EEG) that records changes in the electric field generated by a group of pyramidal neurons, and magnetoencephalography (MEG) that records variations in the magnetic field induced by the variations of the electric field generated by the same neurons. The two phenomena are closely related and they are described by the Maxwell's equations that represent the basis of the electromagnetic theory. In 1929 Hans Berger detected a difference in the electrical potential between two needles in the scalp, or between two electrodes placed on the defatted skin of the scalp. This technique was subsequently consolidated by Jasper Herbert providing the basis for the modern electroencephalography technique. The electric potential differences generated by ionic currents, which cross the synaptic membranes and that induce a flow of external charges in the extracellular space, reach the surface of the scalp being measurable via electrodes placed on the scalp. However, the acquired signal will be attenuated, due to the low conductivity of the skull. In the EEG signal is possible to identify and classify six main frequency ranges:

- Delta is the frequency range up to 4 Hz. It tends to be the highest in amplitude and the slowest waves. It is seen normally in adults in slow wave sleep. It is also seen normally in babies. It may occur focally with subcortical lesions and in general distribution with diffuse lesions, metabolic encephalopathy hydrocephalus or deep midline lesions.
- Theta is the frequency range from 4 Hz to 7 Hz. Theta is seen normally in young children. It may be seen in drowsiness or arousal in older children and adults. It can be seen as a focal disturbance in focal subcortical lesions, in generalized distribution in diffuse disorder or metabolic encephalopathy or deep midline disorders or some instances of hydrocephalus. On the contrary this range has been associated with reports of relaxed, meditative, and creative states.
- Alpha is the frequency range from 7 Hz to 14 Hz. It emerges with closing of the eyes and with relaxation, and attenuates with eye opening or mental exertion. The posterior basic rhythm is actually slower than 8 Hz in young children (therefore technically in the theta range). In addition to the posterior basic rhythm, there are other

normal alpha rhythms such as the mu rhythm (alpha activity in the contralateral sensory and motor cortical areas that emerges when the hands and arms are idle; and the "third rhythm" (alpha activity in the temporal or frontal lobes).

- Beta is the frequency range from 15 Hz to about 30 Hz. It is seen usually on both sides in symmetrical distribution and is most evident frontally. Beta activity is closely linked to motor behavior and is generally attenuated during active movements. Low amplitude beta with multiple and varying frequencies is often associated with active, busy or anxious thinking and active concentration.
- Gamma is the frequency range approximately 30–100 Hz. Gamma rhythms are thought to represent binding of different populations of neurons together into a network for the purpose of carrying out a certain cognitive or motor function.
- Mu ranges 8–13 Hz, and partly overlaps with other frequencies. It reflects the synchronous firing of motor neurons in rest state. Mu suppression is thought to reflect motor mirror neuron systems, because when an action is observed, the pattern extinguishes, possibly because of the normal neuronal system and the mirror neuron system "go out of sync", and interfere with each other(Niedermeyer and Silva, 2005).

The simultaneous activation of a whole population of neurons is necessary in order to measure the cortical electrical activity on the scalp.

This is what happens with the efferent impulses from the thalamus, which can activate simultaneously hundreds of cortical neurons. The sum of postsynaptic potentials (PPS) is sustained by a flow of ionic currents (primary currents) that cross the synaptic membranes. These currents induce a similar flow of external charges (secondary currents) in the extracellular space, which, flowing through all tissues encephalic, finally reach the surface of the head, although they are substantially attenuated by the low conductivity of the cranial bones. Specific electrodes placed on the scalp can detect the electric potential differences generated on the surface of the skull; usually the electrodes consist of metallic discs. In order to ensure a good contact between the skin and the electrode, usually a conductive gel is used. The electrode can be glued directly on the scalp, by means of special adhesive substances, or fixed on a headset or bandage that is subsequently placed on the scalp. The electrodes transduce the ionic current present in the tissues in the electric current compatible with the measurement electronic equipment. This signal transduction takes place by means of a chemical reaction between the metal of which the electrode is made and its salt which is present in the conductive paste.

In order to reconstruct the cortical activity, the EEG signal is measured simultaneously at different points of the scalp. The 10–20 system or International 10–20 system is an internationally recognized method to describe and apply the location of scalp electrodes in the context of an EEG test (Jurcak et al., 2007). The "10" and "20" refer to the fact that the actual distances between adjacent electrodes are either 10% or 20% of the total front–back (nasion - inion) or right–left distance of the skull (ear lobes).

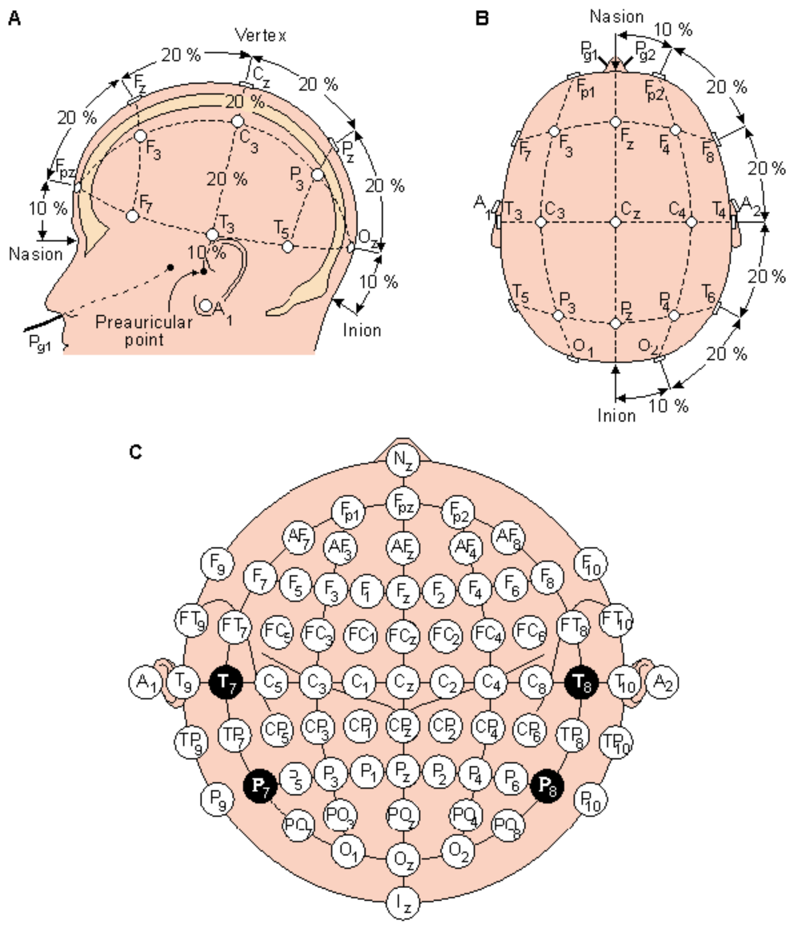

Figure 2.2 A-B) Standard international 10-20 electrodes placement system; C) Standard international 10-5 electrodes placement system

2.3 Event-related potentials

The evoked potential (EP) is a response induced by the presentation of an external stimulus that can be isolated from spontaneous EEG activity. This means that, for any external sensory stimulation, the brain responds with a specific wave, characterized by a particular latency, amplitude, and polarity. A given evoked potential appears at approximately constant intervals from the presentation of the stimulus. Since the amplitude of each response is low compared to the amplitude fluctuations of spontaneous EEG, the answer is extracted from the background activity as the average of a series of single responses (synchronized averaging) in order to significantly reduce the random fluctuations of the EEG. In fact, the variations of potential not synchronous with the stimulus are deleted, while those synchronous are added. The average evoked waveform is clearly evident respect to the background of continuous electrical activity and is characterized by a peak that can have positive or negative polarity (P or N). From the physiological point of view the evoked potentials are defined as electrical changes that occur in the central nervous system in response to an external stimulus: in this sense evoked potentials, such as acoustic and short-latency somatosensory and visual stimulus pattern reversal, represent the induced response of neuronal pool to a particular stimulus. In fact, their identifying features (latency and amplitude), depend on the physical characteristics of the stimulus (for example, tone and volume for the auditory system; contrast, luminance and spatial frequency for the visual system; intensity and stimulation mode for the somatosensory system). The evoked potentials can be divided into two main categories: the stimulus-related evoked potentials, which depend on the physical characteristics of the stimulus, and event-related potentials (ERPs). The generation of the ERPs is a function of the psychological context (event) in which stimulation occurs and their latency depends on cognitive and attentional phenomena(Picton et al., 2000).

2.3.1 Stimulus-related evoked potentials

The stimulus-related evoked potentials can be divided into short, medium and long latency components. For instance in the context of acoustic stimulus-related potentials, the short- latency components, generated by the cochlear nerve and brainstem structures, take place within the first 10ms from the stimulus (BAEPs : brainstem auditory evoked potentials). The medium latency components appear 10 - 50 ms after the onset of the stimulus, and consist of a series of deflections (N0, P0, Na, Pa , Nb , Pb or P1) that are believed to be a combination of potentials related to a reflex muscle activity and neuronal activity, probably generated by the thalamocortical radiation, the primary auditory cortex and associative areas. The long latency components (greater than 50 ms) are N1 and P2, which have their maximum amplitude at the top and for this reason are called " Vertex Potential". The N1 and P2 components are composed of both exogenous and endogenous attributes, i.e. not solely reflect the physical nature of sensory stimuli that have been used to evoke them, and for this reason are also called "mixed". The N1 component appears after about 100ms from the stimulus presentation; has been shown that the amplitude of this wave is increased in response to the expected stimuli compared to those not expected. The P2 is a positive-vertex wave with a latency of about 165ms from the onset of the stimulus. In an oddball task it is related to the rare and unexpected stimulus. It is probably an early manifestation of the decision-making processes related with the subsequent endogenous components N2 and P300.

2.3.2 Event-related potentials

The event-related potentials (ERPs), unlike the stimulus related potentials, depend on the informative content of the stimulus, in fact, they appear only when the subject pays attention to the stimulus and when a "meaning" has been attributed to it. One of the main characteristics of the ERPs is a close temporal relationship between the stimulation and the response to the stimulus itself. An ERP is the result of a phase correction of the background oscillations related to the event of interest and an increase in power. The changes produced by the external event (visual stimulus, sound or movement) are always at a fixed distance in time (latency) of the event of interest. The temporal relationship between the stimulus (or movement) and the oscillatory activity is very stable and highly repeatable over several repetition of the stimulus. The repeatability is around a few milliseconds. Both processes (phase correction and power increment) may be restricted to a specific frequency ranges, but generally, the ERPs cover the whole spectrum from 4-6 Hz to 60-70 Hz. Usually, the fastest oscillations (higher frequencies) occur just after the stimulus, whilst the ERP signal slows down and returns to the background later with respect to the event. The ERPs (due to both external stimuli and motor activity) consist of electrical potential oscillations and the waveforms are characterized by a series of positive or negative deflections. These deflections are normally defined components. The polarity of the ERP components depends on the EEG electrode position respect to the distribution of the electric field. In turn, the distribution of the superficial fields depends on the cortical area activated and its orientation with respect to the scalp. Several ERPs have been identified according to their latency, to the characteristics of the stimuli and to the paradigms used to elicit them. Considering their latency, the first component of the ERPs is vertex-negative wave with a latency of about 200ms called N2. The N200 potential is induced by the rare stimuli, expected or less, and it is followed by P300 when the subject performs a specific task of recognition, such as in the "oddball" paradigm. P300 is a positive deflection (around 10μV) recorded over the scalp central-parietal regions and occurring 250-400 ms after the recognition of a rare or relevant stimulus (Target) within a train of frequent stimuli (Non-Target - Fabiani et al., 2000).

2.3.3 The P300 potential

ERPs are affected by several processes relating to the selective attention, memory, semantic comprehension, elaboration of information.

Duncan-Johnson and Donchin (1977) operationally defined the P300 as a component with a latency longer than 275ms, positive in polarity al all midline electrode locations (in comparison with a “neutral” reference), with a maximum positivity at parietal and central locations, elicited by task relevant stimuli, and whose amplitude is affected by the subjective probability and task relevance of the stimulus. It is endogenous and is related to the attention levels, which affect the subject's concentration and the information processing mechanisms. The P300 component indicates the level of activation of the central nervous system that processes the stimulus. The P300 amplitude increases when the probability of presentation of the target stimuli decreases and is directly proportional to the interval between two consecutive stimuli (Inter Stimulus Interval - ISI; Duncan et al., 2009). A longer interval between target stimuli (Target to Target Interval - TTI) corresponds to higher P300 amplitude and a shorter latency (Gonsalvez and Polich, 2002). The latter increases when the target stimuli are difficult to distinguish from the standard ones and it is a measure of the classification speed of the stimulus, in addition the P300 latency can be used as an index of cognitive efficiency: in fact low latency values correlate with high performance in neuropsychological tests related to temporary memory abilities in healthy subjects. The P300 evoked potential has two subcomponents: P3a and P3b. The P3a exhibits the maximum amplitude in the frontal/central areas with a latency peak falling in the range between 250 and 280ms. The P3b has a peak at around 300ms, although it presents a variable latency between 250-500ms (Polich, 2007). The amplitude of the P3a component increases as decreases the discriminability between stimulus target and non-target stimulus. The P3b component is related to the storage within the parietal cortex of the information about the rare stimulus, which occurred after a comparison between the target and non-target stimuli ("working memory"). The P3a component is generated when the stimulus activates the area of the frontal lobe; P3b component is produced when the attentional resources are used for the evaluation of the stimulus (Figure 2.3). As soon as the stimulus arrives the first input reaches the right/central cortex and the amplitude of the ERPs components is higher in the central and frontal areas of the right hemisphere rather than in the left. After the stimulus storage, the activation is propagated in the corpus callosum: more the fibers of the corpus callosum are broad more the amplitude of the P300 component is high and its latency short (Polich and Criado, 2006; Polich, 2004). The posterior-parietal and the inferior temporal cortex generate the P3b component, while the anterior cingulate cortex produces the P3a component (Linden, 2005).

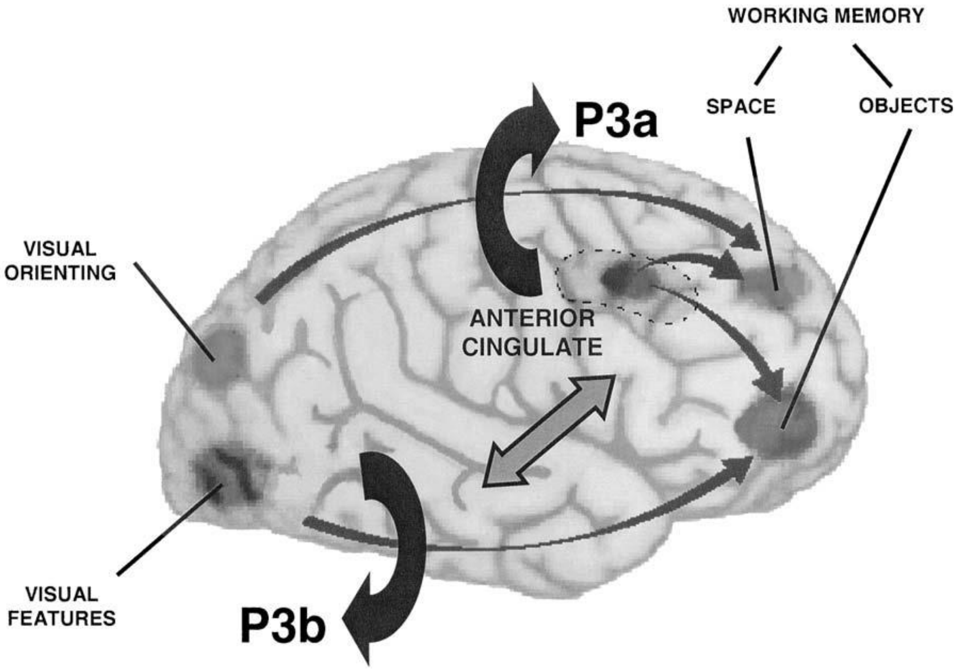

Figure 2.3 The P3a component is related to the anterior cingulate cortex activity during the elaboration of the stimulus performed by the working memory. The P3b component is related to the following activation of the hippocampus when the frontal region communicates with the temporal/parietal region.

2.3.4 Variability of the P300 potential

As explained in the previous section, the P300 morphology is affected by several factors, such as stimulus type, timing and paradigm. Other affecting factors are fatigue and stress (Yagi et al., 1999), attention level (Lutz et al., 2009), aging (Goodin et al., 1978) and drugs (Polich and Kok, 1995). The above mentioned factors produce an intra-subject variability of P300 potential even during the same day.

Ravden and Polich, (1999) carried out a study involving 20 healthy subjects, in order to characterize the variability of the P300 component across repeated measures, they found that the amplitude of the P300 component recorded by the electrodes Fz, Cz, and Pz and evoked by visual stimuli (delivered in blocks of 10 trials, each consisting of 50 trials and presented at intervals of 10 minutes) exhibited a reliable cyclical fluctuation across trial blocks, although P300 latency did not. Horne and Ostberg, (1976) correlated the ability of a person to perform a task with the different hours of the day, distinguishing between Morning-types subjects who perform better the required tasks in the morning, and the Evening-types who reach better performance in the late afternoon. Huang et al. (2006) by mean of an experimental protocol based on an auditory paradigm involving 13 healthy subject, pointed out a significant difference in P300 amplitude of the same subject at different moment of the day: the Morning-Types, who prefer to be active in the early hours of the day, exhibited a higher P300 amplitude at 9:00AM rather than 5:00PM, on the contrary for the Evening-type, usually active later in the day, the P300 amplitude was higher at 5:00PM than at 1:00AM. Such diurnal variation of P300 amplitude has been attributed to the decrease in the excitability of the central nervous system and in cognitive processes experienced by the Morning-type and Evening-type at 5:00PM and 01: 00AM respectively. Since this phenomenon is not accompanied by a decrease in the P300 latency, it has not been attributed to a variation of the task difficulty that the subject has to perform, which also would cause a change in P300 amplitude, but to a lower attention level of the subject. Although it is known that the Reaction Time of a person subjected to the stimulus, and the latency of the P300 component, which reflects the speed of information storage, are not related to each other, the study by Huang et al. showed that they have opposite trends during the day depending on the type of subject.

2.4 Brain Computer Interface

The central nervous system manages interaction with the external world by producing neuromuscular or hormonal outputs. Wolpaw and Wolpaw (2012) defined a Brain-Computer Interface as *"a system that measures CNS activity and converts it into artificial output that replaces, restores, enhances, supplements, or improves natural CNS output and thereby changes the ongoing interactions between the CNS and its external or internal environment"*. In other words a BCI records and elaborates brain signals (such as electric or magnetic field, hemoglobin oxygenation, etc.) looking for specific measures (or features), then the latter are converted into artificial outputs that act on the outside world. Figure 2.4 summarizes the main applications of a BCI:

- Replace: A BCI output can replace natural output that has been compromised by a disease or an injury. For instance BCI can support and restore communication and interaction with external world in people with severe motor impairment.
- Restore: A BCI output can restore lost natural output. For instance, a person with spinal cord injury whose arms and hands are paralyzed can use a BCI to stimulate the paralyzed muscles via implanted electrodes so that the muscles move the limbs.
- Enhance: A BCI output might enhance natural CNS output. For instance a person performing a demanding and critical task (such as a surgeon or a driver) can use BCI in order to detect brain activity preceding lapses in attention and provide and alarm that alerts the person and restores attention
- Supplement : A BCI output might supplement natural CNS output, for instance a person might use a BCI to control a robotic arm and hand, supplementing natural neuromuscular output with and additional, artificial output;
- Improve: A BCI output might improve natural CNS output. For instance a BCI can be used as a rehabilitation tool for stroke

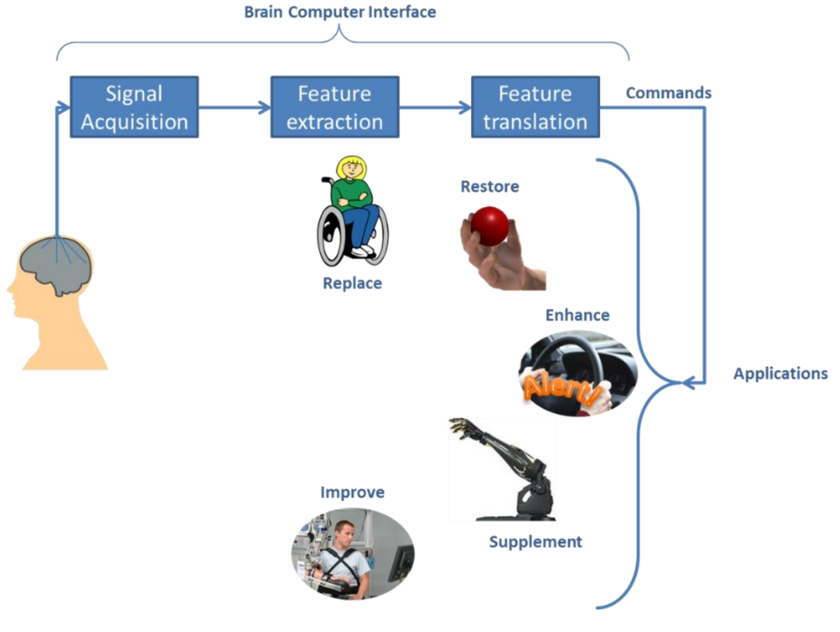

Figure 2.4: the five main application of Brain Computer Interface systems (Wolpaw and Wolpaw 2012)

BCI system can be distinguished in dependent and independent BCI, both use brain signals to control their applications but they differ in their dependence on natural CNS output. A dependent BCI relies on brain signals that depend on muscle activity so they are an alternative method for detecting messages carried in natural CNS outputs. On the contrary, an independent BCI does not depend on natural CNS output.

Like any communication or control system, a BCI consist of an input (e.g. electrophysiological activity from the user), an output (i.e. device commands), components that translate input into output, and a protocol that determines the onset, offset, and timing of operation. In non-invasive BCIs the input is EEG recorded from the scalp. The digitized signals are then subjected to one or more of a variety of feature extraction procedures, such as spatial filtering, voltage amplitude measurements, spectral analyses. This analysis extracts the signal features that encode the user's messages or commands. BCIs can use signal features that are in the time domain (e.g. evoked potential amplitudes) or the frequency domain (e.g. mu or beta-rhythm amplitudes). The first part of signal processing simply extracts specific signal features. The next stage, the translation algorithm, translates these signal features into device commands-orders that carry out the user's intent. This algorithm might use linear methods (e.g. classical statistical analyses) or nonlinear methods (e.g. neural networks). Whatever its nature, each algorithm changes independent variables (i.e. signal features) into dependent variables (i.e. device control commands). For most current BCIs, the output device is a computer screen and the output is the selection of targets, letters, or icons presented on it. Selection is indicated in various ways (e.g. the letter flashes). Some BCIs also provide additional, interim output, such as cursor movement toward the item prior to its selection. As mentioned before BCI can also control a neuroprosthesis or orthesis that provides hand closure to people with cervical spinal cord injuries, or other electrical devices such as domotic appliance or wheelchair. Finally, each BCI has a protocol that manages its operation. This protocol defines how the system is turned on and off, whether communication is continuous or discontinuous, whether message transmission is triggered by the system (e.g. by the stimulus that evokes a P300) or by the user, the sequence and speed of interactions between user and system, and what feedback is provided to the user.

The control features mainly exploited by non-invasive BCI are 4:

- Slow Cortical Potentials are gradual changes in the membrane potentials of cortical dendrites that last from 300ms to several seconds (1-2Hz). Decreases in the cortex potential are associated with motor activity, while positive variations correspond to a decrease of cortical activity.
- Sensorimotor rhythms (SMRs) are associated with the cortical areas related to natural motor channels of the brain and include the sensory-motor frequencies (12 - 15Hz), the Mu rhythm (8-13Hz), and the central beta (13 -26Hz) and gamma (up to 32Hz) bands. In correspondence with the movement or with its preparation a decrease in Mu and Beta rhythms called "event-related desynchronization" (ERD) occurs, while with the relaxation such rhythms increase and this phenomenon is defined as "event-related synchronization" (ERS). The peculiarity of the sensors motors rhythms is that they are generated rather in correspondence of the performed action and with its imagination. This aspect is of fundamental importance for the use of these rhythms as a BCI control feature for subjects with impaired motor abilities. Moreover, they can be enhanced through training. This type of BCI allows the user to move a cursor or a "joystick" on a screen. Recently they have been successfully applied for rehabilitation of patients with stroke outcome
- The Steady State Visual Evoked Potentials (SSVEPs) represent the neuronal responses to visual stimulation at a specific frequency. They are widely used in this research field thanks to the excellent signal to noise ratio and to the considerable immunity to artifacts.
- The Event-related potentials, such as P300 potential that has been widely described in previous sections.

2.4.1 P300-based Brain computer interface

The P300 event related potential is one of the most effective control features for communication and control since it does not requires a long training and allows a relatively quick selection of a target within a group of possible choices. A classic BCI that relies on P300 potential appears like a keyboard, allowing the user to choose between a finite number of options (Farwell and Donchin, 1988). In a P300 based BCI system the main issue is the recognition of the Target stimulus. Usually each item on the interface is presented several times to the subject who is required to pay attention to the presentation of the item of interest (Target). Only the presentation of the Target item will elicit a P300 potential, the latter usually is identified by mean of a synchronized average and of statistic elaborations of the EEG signal.

2.5 State of the art of BCI as assistive technology and users' needs

Since BCI system might allow communication and interaction with external world by-passing the usual pathway of the nervous system such as muscle and nerves (Wolpaw et al., 2002), people with amyotrophic lateral sclerosis (ALS) or locked-in syndrome (LIS) have been appointed as target users for BCI system for communication and control. However, considering the complexity and the heterogeneity of the deprivation pattern of motor abilities experienced by person with ALS, an assistive technology complying with their needs can be successfully adopted by people affected by other diseases (spinal muscular atrophy, multiple sclerosis, etc.). In patients with ALS, communication difficulties usually result from progressive dysarthria, while language functions remain largely intact. When this status progresses, augmentative and alternative communication (AAC) systems that can substantially improve the quality life are needed (Andersen et al., 2005). For ventilated patients, eye-pointing and eye gaze based high-tech assistive technologies have been proven to be useful. Similarly, a BCI could help users communicate with devices and other people. Professor Birbaumer established the first communication with a locked-in patient in the 90s (Birbaumer et al., 1999). Later, several studies aimed to show the feasibility and to compare the performances with healthy subjects using either slow cortical potentials (Kübler et al., 2004) or cognitive evoked potentials like P300 (Piccione et al., 2006) or motor imagery (MI) (Kübler et al., 2005). Later research has further shown that persons, even despite severe disabilities, may interact with computers by only using their brain—in the extreme case using the brain channel as a single switch, just like a hand mouse. Research on establishing communication functions were mostly focused on writing (spelling) applications and surfing (browsing) the Internet. Several spelling devices based on the voluntary modulation of brain rhythms have been demonstrated. These systems can operate synchronously (Birbaumer et al., 1999; Parra et al., 2003) or asynchronously (del Millan et al., 2002; Millán et al., 2004; Muller and Blankertz, 2006; Scherer et al., 2004; Williamson et al., 2009). Mostly binary choices of the BCI were used to select letters, e.g. in a procedure where the alphabet was iteratively split into halves (binary tree). The big disadvantage of all these systems is that the writing speed is very slow. Particularly relevant is the spelling system called Hex-O-Spell (Williamson et al., 2009), which illustrates how a normal BCI can be significantly improved by state-of-the-art human-computer interaction principles, although the text entry system is still controlled only by one or two input signals (based on motor imagery). The principle of structuring the character locations based on an underlying language model speeds up the writing process.

Other kinds of BCI spelling devices, especially those mostly used by disabled people, are based on the detection of potentials that are evoked by external stimuli. The most prominent is the approach that elicits a P300 component (Farwell and Donchin, 1988). In this approach, all characters are presented in a matrix. The symbol on which the user focuses her/his attention can be predicted from the brain potentials that are evoked by random flashing of rows and columns. Similar P300-based spelling devices have extensively been investigated and developed since then (e.g. Allison and Pineda, 2006; Nijboer et al., 2008; Piccione et al., 2006; Sellers et al., 2006; Silvoni et al., 2009). Additionally, steady state visual evoked potentials (SSVEPs) can be used for virtual keyboards. Either each character of the alphabet or each number on a number pad is stimulated with its own frequency and can be selected directly (Gao et al., 2003), or additional stimulation boxes (like arrows) are placed aside the keyboard and are used for navigating left/right/up/down and selecting the letter (Valbuena et al., 2010).

The first application to access the Internet via the BCI was a very simple solution, by displaying web pages for a fixed amount of time (‘Descartes’ by Karim et al., 2006), but later browsers allowed a more flexible selection of links (‘Nessi’ by Bensch et al., 2007). The challenge of selecting a large amount of links with only a limited amount of BCI commands (mostly two) can be overcome by applying scanning techniques, which allow a sequential switching or auto-switching between them. Even functions like zoom in/out, scroll up/down, go back/forward can be added in the user interface and selected by the BCI via a hierarchical approach. Nevertheless, users reported that the correct selection can be quite demanding (Leeb et al., 2011). More recently, different groups have developed Internet browsers based on P300 potentials. In the first one, all possible links are tagged with characters, and a normal character P300 matrix (6x6 matrix) was used on a separate screen for selection (Mugler et al., 2010). In a more recent approach, an active overlay was placed over the web site that elicited the P300 by directly highlighting the links. Hence, switching between the stimulation device and the browsing screen was not necessary (Riccio et al., 2011). After nearly 20 years of research a first commercial BCI system for typing was released recently, called IntendiX® (g.tec medical engineering, Schiedelberg, Austria). The system relies on VEP/P300 potentials to use for patients with motor disabilities. However, despite the state of the art demonstrates that research is going on providing several applications and solutions, BCI systems are still mainly research prototypes. There are only few cases of people using BCI system in their daily life (Sellers et al., 2010) and this is due to their low information transfer rate and reliability. A BCI is a complex system requiring intricate procedures for their configuration and calibration, and specific knowledge is necessary to understand if everything is well arranged. Recently, considering the example provided by the human-computer interaction field, the user-center design approach has been applied in order to provide solutions that can satisfy the users' needs. The study by Huggins et al. (2011) reports the results of a telephone survey about the opinions and priorities of people with ALS regarding BCI design. The 61 people with ALS that were interviewed expressed a strong interest in obtaining BCIs, but, at the same time they pointed out that current BCIs do not yet provide desired performance. In fact, with regard to BCI design, participants prioritized accuracy of command identification of at least 90% (satisfying 84% of respondents), speed of operation comparable to at least 15-19 letters per minute (satisfying 72%), and accidental exits from a standby mode not more than once every 2-4 h (satisfying 84%). While 84% of respondents would accept using an electrode cap, 72% were willing to undergo outpatient surgery and 41% to undergo surgery with a short hospital stay in order to obtain a BCI. Blain-Moraes et al. (2012) carried out a focus group with 8 individuals with ALS and their caregivers (n = 9) to determine the barriers to and mediators of BCI acceptance in this population. Focus group participants expressed their interest in BCI technology and its potential, since BCI can give users freedom from the confining effects of their condition by enabling them to stay connected to their family and allowing them to remain independent as their condition progressed. However, participants also expressed that BCI technology in its current form would not be acceptable or appropriate for day-to-day use in practical, real-world situations. Participants identified a number of psychological, physiological and physical personal issues that could present as unyielding barriers towards the successful, real-world use of BCIs. For example, if individuals are unable to manage the physiological and psychological fatigue associated with BCI use, it is unlikely that this technology will be successfully integrated into their day-to-day communication and activity patterns

3 AIMS

The general objective of this PhD thesis is to investigate factors that improve the overall usability and reliability of P300-based BCI systems in order to transfer them from research prototypes to assistive technologies. According to ISO recommendations (ISO 9241 1998), the concept of usability (i.e. “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use”) can be split into three different measures: *(i) effectiveness*, which estimates the accuracy and the completeness with which intended goals are achieved; *(ii) efficiency*, which is the measure of the amount of human, economic and temporal resources expended in attaining the required level of product effectiveness; and *(iii) satisfaction*, a measure of the immediate and the long-term comfort and acceptability of the overall system.

With regard to effectiveness of the system new algorithms to improve the system accuracy will be proposed to: *i)* avoid false positives and misclassifications due to subject distractions or voluntary no-control periods; *ii)* ensure high reliability of the system for long term usage; *iii)* detect and report possible failures of the system without the need for the presence of a specialized technical staff.

The efficiency of the system will be improved introducing innovative classification algorithms able to automatically and continuously adapt to the current user state, in terms of speed of selection and of control parameters.

User satisfaction will be taken into account by: *i)* reducing the configuration and calibration procedures for an independent usage of BCI systems, *ii)* involving end users in the design, development and evaluation of an assistive technology including a BCI as input device. Needs, opinions and feedback of end users will be collected and analyzed in order to improve their satisfaction concerning the usage of the BCI systems.

4 ASYNCHRONOUS CLASSIFICATION IN P300-BASED BCI

In the past twenty years, a great body of BCI research has addressed relevant issues for successful deployment of P300-based BCIs that encompasses technological aspects such as stimulation framework (Brouwer and van Erp, 2010; Townsend et al., 2010; Treder and Blankertz, 2010), signal processing and classification (Lotte et al., 2007) and, more recently, aspects related to system usability outside of research laboratories (Sellers et al., 2010). Whereas the former aspects can exploit knowledge emerging from BCI system testing with the participation of healthy volunteers, the validation of devices with actual end-users remains essential to address and quantify their usability and usefulness. In spite of the amount of research devoted to P300-based BCIs, one aspect concerning the modality of functioning and in turn their usability, still awaits for further implementation. Indeed, P300-based BCI systems typically work in a synchronous mode: i.e., they provide a well-defined sequence of stimuli lasting a predefined time, after which the system always “makes” a decision. This operating mode causes obvious drawbacks when using these systems in everyday life. First, in synchronous modality it is assumed that the user is continuously in control of the interface. Second, several factors such as fatigue and stress (Yagi et al., 1999), attention level (Lutz et al., 2009), aging (Goodin et al., 1978) and drugs (Polich and Kok, 1995) have been described to influence the amplitude and/or latency of the P300 potential. These factors should be taken into account when considering the system performance (i.e., time needed for a given selection) under the real-life conditions. To cope with the above mentioned issues it is necessary to empower P300-based BCI systems for their usage in real-life contexts under which the potential users may or may not intend to send a command and/or she/he could be “distracted” by external events other than BCI control. In other words, the systems should be endowed with an operability mode which “avoids” possible false classifications; in fact, transferring a BCI from the laboratory environment to real-life settings requires a BCI system to recognize the user’s intent without any additional external input (Vaughan et al., 2006). This is the reason for the growing interest regarding the issue of asynchronous (self-paced) BCI control design. Several studies have addressed the issue of the asynchronous BCI in the domain of sensorimotor rhythms (Del R Millan et al., 2009; Mason and Birch, 2000; Townsend et al., 2004). With regard to P300-based BCI several classification algorithms have been proposed to provide a solution to the issue of dynamically adapt the number of stimuli repetitions, and then the communication speed of the system, to the current user's state (Jin et al., 2011; Lenhardt et al., 2008; Schreuder et al., 2011). However the mentioned algorithms cannot recognize when the user is not paying attention to the stimulation interface and then they do not automatically suspend the control producing false positives during voluntary no control periods. The latter features was first provided by Zhang et al. (2008), who proposed a computational approach to implement an asynchronous mode of control for the P300-based BCI. Using statistical and probability models about control and no-control user's state, they developed an algorithm that first recognizes the control state and then looks for the target, giving a classification after at least three stimulation sequences, as a result. They used a stimulation interface containing nine numeric items, and stimuli were provided, intensifying each single button.

This chapter presents an innovative approach for asynchronous classification in P300-based BCIs; in fact it relies on the widely validated row/columns paradigm allowing for higher information transfer rate with respect to the single character approach. The first evaluation study was conducted involving healthy users and allowed to perform preliminary studies on EEG data regarding periods in which the user was attending to the stimulation and periods in which he was engaged in different tasks. On the results of preliminary study the algorithms for asynchronous classification was defined and validated for environmental control. In fact environmental control is an important challenge for people who have (partially) lost motor ability. The opportunity to independently perform simple everyday actions, such as turning on/off lights, opening a door or changing TV channels, might represent a significant improvement in their quality of life, by reducing the level of dependency from the caregivers. In this respect, the recent advance in the field of domotics (a set of methods and techniques for the automation of the home) has remarkably augmented the potential to interact with the environment and thus, the management of everyday life activity in case of disability (as a forefront example see SM4ALL - Smart Homes for All European project – Baldoni et al., 2009).

In a later study the same algorithm was validated involving potential end users (people with muscular and motor diseases) for controlling virtual apartment. Finally a comparison of the proposed asynchronous system with a synchronous one will be reported in terms of communication efficiency as defined by Bianchi et al. (2007).

4.1 Asynchronous classifier and evaluation with healthy users

4.1.1 Study Design and Participants

Eleven healthy volunteers (4 females, 7 males; mean age and std 26,45±4,05 years) were involved in the study. Five out of 11 subjects were naive to P300-based BCI context.

The acquisition protocol was based on the P300 Speller application (Farwell and Donchin, 1988) within the BCI2000 framework (Schalk et al., 2004). This application was adapted to control a home automation system by using a 4 by 4 matrix which allowed a total of N_s = 8 stimulation classes. As shown in Figure 4.1, the stimulation interface consisted of 16 B&W icons representing the achievable actions on the environment. An asynchronous operating mode was implemented within the BCI2000 framework (see below). The fixation cross was static in the middle of the interface and never flashing, also during the stimulation.

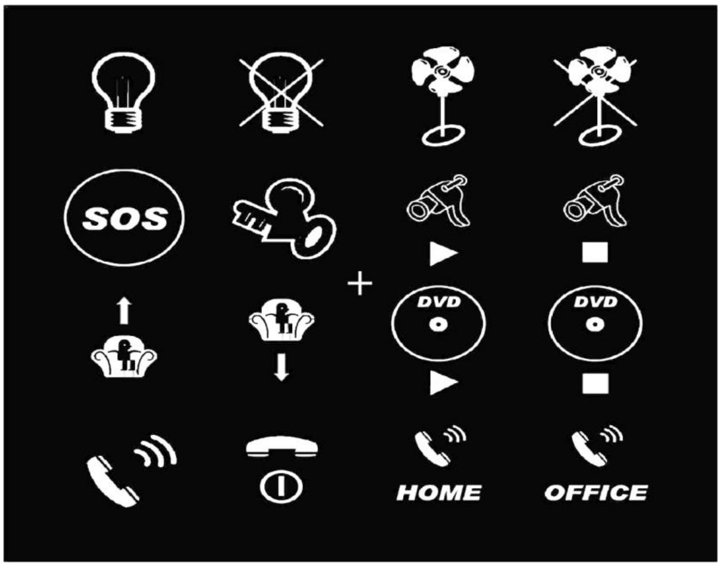

Figure 4.1 Stimulation Interface composed by several icons coding for domotic control. Stimuli consisted in an intensification of rows and columns. A no flashing fixation cross was placed in the center of the screen and used in some No-Control trials.

Scalp EEG potentials were recorded (g.MobiLab, gTec, Austria, sampling rate 256 Hz) from 8 positions according to 10-10 standard (Fz, Cz, Pz, Oz, P3, P4, PO7 and PO8; (Krusienski et al., 2008)). Each channel was referenced to linked earlobe and grounded to the left mastoid. Stimulation was provided through a 22" LCD monitor: on one half of the screen there was the stimulation matrix, while on the other half there were movies played in a DVD player.

4.1.1.1 *Data Acquisition Protocol*

Stimuli were provided by the BCI2000 framework through the random intensification of rows and columns in the matrix. Each stimulus was intensified for 125ms, the inter stimulus interval (ISI) was set at 125ms, so that the stimulus onset asynchrony (SOA) interval lasted 250ms. The EEG signal was reorganized in overlapping epochs lasting 800ms and following the onset of each stimulus. Epochs were then grouped into sequences. A *Sequence* consisted in a single flash relative to each row and column on the control interface. A set of sequences in which the target icon was the same composed a *Trial*. Finally, a *Run* comprised a trial series.

In the synchronous system after a number of stimulation sequences fixed *a priori*, a selection was always made. On the contrary, the asynchronous mode allowed a selection only when the thresholds were exceeded. When the given thresholds were not reached after a fixed number of stimulation sequences, a new trial began with no selection occurring. During the experimentation, the system indicated the Target icon for the next trial, or presented the classification result within the 4 seconds between two trials. All subjects underwent a total of 3 recording sessions over two weeks (3 or 4 days elapsed between two sessions). The first 2 sessions were defined as *Training sessions* and the last one as *On-line session.*

4.1.2.1 *Training Session*

The aim of the first 2 recording sessions was to collect data for the off-line analysis and to extract parameters (features and thresholds) for the On-line session.

Figure 4.2 illustrates the Training sessions composition. Each session was composed by 8 runs. The first 2 runs were defined as Control runs and they consisted of 8 Control trials of 10 stimulation sequences each. During the Control runs, the subject had to mentally count the occurrences of Target icon that was cued by its intensification before the beginning of each trial. Throughout the training sessions all icons were presented as a Target in random order spanning each position on the matrix. This dataset was used both for off-line analysis about synchronous system and to extract control parameters for synchronous control runs in the On-line session. During the Alternate runs, data were acquired under Control and No-Control condition. In these runs, Control and No-Control trials alternated for a total of 10 Trials composed of 10 stimulation sequences each. During Control task, the subjects were asked to count Target icon flashing (as in the first 2 runs of the session). No control parameters were set during Control trials and thus, no feedback was provided about the on-line classification results. As for the No-Control condition, subjects attended 3 different tasks while the stimulation (icon flashing) was running:

- Fixation Cross, Training Session 1, Alternate Runs [3-8]: Subjects were instructed to fixate the cross in the center of the interface and to ignore the stimulation;
- Watch & Listen, Training Session 2, Alternate Runs [3-5]: Subjects were instructed to watch a movie displayed on the half of the screen beside the matrix;

• Computation, Training Session 2, Alternate Runs [6-8]: Subjects had to answer simple arithmetic questions posed by the operator while fixating the cross.

By doing so, It was ensured that No-Control data set would contain trials mimicking real life situation wherein the users could direct their attention elsewhere or could interact with other persons. Furthermore, under No-Control experimental condition, the user's visual field was not immune from the random stimuli thus, allowing to test the system robustness to artifacts and to reduce potential visual misclassification.

Each icon on the interface was presented as a Target with the same frequency, both for Control and Alternate runs, over the two training sessions with the aims to make all icons equal likelihood for subsequent analysis. Only one icon (the 16th) was never suggested as a Target during Alternate runs because we used it to indicate the No-Control trials to the users.

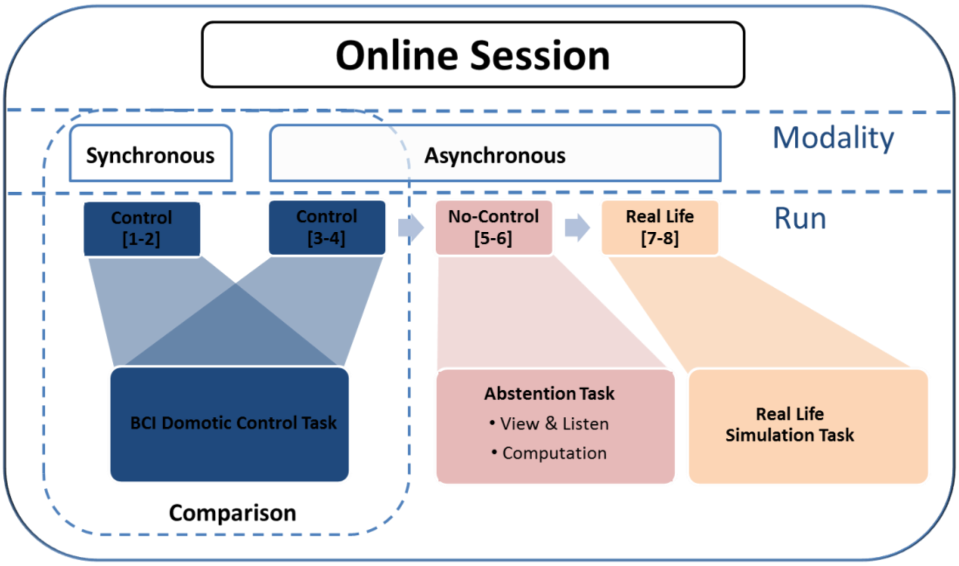

Figure 4.2 Organization of the training sessions. Each session includes 2 Control runs and 6 Alternate runs. The Control runs consist of 8 Control trials. The subject is always attending to the stimulation, mentally counting Target occurrences. The Alternate runs consist of 5 Control trials and 5 No-Control trials. During No-Control trials subject diverts his attention from the stimulation doing different actions (Fixation Cross, View & Listen, Computation).

4.1.3.1 *On-line Session*

Figure 4.3 illustrates the adopted scheme for the On-line sessions. The first 4 Control runs were performed in order to compare synchronous versus asynchronous system in terms of the time required to perform a previously defined list of actions. These actions were selected in order to test each device in the environment. The goal of each run was to complete 5 actions; the Target icons were cued at the beginning of each run and the subjects were also informed on how to correct each potential error by selecting the complementary action. The number of trials was not fixed a priori but depended of number of errors (synchronous and asynchronous modalities) and abstentions (asynchronous modality). Classification results were shown to the subject by intensifying a single icon, and through the corresponding device operation. Abstentions with the asynchronous system were fed back by intensifying all together the icons on the interface. The purpose of the 2 No-Control runs was to assess system reliability avoiding false positives. Each No-Control run took 5 minutes during which the stimulation was kept on. During the first and second No-Control run, the subjects were asked to refrain from the control by watching a movie or by answering to arithmetic questions looking at the fixation cross, respectively.

Finally, the last on-line runs were devoted to test the usefulness of asynchronous BCI in everyday life. Two different scenarios were simulated in order to quantify errors and false positives occurrence:

- Scenario 1, On-line session, Real Life [7]: someone rings the doorbell, immediately after the user turns on the interphone's video camera and waits until that the image appears (about 30 seconds). He's one of his friends. He decides to open the door and to turn on the light for his guest. Subsequently, he turns on the DVD and they watch a video together. After 1 minute the user turns off the DVD player (5 BCI commands).

- Scenario 2, On-line session, Real Life [8]: the user is tired and wants to relax, he lowers the chair and turns off the light. After 1 minute the phone rings and he answers. It's his sister and he talks with her for 1 minute, then he hangs up, turns on the fan and starts to sleep (5 BCI commands).

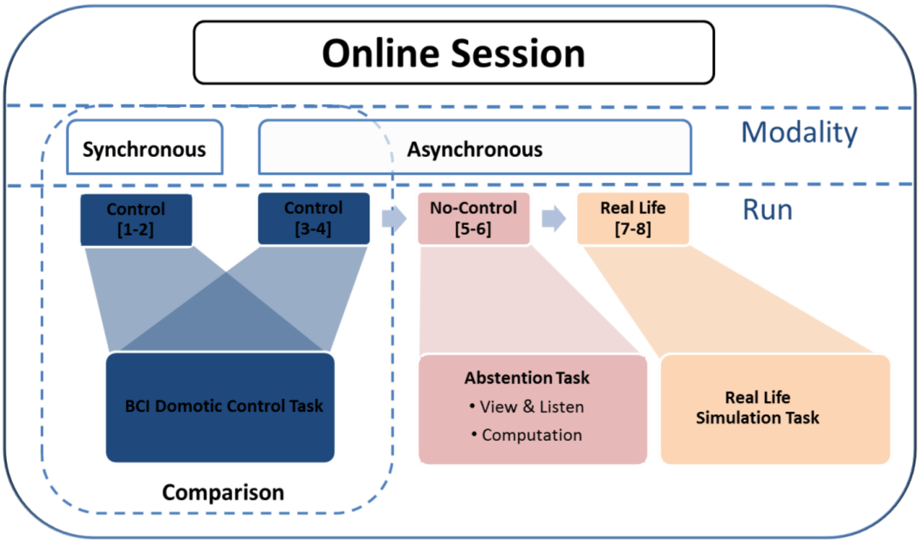

Figure 4.3 On-line session scheme. During Control runs subject had to achieve a previously defined list of actions, under both synchronous and asynchronous modalities. The No-Control runs consisted of "abstention" tasks (5 min lasting) in which different actions were required to perform. The last two "Real Life" runs simulated two real-life scenarios

4.1.2 EEG preprocessing, features extraction and classification

The EEG signal was divided into 800 ms epochs starting from the onset of each stimulus. It was possible to distinguish Target and No-Target epochs related to Control trials, and No-Control epochs related to No-Control trials. EEG epochs were subsequently reorganized into a three-dimensional array: each 2D matrix of the array represents a single epoch, where rows stand for acquisition channels and columns are correspondent to samples of each epoch. Despite the previous decimation (decimation factor = 3), the amount of data was still demanding and a further reduction in features space was performed by using the Stepwise Linear Discriminant Analysis (SWLDA; Krusienski et al., 2006). We divided the dataset related to Alternate trials in two parts: training and testing data set; the first one contains 3 runs from the first training session and 4 runs from the second one. In this way No-Control trials related to all the 3 different No-Control tasks were included in the training data set. The same runs were used for every subject to extract thresholds and features for asynchronous runs in the On-line session:

- Training session 1: Runs 3-4-5, Fixation Cross;
- Training session 2: Runs 3-4, View & Listen; Runs 6-7, Computation.

A SWLDA was applied on training data set including No-Control trials, assigning a label equal to zero to the No-Target and the No-Control epochs while label was equal to 1 for Target epochs. For the synchronous system all the Control runs of the training sessions were used to extract the features for the On-line session, while for the Off-line analysis a cross-validation was performed, using 2 runs to extract significant features and other 2 to test them. Finally for both asynchronous and synchronous modalities, the final discriminant function was restricted to contain a maximum of 60 features by SWLDA. Nonzero weights are assigned to these features, w. The score values for each epoch are then calculated as:

$$y_i = \sum_e w \cdot f_e + b \quad (1)$$

Where e denotes all features related to single stimulus i and b represents the intercept. For the classification it is assumed that a P300 is elicited for one of the four row/column intensifications during control period, and that the P300 response is invariant to row/column stimuli, the resultant classification in synchronous system taken as the maximum of the scored feature vectors for the respective rows, as well as for the columns. The icon that appears at the intersection of the predicted row and column in the matrix is the one chosen.

4.1.3 Threshold values extraction

The threshold values were chosen through a procedure that relies on the use of ROC curves (Schulzer, 1994). The score values on the Target, No-Target and No-Control epochs were assessed using data of Alternate runs; training data set to extract features and testing data set to estimate score value as explained in the previous section. Figure 4.4 shows the trend of this three normalized distributions for a representative subject.

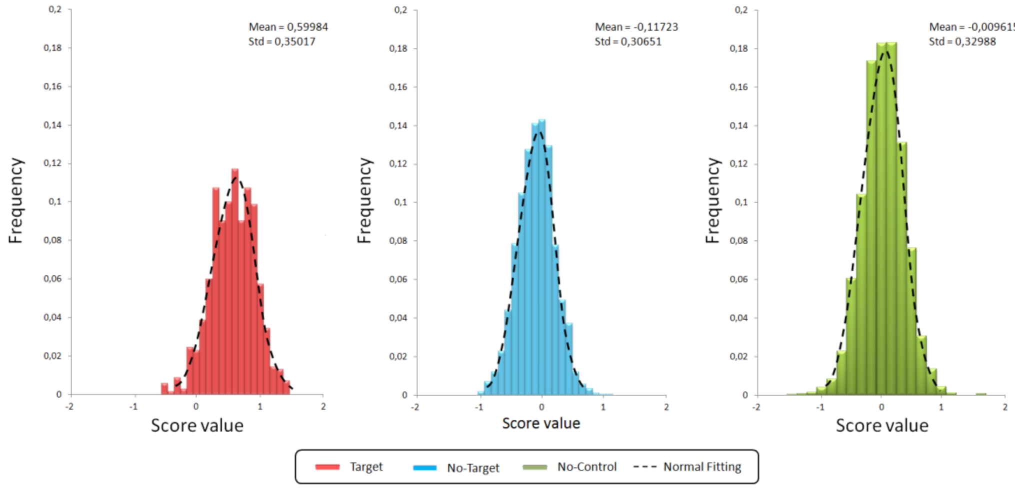

Figure 4.4 Score distributions for a representative subject. (a) Target score distribution; (b) No-Target score distribution; (c) No-Control score distribution. T-test results have shown that the hypothesis of normal distribution is true with 95% confidence level. The dotted-line denotes normal fitting

As it can be seen in Figure 4.4, a normal distribution well fits the score distributions. In fact, we ran a t-test on the three different score distributions for each subject, and its results have shown that the hypothesis of normal distribution is true with 95% confidence level.

Next, a Kolmogorov-Smirnov test was performed on each pair of sample. The value of the statistic test and the corresponding p-value are reported in Table 4.1. The hypothesis of different distribution was confirmed with the 95% confidence level for all the subjects except for the subject 1. For this reason, it is necessary to take into consideration the No-Control trials to estimate control parameters and thresholds.

Table 4.1 Kolmogorov-Smirnov's test values comparing pair wise distribution of Target, No-Target and No-Control score distribution. The hypothesis of different distribution was confirmed with the 95% confidence level for all the subjects except for the subject 1.

	Target vs NoTarget		Target vs NoControl		NoTarget-NoControl	
	ks-value	*p*-value	ks-value	*p*-value	ks-value	*p*-value
SUBJ 1	0.64	<0.001	0.63	<0.001	0.02	0.44
SUBJ 2	0.58	<0.001	0.53	<0.001	0.06	<0.001
SUBJ 3	0.74	<0.001	0.65	<0.001	0.14	<0.001
SUBJ 4	0.68	<0.001	0.61	<0.001	0.10	<0.001
SUBJ 5	0.66	<0.001	0.57	<0.001	0.10	<0.001
SUBJ 6	0.74	<0.001	0.65	<0.001	0.13	<0.001
SUBJ 7	0.78	<0.001	0.70	<0.001	0.15	<0.001
SUBJ 8	0.62	<0.001	0.52	<0.001	0.14	<0.001
SUBJ 9	0.57	<0.001	0.47	<0.001	0.11	<0.001
SUBJ 10	0.62	<0.001	0.51	<0.001	0.15	<0.001
SUBJ 11	0.67	<0.001	0.67	<0.001	0.05	0.01

The threshold values were chosen according to the number of stimulation sequences accumulated in the trial. In fact, the scores for the general stimulus *i* at the sequence s will be defined as:

$$yacc_i^s = \sum_{n=1}^{s} y_i^n \quad i = 1, 2, \dots N_s \quad (2)$$

Where y_i^n is given by (1) for the sequence *n*, N_s is the number of stimulation classes for the domotic interface (N_s = 8) and the maximum number of stimulation sequences in a trial was fixed to 10. Figure 4.5 shows the box plot of scores distributions earned on the basis of accumulated sequence. The Target distribution deviates from the No-target and No-Control distributions as more as the number of accumulated sequences increases while difference between No-Target and No-Control remains quite the same with the addition of new sequences.

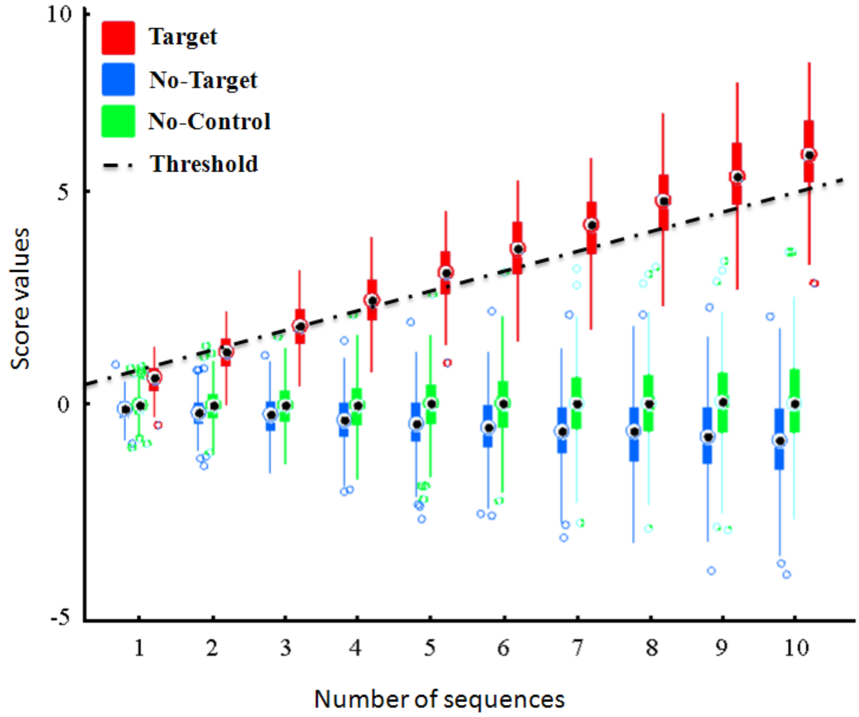

Figure 4.5 Distributions of Target, No-Target and No-Control accumulated scores. Each Box Plot is related to the number of sequences elapsed in a Trial. The dotted line denotes the thresholds trend. The difference between Target distributions and No-Target and No-Control ones increases with sequences accumulated in the Trial

Afterwards, the maximum score of the row stimuli and the maximum score of the column stimuli were assessed for each sequence, then it was assigned to them a label equal to 1 if the maximum scores were relative to a target stimulus and equal to 0 if they referred to No-Target or No-Control stimuli. In this way the maximum score values related to No-Control trials were included in ROC curves training, so that threshold values took into account possible artifacts that could occur when the subject was not engaged in BCI control. From this point on ROC curves could be plotted for each sequence using the corresponding scores. In Figure 4.6 there is an example about ROC curves trend. It is evident that when the number of the elapsed sequences in the trial increases the ROC curves assume an ideal tendency. To choose the threshold value it is necessary to find a tradeoff between false positive rate (FPR) and true positive rate (TPR). The maximum FPR was set to 0.05 and the lowest TPR to 0.5, so that the threshold will be chosen at the intersection of the ROC curve to the straight line joining points (0,1) and (0.05 ,0.5). The choice of these values was based on empirical considerations. The aim was to control home automation, so the specificity (low FPR) was preferred with respect to sensitivity. In any case, as it can be seen from Figure 4.6, this affects threshold values only for the firsts sequences in the trial, because as the number of sequences accumulated grows the ROC curves tend rapidly to the ideal trend.

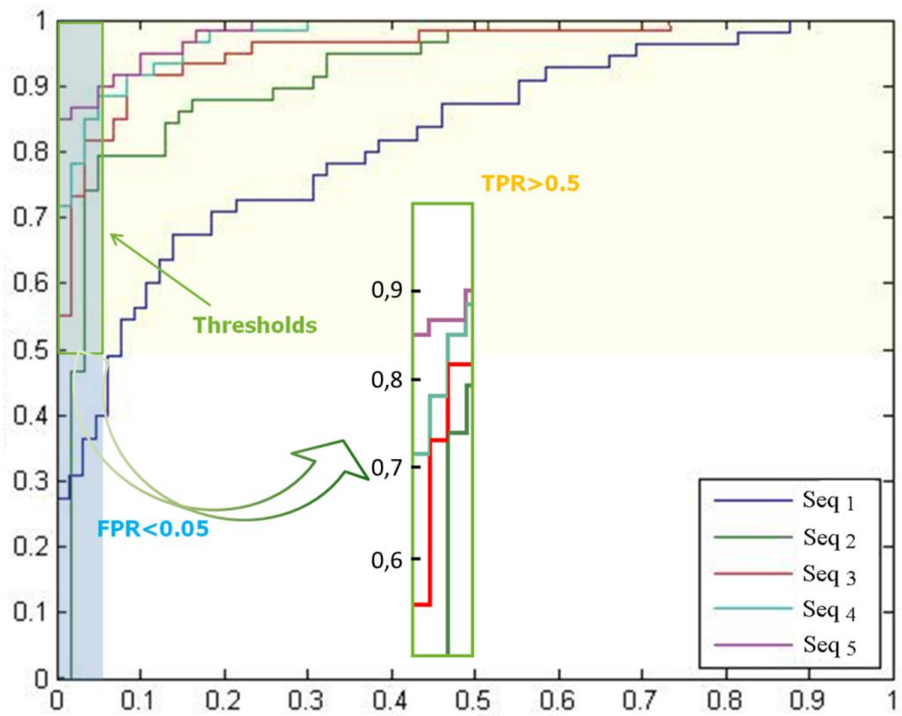

Figure 4.6 Thresholds extraction process. On the right Roc Curves plotted as a function of the number of sequence elapsed. The ROC curves show a trend closer to the ideal when the number of sequences accumulated in the Trial increases. The box on the left gives a zoom of the area representing the chosen tradeoff between FPR and TPR values. Threshold values are taken at the intersection of the ROC curve to the straight line joining points (0, 1) and (0.05 ,0.5).

The classification process in the BCI2000 framework was modified. The score values were computed at each new sequence and accumulated to the previous ones in the current trial. Threshold values were related to the number of elapsed sequences in the trial. At the end of each sequence, the maximum rows and columns values were compared to the specific threshold. If threshold exceeded simultaneously because of the maximum row and column values, the system classified the icon at their intersection. Conversely, if the threshold values did not exceed throughout the maximum number of stimulation sequence fixed a priori (Reset-Value), the system refrained from making a selection.

4.1.4 Results

4.4.1.1 *Synchronous System Off-line performances*

The Control runs data from training sessions were used to evaluate synchronous system performances. A cross-validation was operated using 2 runs to extract significant features and the other 2 to test them; the results of classification for each possible combination of training and testing data set (a 6 rounds cross-validation) were averaged. The trials used for both train and test were almost half of the ones used for Asynchronous system, even if 16 trials were enough to train SWLDA for P300-based BCI (Guger et al., 2009). Instead a greater amount of data was necessary to achieve good resolution in ROC curves plotting; this is the reason why the Asynchronous data set was larger than the Synchronous one.

Figure 4.7 shows the trend of the percentages of correct classification based on the number of stimulation sequences averaged for each subject. The black dotted line represents an accuracy of 95% in according to the value of false positives set in the asynchronous system through ROC curves.

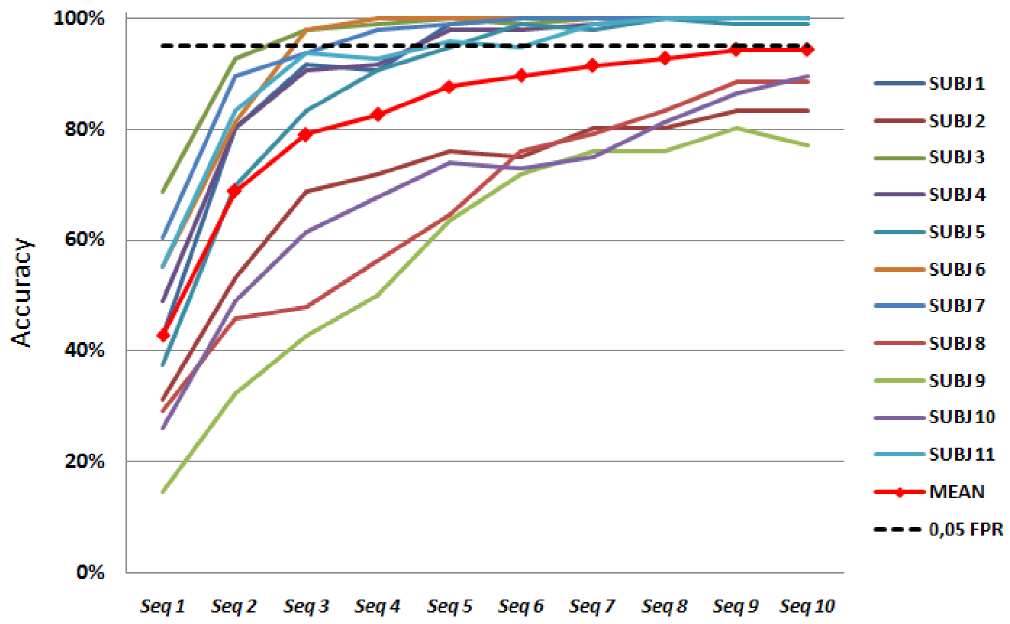

Figure 4.7 Off-line cross validation for synchronous system using data from Control trials in the training sessions. Accuracy values bases on the number of stimulation sequences averaged for each subject.

These results were used to set the maximum number of sequences within the synchronous runs and the Reset-Value for the asynchronous runs during the On-line session. In the synchronous mode it has been set the number of sequences equal to the number of stimulation sequences needed to reach 95% of accuracy. If it failed the maximum number of sequences was left to 10. For the asynchronous system the Reset-Value was chosen according to the maximum number of sequences needed to take a selection during Control periods (see Table 4.2). At the same time, the accuracy values and the number of sequences for both synchronous and Asynchronous system was used for bit-rate evaluation in the Information Transfer Rate section.

4.4.2.1 *Asynchronous system Off-line performances*

An off-line 6 rounds cross-validation was performed on the data acquired during the Alternate trials in the training session. 6 different training data sets were used to extract features defining threshold value (3 Fixation Cross, 2 Watch & Listen and 2 Computation Alternate runs), and the complementary testing data sets to assess off-line performances for the asynchronous system. A 6 rounds cross-validation was performed in order to match the maximum number of rounds achievable with the Control Runs data. Depending on the user's state (Control or No-Control) there could be five different classification outcomes.

Control state classification outcomes:

- Correct Classification: the target was correctly recognized;
- Wrong Classification: there was a target misclassification;
- Missed Classification: the thresholds were never exceeded, and the system abstained from taking a decision.

No-Control state classification outcomes:

- Abstention: the system properly refrained from taking decisions
- Missed Abstention: the thresholds were exceeded and the system made a wrong choice.

Figure 4.8 reports cross-validation results: it can be seen how the system was proved to be robust in avoiding false positive during No-Control trials, in fact Abstentions reached on average 98.61%. On the other hand, the 11.39% on average of Missed Classification represented the system's ability to avoid misclassification, because the percentage of wrong classification did not exceed an average of 1.5%. Table 4.2 reports the maximum, mean and standard deviation values of the number of stimulation sequences needed for each subject to exceed thresholds during Control trials.

Table 4.2. Number of stimulation sequences needed to exceed thresholds in the asynchronous modality

NSeq	**S 1**	**S 2**	**S 3**	**S 4**	**S 5**	**S 6**	**S 7**	**S 8**	**S 9**	**S 10**	**S 11**	**MEAN**
Max	9	10	8	9	9	10	8	9	10	9	8	9
Avg	3.61	4.63	3.08	3.95	3.52	4.70	3.31	4.80	5.56	4.42	3.31	4.08
Std	2.07	2.31	1.78	2.17	1.98	2.69	1.56	2.12	1.97	2.15	1.72	0.79*

* Standard deviation of inter-subject mean values

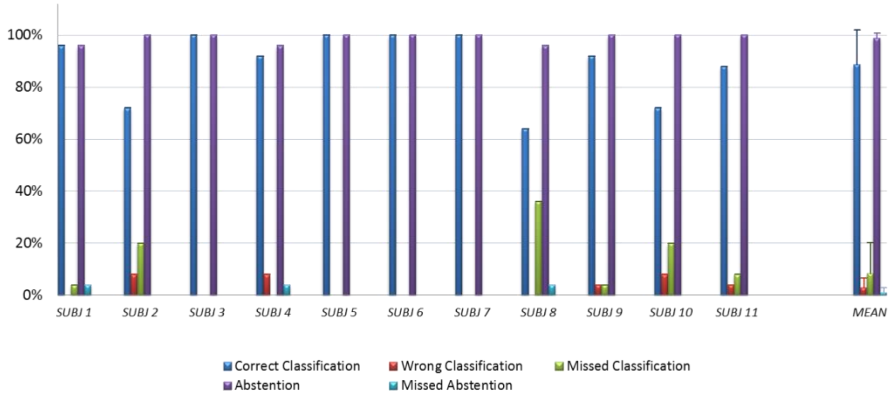

Figure 4.8. Results of off-line cross-validation for the asynchronous system using data from Alternate sessions. Results refer to Control and No-control trials. Vertical bars on the mean values denote the standard deviation of subjects performance. The system demonstrated high reliability during No-Control trials (Abstention mean = 98.91%) and at the same time on average the 88.73% of Control Trials was correctly classified. The error bars on the mean values denote the inter-subject variability

4.4.3.1 *Information Transfer Rate*

In order to assess the efficiency of the two systems in terms of an information transfer rate, it was used the definition of bit rate given by Wolpaw et al. (2000) and widely used in BCI community. The latter is based on the definition of information rate proposed by Shannon for noisy channels with some simplifying assumptions; all of the symbols have the same *a priori* occurrence probability $p = 1/Ns$, the classifier accuracy P is the same for all target symbols and the classification error $1–P$ is equally distributed amongst all of the remaining symbols:

$$B_{Wolpaw} = \log_2 N_s + P \log_2 P + (1-P) \log_2 \frac{1-P}{N_s - 1} \quad (3)$$

This expresses the bit rate or bit/trial for each selection. The information transfer rate (bits per minute) is equal to B_{Wolpaw} multiplied by speed of selection S (selection per minute). In turn, the speed selection for the P300-based system depends on the number of stimulation sequences used, and, in this case the 4s between two Trials have been taken into account. Table 4.3 shows the values of the information transfer rate for each subject calculated, using the number of sequences and the percentage of accuracy obtained by off-line analysis. For the asynchronous system, only the results from the Control trials were considered. A t-test was run to evaluate the differences between the two distributions. It did not show any statistical significance between the two distributions ($t=-0.81$, p-value =0.62); however, the asynchronous system exhibited on average an information transfer rate higher than the synchronous one.

Table 4.3 Information Transfer rate for each subject evaluated for synchronous system and for the asynchronous one. These values refer to Off-line analysis.

Information Transfer Rate (bit/min)												
	S 1	**S 2**	**S 3**	**S 4**	**S 5**	**S 6**	**S 7**	**S 8**	**S 9**	**S 10**	**S 11**	**MEAN**
SYNC	12.39	4.66	16.81	12.01	11.03	16.81	14.01	5.48	3.94	5.63	11.34	10.37
ASYNC	14.55	5.60	17.06	12.62	13.99	11.51	14.95	4.15	8.69	6.54	13.41	11.19

4.4.4.1 *On-line Results*

As mentioned before, during the On-line session subjects were asked to manage some devices using the BCI in both synchronous and asynchronous mode.

Figure 4.9 reports the total time needed by each subject to complete the 2 Control runs. Results are on average consistent with off-line bit rate values but some subjects (3, 4 and 8) exhibited different performances with respect to those expected from the off-line analysis. T-test did not present significant difference between the 2 distributions (p-value = 0.74; t = 0.33).

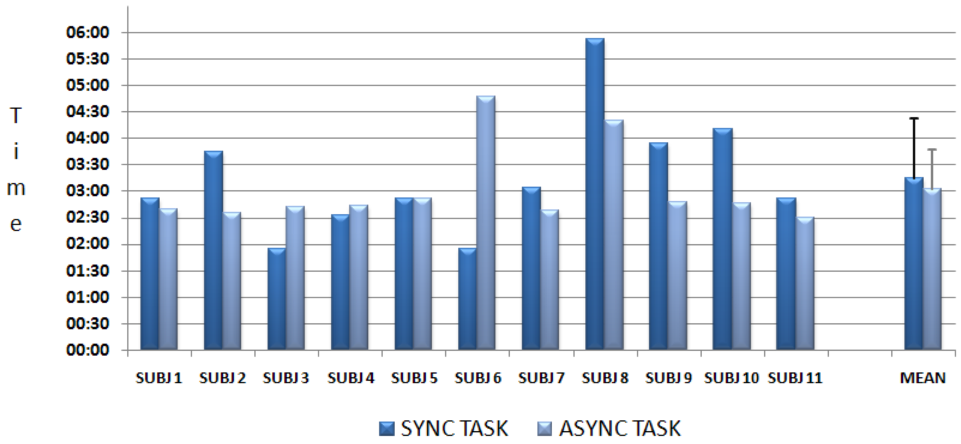

Figure 4.9 Results for Comparison in the On line sessions. Time needed to complete ten different actions with the synchronous and the asynchronous system. Values refer to the sum of time needed to complete Control runs 1-2 for synchronous system, and Control run 3-4 for asynchronous one.

With regard to the 2 No-Control trials in the On-line session, on average 0.26 false positive/min (std = 0.4) were detected. Furthermore each subject was able to complete the real-life runs. On average, over the 2 *scenarios*, there were 1.73 false positives (std = 2.14) during No-Control actions, while during Control actions the asynchronous system achieved on average 87.46 % (std = 13.34%) of Correct Classifications.

4.2 Evaluation with potential end users

4.2.1 Participants

Seven elderly (4 males, 3 females; mean age=64.85 ± 5.81 years) individuals who are clients of the Frisian home care organization (THFL) joined the recording protocol. Four of them suffered from chronic neurological disorders: one affected by amyotrophic lateral sclerosis (ALS), two by multiple sclerosis (MS) and one by stroke. The degree of functional motor impairments was defined on the basis of the ALS Functional Rating Scale-Revised (ALSFRS-R – Cedarbaum et al., 1999), the Kurtzke Expanded Disability Status Scale (EDSS – Kurtzke, 1983), the Rankin Scale for Stroke Disability (RSSD – Rankin, 1957). In addition, the Barthel Index (BI – F I

Mahoney and Barthel, 1965) was administered to all participants to estimate a global degree of independence in performing daily activities. Functional scores revealed a moderate to severe motor disability (EDSS score = 3 for MS patients; ALSFRS-R score = 10 for ALS patient; RSSD score = IV for the stroke patient; BI 85.71±19.24; n=7). Cognitive functions were preserved in all of them as indicated by the MMSE scores (between 27-30 – Folstein et al., 1975). Each participant had previous experience in exploring the virtual environment (see below) by means of classical input devices such as mouse or joystick. All of them were naïve to the BCI control.

4.2.2 Experimental Setup

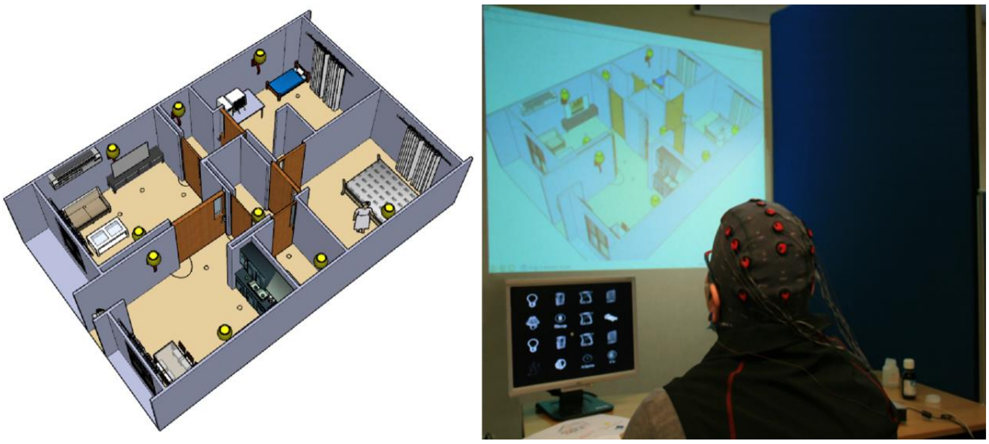

Figure 4.10 A) Illustration of the virtual environment. The users can operate lights, doors, curtains, windows, bed, TV, air condition, alarm and pause applications by means of the BCI system. B) A moment of the experimental session. The user's feedback consisted of the actuation of the corresponding device in the virtual apartment.

The domotic environment to be controlled was based on a virtual reconstruction of a real apartment that is built at the premises of the Fondazione Santa Lucia in Rome (see Figure 4.10A). The apartment consisted of four rooms: two bedrooms, a kitchen and a living room, and the devices operable by means of the BCI were lights, doors, curtains, windows, bed, TV, air conditioning, SOS and "sleep macro." The "sleep macro" arranges the environment in sleep modality (e.g., turns off the lights, closes the curtains). Detailed description of the environment software simulator is reported in

Kaldeli et al. (2010). In the proposed experiments, a beamer was used to present the apartment in a bird-eye's view to the participant; a laptop processed the instructions coming from the BCI software. Figure 4.10B shows a moment of the actual experimentation. The acquisition protocol was based on the Speller paradigm (Farwell and Donchin, 1988), implemented in the P3Speller application within the BCI2000 framework (Schalk et al., 2004). Stimuli were provided by row and column flashing in a matrix and a static fixation cross was placed in the middle of the interface. The first recording session consisted in a text input task, and for this reason we used a 6 by 6 matrix containing alpha numeric items. For the second recording session, which concerned control of the domotic appliances the Speller was reduced to a 4 by 4 matrix and letters were replaced with the icons representing the devices available in the virtual apartment. In order to perform the asynchronous control, we used a modified version of the P3Speller application (Aloise et al., 2011). When the BCI system recognized a Target, a message was sent to an external application that parsed the information and generated a call request using the XML based Simple Object Access Protocol (SOAP) protocol to the virtual environment (Warriach et al., 2010). Finally, the software provided a feedback to the user through the animation of the selected device in the virtual environment. For instance, commanding the closing of a curtain resulted in the actual movement of the curtain fabric in the virtual environment. Scalp EEG potentials were recorded from 16 positions according to the 10-10 standard (Fz, FCz, Cz, CPz, Pz, Oz, F3, F4, C3, C4, CP3, CP4, P3, P4, PO7, PO8) with g.LADYbird active ring electrode using the g.USBamp amplifier (g.tec medical engineering GmbH, Austria). The EEG was sampled at 256 Hz. Each channel was referenced to the left earlobe and grounded to the right mastoid. Stimulation was provided to the subjects using a 17''LCD monitor placed at 1 meter of distance from him/her.

4.2.3 Recording Protocol

The purpose of the first recording session was to investigate the subjects' ability to control a P300-based BCI and to allow them to familiarize with the system. Each participant underwent two recording sessions on different days. In the first recording session (Speller), the subjects were asked to perform 5 runs. During the second session (Domotic), in order to train and test the asynchronous classifier, subjects performed also No-Control runs for a total of 10 runs. A run consisted of 5 trials and every trial was composed of 12 stimulation sequences (each composed by the single flash of each row and column on the interface). Each stimulus was intensified for 125ms, the inter-stimulus interval (ISI) was set at 125ms.

4.3.1.2 *Speller session*

The Speller session consisted of 5 Copy mode runs in synchronous mode, i.e., the user was cued with a target letter to concentrate on, and a fixed-length train (12 sequences) of random stimuli was presented. Subjects were asked to spell five common words: WATER, WATER, KOPJE, BROOD and KLEIN (these words are in Dutch and mean water, water, cup, bread and small). The word WATER was repeated twice as previously reported by Guger et al., 2009. In order to extract the most significant features we applied a Stepwise Linear Discriminant Analysis (SWLDA – Krusienski et al., 2006) on the first two Copy runs, during which subjects did not receive any feedback regarding classification outcome. During the last three runs, the selected letter was presented to the subject at the end of each trial. Since on-line results referred to 12 stimulation sequences, an off-line cross-validation was performed to provide accuracy as a function of stimulation sequences. In particular 10 crossvalidation rounds have been performed using all possible combinations of 2 runs to extract control parameters, whereas the remaining 3 runs were used for testing them.

4.3.2.2 *Domotic session*

The Domotic session was organized in two parts; synchronous and asynchronous applications. The synchronous part consisted of 4 Copy mode runs and 2 NoControl runs. During each Copy mode run the subjects had to operate five different devices in the virtual environment. The system suggested the Target icon at the beginning of each trial in a pseudo random order, ensuring that all the items of the matrix were presented at least once. The first two runs were used to extract control parameters through

SWLDA. The parameters estimated from the Speller session were not used since the matrix was different in size, and thus implied changes in the Target to Target Interval (TTI). The TTI can affect P300 morphology and, as a consequence, the control parameters (Gonsalvez and Polich, 2002). From the third run, feedback about the outcome of classification was provided to the user: at the end of every trial the selected icon was intensified and the subjects could see the corresponding device change its state in the virtual environment. This occurred within 5 seconds after each trial. The inter-trial time was extended to 10 seconds in order to avoid artifacts due to subject's movements looking at the image of the virtual apartment. As for the data acquired in the Speller session, a 6 rounds cross-validation was performed on the Domotic session data: data relating to the copy mode runs were divided into a training data set composed of two runs and a testing data set including the remaining runs. During NoControl runs no target icons were provided to the subjects who were required to ignore the stimulation and execute two different No-Control Tasks: during the No-Control Task 1 the subject had to gaze at a fixation cross in the middle of the interface while the stimulation was on, during the No-Control Task 2 the subject had to gaze at the fixation cross and talk with the operator (if he/she was able to verbally communicate) while the stimulation was on.

The asynchronous part was composed of 1 Control run and 2 NoControl runs operated by the asynchronous classifier. The asynchronous system is based on the introduction of threshold values in the on line classifier: at the end of each sequence, the maximum row and column score values were compared to the specific threshold. If the threshold was exceeded because of the maximum row and column values, the system classified the icon at their intersection. Conversely, if the threshold values did not exceed the maximum number of stimulation sequence fixed a priori (reset value), the system refrained from making a selection. After the reset, the system set to zero the scores values accumulated for each stimulus class, and a new trial began. Features and thresholds were extracted using data acquired during the synchronous part. In order to make the system robust to possible artifacts that can occur during No-Control tasks, the SWLDA was applied also on the No-Control data which were labeled as a NonTarget. Threshold extracting relies on ROC curve plotting of score values, and the latter were dependent on the number of stimulation sequences accumulated in a trial

(Aloise et al., 2011). During the Control run, subjects were asked to operate five different devices in the environment using the BCI. At the beginning of each trial, the operator suggested the icon on which the user had to focus. There could be three different classification outcomes: 1) correct classification: the system correctly recognized the target icon; 2) misclassification: the system selected an unwanted item; 3) abstention: the thresholds were not exceeded throughout the reset value of stimulation sequences, so a new trial began without selections. Subjects had 10 minutes to complete the task, otherwise the task was considered incorrect. If an unwanted abstention or a misclassification occurred, the operator invited the subject to again select the desired icon. This run allowed quantification of the accuracy of the asynchronous system when the subject intends to operate a control on the environment. In order to quantify the robustness of the asynchronous system with respect to false positives, subjects performed two No-Control runs lasting 2.5 minutes, during which subjects repeated the two No-Control tasks.

4.2.4 Intra-subject variability analysis

In order to quantify intra-subject variability of the stimulation sequences needed to achieve a correct classification for end-users, the data acquired during the domotic session were used. In particular, from the 6 rounds off-line cross-validation, it has been collected the number of stimulation sequences at which the correct classification was achieved. The results obtained from data acquired in this study were compared with those obtained from the dataset collected from the 11 healthy young subjects of the previous section. This latter Control dataset related to 4 Copy mode runs of 8 trials each in which subjects were engaged in a similar experimental task. The stimulation modalities were the same as the actual study: stimuli were provided through a 4 by 4 matrix containing 16 black and white icons (Stimulus Duration = 125ms, ISI = 125ms), representing some device that can be really operated by BCI (see Figure 4.1). The icons presented in this interface are slightly different with respect to the icons used in the previous study. In order to make the two datasets comparable, the end-users' dataset was reduced to 8 EEG acquisition channels (Fz, Cz, Pz, Oz, P3, P4, PO7, PO8) and it were considered only the first 10 stimulation sequences as it was in the Control dataset. Also for the latter only the first 5 trials of each run were used.

4.2.5 Results

4.5.1.2 *On-line results*

Figure 4.11A illustrates online results of the Speller and Domotic sessions. The bars represent the mean accuracy reached during the runs in Copy mode with on-line feedback. During the Speller session, all subjects except Subject 3 exceeded the level of 80% accuracy. During the Domotic session mean accuracy was lower than in the Speller session. Subject 3 improved his performance with respect to the Speller session. Results related to the control run operated by the asynchronous application are presented in Figure 4.11B. The asynchronous system was strong in avoiding misclassifications because they occurred only with Subject 3 (7% of errors), in all other cases the misclassification rates were zero. Three subjects achieved 100% correct classifications consistent with the results obtained in synchronous mode. On average there were the 26.33% of abstentions, but during the synchronous part of the domotic session there were on average 18.57% wrong classifications while in the asynchronous run error rate dropped to 1%. For the 2 NoControl runs in asynchronous modality, the system exhibited strong reliability in avoiding false positives when the subjects were engaged in other tasks. If thresholds were not passed, 2 abstentions could be collected in a minute; on the contrary, if the thresholds were incorrectly exceeded after the first stimulation sequence a maximum of 30 false positives could occur in a minute not considering the inter-trial interval. On average 0.225 false positives/minute were collected. This value is comparable to the 0.26 false positives/minutes achieved in the previous study with the control subjects.

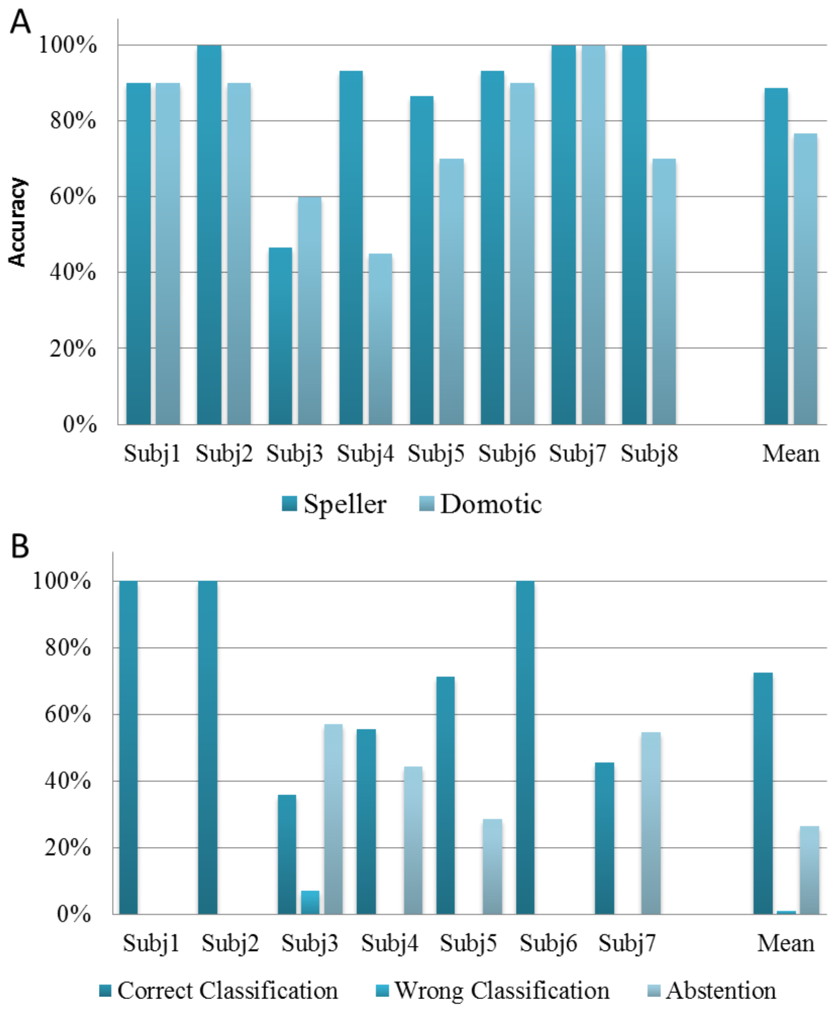

Figure 4.11 A) On-line classification accuracy reached from each subject during the Speller session and the Domotic session, respectively. Subject's error bars denote the standard deviation values of subject's accuracy among the runs considered (3 runs for Speller session and 2 runs for Domotic session). The error bars on the mean values denote the performance standard deviation among subjects. B) On-line results for the Control run with the asynchronous classifier

4.5.2.2 *Off-line results*

Figure 4.12A and B show the outcome of the 10 and 6 rounds crossvalidation performed for Speller and Domotic datasets, respectively. A t-test was used on the accuracy values reached at the 12th stimulation sequence obtained on-line and off-line for the Speller and Domotic sessions. It did not show statistical difference between the on-line and the off-line accuracy (Speller: t-value = 0.082, p-value = 0.93; Domotic: t-value = 0.90, p-value = 0.38). The box plots Figure 4.13A and B represent the distribution of the number of stimulation sequences needed to achieve the correct classification for end- and control-users. Distributions are the outcomes of the 6 rounds cross-validation. It can be seen that end-users exhibited a higher intra-subject variability of stimulation sequences with respect to the control-users. This was confirmed performing a t-test (α=0.05): the two standard deviations distributions of the number of stimulation sequences for end-users and control-users were statistically different (p-value = 0.04, t-value = 2.22); in particular, variability in the end-users distribution (mean value = 2.1 sequences) was higher than in the control (mean value = 1.53 sequences). Figure 4.13C illustrates the mean value and the standard deviation of the number of stimulation sequences to achieve correct classifications for each subject during the online asynchronous control run. The standard deviation shows that the asynchronous system was able to adapt its speed of selection to the intra-subject variability, which typically is in the range of two sequences.

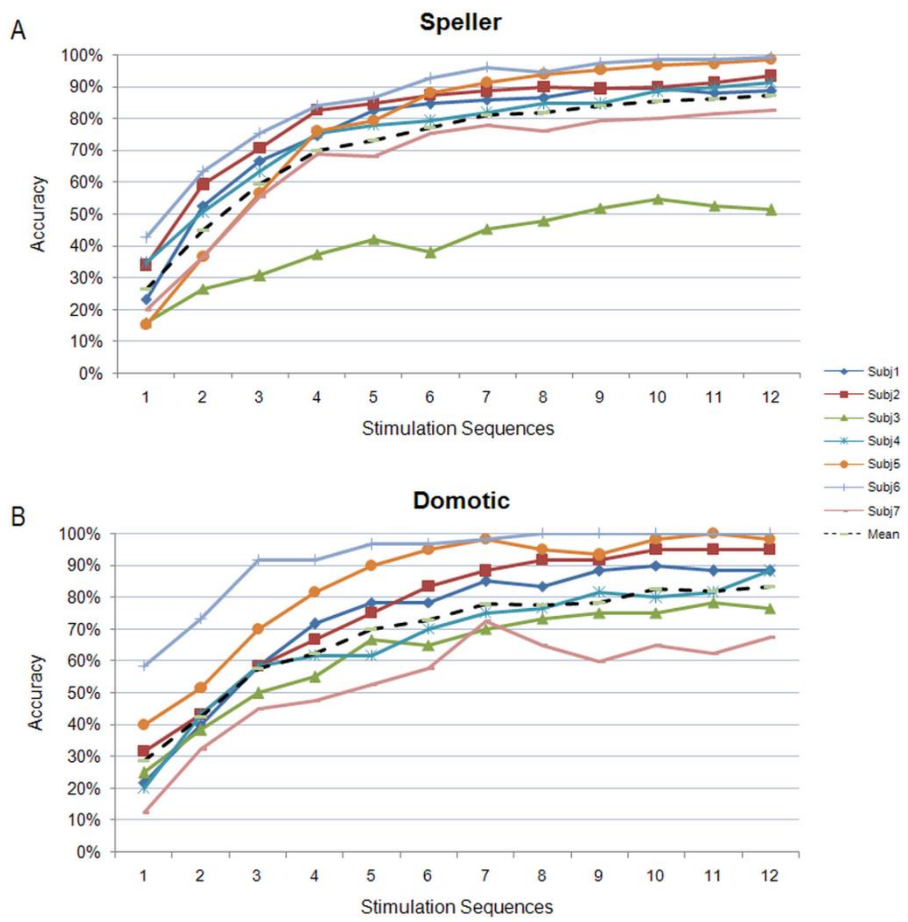

Figure 4.12 Accuracy as a function on the number of stimulation sequences obtained through off-line cross-validation. A) 10 rounds cross-validation outcome for the Speller session; B) 6 rounds cross-validation outcome for the Domotic session

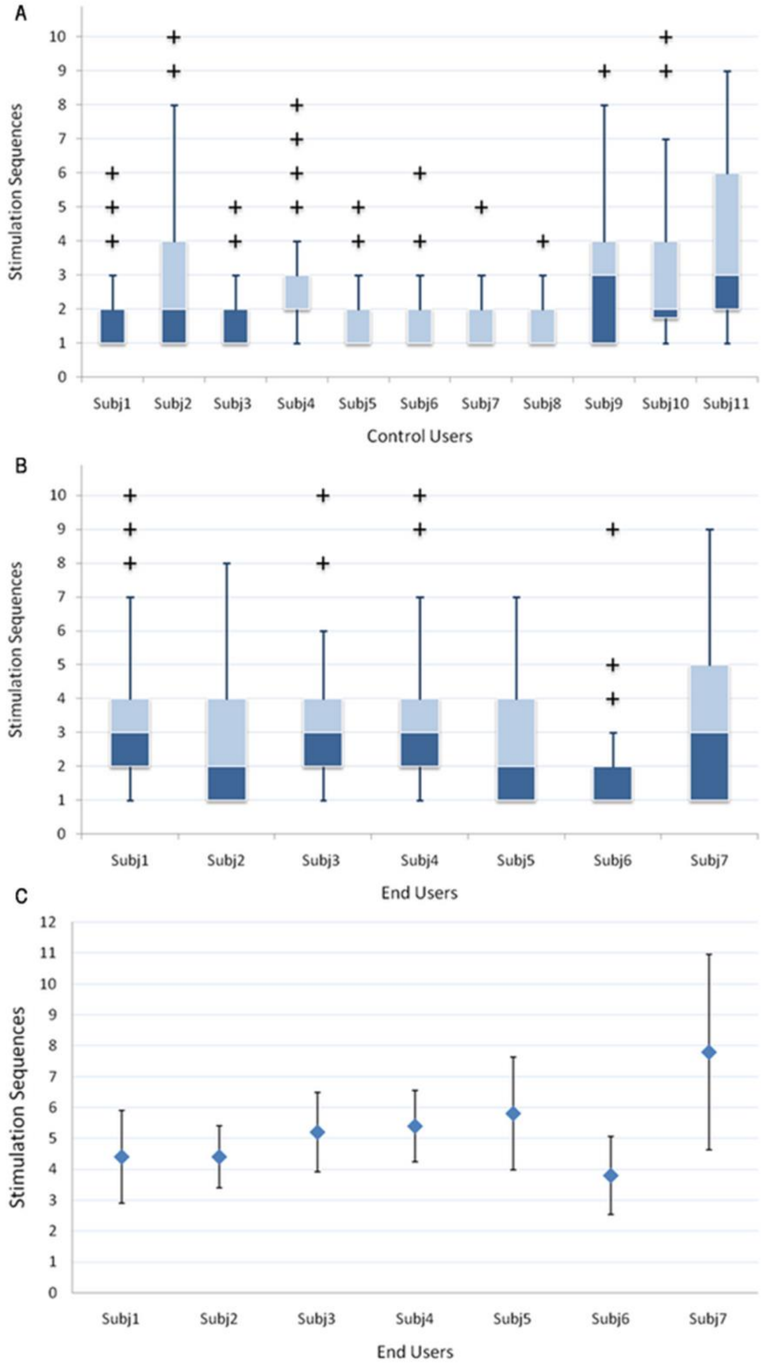

Figure 4.13 Boxplots A) and B) represent standard deviation distributions of the number of stimulation sequences necessary to each subject to achieve a correct classification. A) Control-users B) End-users C) Number of stimulation sequences needed for a correct on-line classification during the Asynchronous control run with end-users

4.3 Communication efficiency assessment

This section aims at demonstrating how the efficiency of communication, evaluated in terms of speed of symbols output and time needed to recover from errors, can actually been increased by an asynchronous classifier. In order to generalize results both a dataset relating to environmental control (database A) and a dataset relating to a spelling task (database B) in different operation conditions have been considered and global performance were compared to the performance of a classical synchronous P00-based BCI.

4.3.1 The efficiency metric

In order to evaluate the efficiency of communication, the metric of Bianchi et al., (2007) and recently validated by Quitadamo et al. (2011) has been applied. This metric predicts the extent to which the accuracy of classification can support communication – i.e., whether the time that is spent in correcting mistakes is shorter than that needed to generate a correct selection. With respect to the written symbol rate (WSR - Furdea et al., 2009) and the Wolpaw's bit-rate metrics (Wolpaw et al., 2000), which take into account recognition accuracy, the efficiency metric allows one to consider the various costs of abstentions and errors – a key point in the correct evaluation of the performance of an asynchronous system. The method by which communication efficiency was evaluated has been reported (Bianchi et al., 2007). Briefly, an initial extended confusion matrix (ECM) is computed, comprising an N by N +1 matrix, where N is the number of the available symbols. The additional column indicates when the classifier abstains from making a decision. To estimate the probability that the classification of a symbol is incorrect or undetermined, a misclassification probability matrix (MPM) can be defined from the ECM. The extended overtime matrix (EOM) is built, representing the costs that are associated with errors and abstentions with regard to additional steps that must be taken to correct mistakes. In this study, the following assumptions about the costs have been made: a cost of 1 was associated with abstentions (the user needs only to repeat the trial, trying to reselect the desired character) and a cost of 2 with misclassifications, considering that in this case, the subject must first delete the incorrect character and reselect the desired one. The latter assumption is also valid if the desired symbol is misclassified with the UNDO item, leading to deletion of a correct symbol. A super tax (ST) vector can be defined as:

$$ST[i] = \sum_{j=1}^{N+1} MPM[i,j] \times EOM[i,j]$$

Where *i* denotes the desired class and j indicates the predicted class. Considering that all symbols in the matrix are equally probable, the mean expected selection cost (ESC) can be defined, which is the mean number of selections that is needed to generate a correct symbol, taking into account the recovery from errors and abstentions.

$$\overline{ESC} = \sum_{i=1}^{N} \frac{1}{1 - ST[i]}$$

The efficiency of a system, with regard to the time that is needed to achieve a classification, is expressed as a function of the number of stimulation sequences (*NumSeq*):

$$Eff = \frac{1}{NumSeq \times \overline{ESC}}$$

4.1.1.3 *Database A: environmental control*

Eleven healthy volunteers (4 females, 7 males; mean age 26.4 +/- 4 years) were involved in this part of the study. The acquisition protocol was based on the P300 Speller interface (Farwell and Donchin, 1988) adapted to control a home automation system by using a 4 by 4 matrix containing 16 black and white icons representing the available actions on the environment (see Figure 4.1). Stimulation and data acquisition was managed by the BCI2000 framework (Schalk et al., 2004). Stimuli consisted in the intensification of rows and columns of the matrix. Each stimulus was intensified for 125ms with an inter stimulus interval (ISI) of 125ms. Scalp EEG potentials were recorded (g.MobiLab, gTec, Austria, sampling rate 256 Hz) from 8 scalp positions

(Fz, Cz, Pz, Oz, P3, P4, PO7 and PO8). Each channel was referenced to the linked earlobes and grounded to the left mastoid. The term Sequence indicates a complete cycle of intensification of each row and column. Ten Sequences make a Trial. For each subject 4 Control runs were acquired, made of 8 Control trials, and 12 Alternate runs during which Control and No-Control trials alternated for a total of 10 trials per run. During the No-Control trials the subjects voluntarily diverted their attention from the stimulation performing three simple no control tasks:

- Fixation Cross, 30 trials: Subjects were instructed to fixate the cross in the center of the interface and to ignore the stimulation;
- Watch & Listen, 15 trials: Subjects were instructed to watch a movie displayed on the half of the screen beside the matrix;
- Computation, 15 trials: Subjects had to answer simple arithmetic questions posed by the operator while fixating the cross.

Control runs were used to assess the accuracy of the Synchronous system, by repeating 6 rounds of 2-fold crossvalidation. In each round, we used 16 trials to extract significant control features by Stepwise linear discriminant analysis (SWLDA) and 16 trials as testing set. A similar procedure was also applied to the alternate runs, to evaluate the performance of the asynchronous system. In each round, the training dataset was composed of 35 Control trials, and 35 No-Control trials (15 Fixation, 10 Watch & Listen and 10 Computation), while the testing dataset was composed of the 25 remaining Control trials.

4.1.2.3 *Database B: copy spelling*

Nine healthy subjects (5 females, 4 males mean age = 26.4 ± 4.4) were enrolled in the study. All of them had previous experience with P300 based BCI and the GeoSpell interface (Figure 4.14 - Aloise et al., 2012). The latter was designed to be operated in covert attention mode, so that it can be used also if ocular movements are impaired. In the GeoSpell interface the 36 alphanumeric characters of the Farwell and Donchin's Speller were redistributed on the vertices of 12 hexagons hereinafter defined as groups or stimulation classes. Each character belongs to two groups, in which it is displayed on the same vertex. A fixation cross was displayed in the center of the stimulation interface at all times. The distance between the fixation cross and each character was fixed so that the visual angle is lower than 1 degree. Stimulation consisted in the pseudo-random appearance of groups (stimulus duration 125ms and ISI of 125ms) and was provided by a modified version of the BCI2000 framework. Scalp EEG signals were recorded from 8 positions (Fz, Cz, Pz, Oz, P3, P4, PO7 and PO8; gUSBamp, gTec, Austria, sampling rate 256 Hz). Each subject performed 8 runs of 6 trials each. During the first 6 runs called Control Runs all the 36 characters of the GeoSpell interface were presented as Targets to the subject, who had to focus his attention on it mentally counting the number of its occurrences always gazing to the fixation cross in the middle of the interface. During a trial, 10 stimulation sequences were delivered. Also each subject performed 2 No-Control Runs. During the first No-Control run the subjects were required to fixate the cross in the center of the interface, trying to ignore the surrounding stimulations; during the last No-Control run subjects were also required to perform simple mathematic computations.

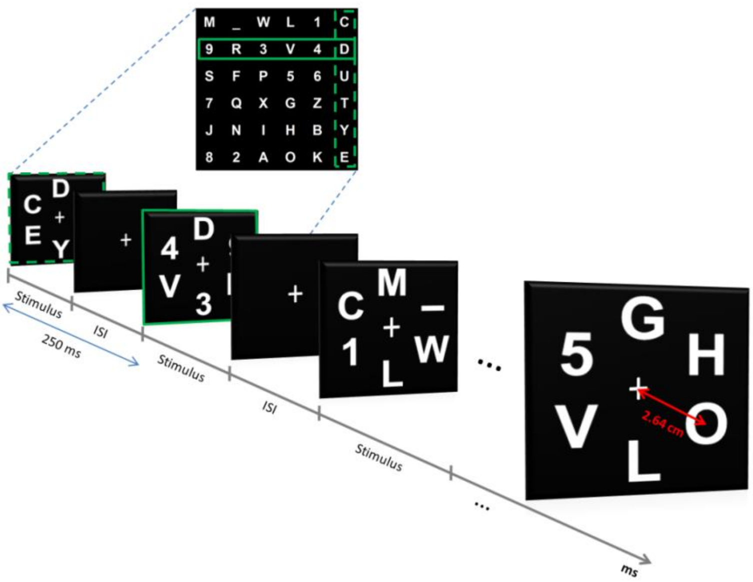

Figure 4.14 GeoSpell interface

To assess performance of the synchronous classifier, a 6-fold cross-validation was carried out using data from the 6 Control runs. Classification accuracy was then assessed as a function of the sequences accumulated in a trial in order to define the number of sequences to be used to define system efficiency. Regarding the asynchronous classifier, a 6-fold cross-validation was also performed. In this case, the 2 No-Control runs from the offline session were introduced in the training dataset and SWLDA was used to extract the control features.

4.3.2 Results

4.2.1.3 *Accuracy for the synchronous and the asynchronous classifier*

Figure 4.15 shows the performance of both synchronous and asynchronous classifier. When a character or an icon was correctly recognized and selected it was defined as a “correct”. If a classification occurred but the system selected an undesired character it was defined as an error, finally an abstention occurred when the threshold was not exceeded, and the latter was only possible for the asynchronous classifier.

Three 2-way ANOVA (CI=.95) were performed considering the paradigms (environmental control/copy spelling) and the classification mode (asynchronous/synchronous) as factors and corrects, errors, and number of stimulation sequences as dependent variables respectively. The synchronous classifier on average exhibited an higher percentage of correct classification with respect to the asynchronous one (93.27% ± 6.53 and 84.49% ± 11.27% respectively; $F(1, 36)=9.7813$, $p=.00348$), however the error rate was lower for the asynchronous classifier than for the synchronous one (2.85% ± 3.07 versus 6.73% ± 6.53; $F(1, 36)=5,.431$, $p=.0294$), since the former avoids errors through the abstentions (12.66% ± 10.48). Furthermore it should be considered that the number of sequences needed to achieve a classification was significantly lower for the asynchronous classifier (4.5 ± 1.06) than for the synchronous one (6.85 ± 2.56), as confirmed by the 2-way ANOVA (CI=.95) on the two distributions ($F(1, 36)=13.870$, $p=.00067$).

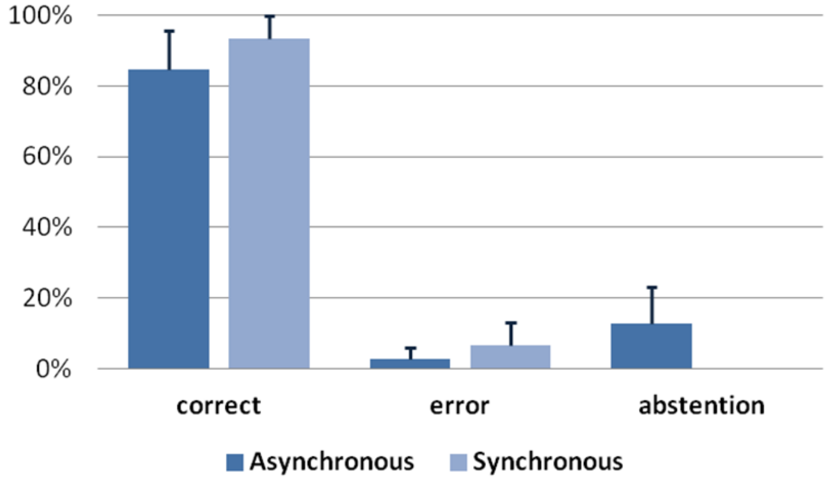

Figure 4.15 Offline performance of the asynchronous and the synchronous classifier

4.2.2.3 *Efficiency*

Errore. L'origine riferimento non è stata trovata. reports the communication fficiency values for the asynchronous and the synchronous classifier for both the considered tasks. The asynchronous control exhibited a higher efficiency (0.21 ± 0.06) with respect to the synchronous classifier (0.17 ± 0.08). However this difference was not statistically significant as assessed by a 2-way ANOVA considering the paradigms (environmental control/copy spelling) and the classification mode (asynchronous/synchronous) as factors and the Efficiency values as dependent variables ($F(1, 36)=3.4542$, $p=.07128$). Considering the Information Transfer Rate (ITR) assessed by the Wolpaw's metric (Wolpaw et al., 2000), which considers errors and abstentions in the same way, the asynchronous system exhibited an higher value (19,8 ± 9.19 bits/min) with respect to the synchronous classifier (17.3± 9.61 bits/min), but this difference was not significant as assessed by a 2-way ANOVA with paradigms and classification mode as factors and ITR values as dependent variables ($F(1, 36)=.72306$, $p=.40076$).

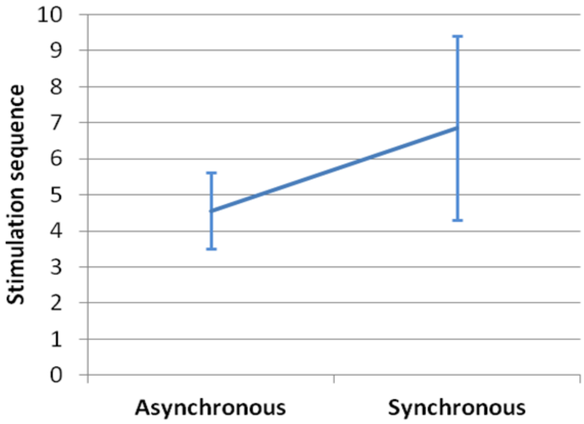

Figure 4.16 mean value of the number of sequences needed to achieve a classification with both synchronous and asynchronous classifier

Table 4.4 communication efficiency values for the asynchronous and the synchronous classifier

	Environmental control		**Copy spelling task**		
	Asynch	*Synch*		*Asynch*	*Synch*
Subj1	0.27	0.20	**Subj12**	0.18	0.25
Subj2	0.14	0.07	**Subj13**	0.18	0.20
Subj3	0.32	0.32	**Subj14**	0.21	0.17
Subj4	0.23	0.19	**Subj15**	0.21	0.14
Subj5	0.27	0.18	**Subj16**	0.17	0.12
Subj6	0.20	0.32	**Subj17**	0.20	0.11
Subj7	0.29	0.24	**Subj18**	0.18	0.14
Subj8	0.13	0.08	**Subj19**	0.15	0.11
Subj9	0.16	0.05	**Subj20**	0.37	0.25
Subj10	0.16	0.08	**Mean**	**0.21**	**0.17**
Subj11	0.27	0.18	**std**	**0.06**	**0.08**

4.4 Discussion

Understanding the user's intentions from the ongoing EEG such as when he/she wishes to suspend the control or when he/she recognizes an error represents an important issue which could improve usability and reliability of BCI system. To this aim, at the state of the art several classification algorithms have been proposed for P300 based BCIs (Jin et al., 2011; Lenhardt et al., 2008; Schreuder et al., 2011). However they only provides solution to dynamically adapt the number of stimuli repetitions and are not able to abstains from taking a decision if the user diverts his attention from the stimulation, or if the EEG signal is not reliable enough. While the statistical approach proposed by Zhang et al. (2008) provides this feature, it should be stressed that i) their test were carried out on a small number of subjects (4); ii) they reported a relatively low robustness to false positives during No-Control periods (0.71 false positives/min) and an Information Transfer Rate of 20 bits/min. Other asynchronous paradigms have been proposed based different control features, Panicker et al. (2011) combined P300 potential with Steady state visual evoked potentials (SSVEPs) for the detection of the control state reporting an ITR of 19.05 bits/min during control periods and false alarm rate of 4.2% during No-Control periods. Diez et al. (2011), with high frequency SSVEPs, reported an ITR varying from 9.4 to 45 bits/min. Zhang et al. (2012) recently proposed an asynchronous paradigm based on the N200 speller and the motion visual evoked potentials (mVEPs). The latter paradigm allowed to reach during on line tests on 9 healthy subjects 70,1% accuracy during control periods, while 2,38 false positives/min were detected during No-Control periods. The asynchronous classifier described in this section exhibited on average 0.26 false positives/min with healthy subjects and 0.225 false positives/min with potentials end users. Considering user needs and requests (Huggins et al., 2011) about 1 false positive every 4 minutes may still be considered unsatisfactory for a continuous use. However this value may be acceptable for short pauses such as waiting for an answer during a talk, or thinking about what we are going to write.

From the other side, abstentions may also occur during a control period, thus reducing the system's accuracy with respect to a classic synchronous classifier. As it was demonstrated in the communication efficiency evaluation section, considering that error recovery have a higher cost than abstentions, the asynchronous system exhibits higher communication efficiency because of its lower error rate with respect to the synchronous classifier. Moreover, with regard to the evaluation with potential end users, it is possible to hypothesize that ageing (even normal) would play a role in determining the level of intra-subject variability which in turn might affect the number of stimulation sequences needed to achieved a correct classification. In the case of young volunteers with a lower intra-subject variability, the feature of the asynchronous system to adapt its speed of selection to the current user state was not enhanced. In the evaluation with potential end-users, aged potential end-users showed a higher intra-subject variability in terms of time (number of stimuli) to achieve correct classification. Under this condition, a synchronous system would cause obvious uncertainty in deciding the number of sequences to be used for the online control of domotic and/or communication appliances. Choosing a higher value can improve performance in terms of system accuracy, but would also lead to a slower system. The asynchronous system can provide a solution by continuously adapting its speed to the most effective number of stimulation sequences, thus, maintaining high accuracy without lowering the system's bit-rate. The present findings point out the usability and reliability of an asynchronous BCI system for environmental control, indicating how these systems could be considered as input devices to interact with the external world and to restore the personal independence of people with severe motor disabilities.

To summarize, this section addresses the issue of an asynchronous P300-based BCI capable of understanding the user's intent and refraining from selections when the user is engaged in another task or is distracted by the surrounding events. The introduction of a threshold-based classification approach might allow the user to divert her/his attention from the control interface at any time and without the use of external inputs. A further advantage consists in increasing the accuracy of the system; an asynchronous BCI may prevent errors through abstentions. The advantages of the asynchronous classifier with respect to the synchronous classifiers used in experimental settings were first investigated by an experimental protocol involving healthy users and confirmed with potential end users both for environmental control and communication applications. Altogether, these features would allow for a more independent use of a BCI system by people with severe disabilities.

5 Self-calibration algorithm in asynchronous P300-based BCI

To reduce the gap between BCI systems and other alternative-augmentative communication (AAC) technologies BCI systems should ensure high reliability and should not require complex configuration and calibration procedures (Cipresso et al., 2012). Moreover, the throughput speed should be increased and the operation mode should match the daily life necessities. Among the physiological signals usable as control features for a BCI, the P300 is an event-related potential (ERP) widely used for communication and environmental control since it allows selecting an item of interest between a set of available choices with a relatively low effort (no user training, short calibration sessions, possibility to display several items at once, etc.). ERPs show a wide variability, both between different subjects and within the same subject (Polich and Kok, 1995; Ravden and Polich, 1999). In fact, external factors such as light, noise, stimulation modalities (Cano et al., 2009; Polich and Bondurant, 1997), and "internal" factors as the attentional level or fatigue may affect the morphology of these potentials (Geisler and Polich, 1992; Polich, 1997). As pointed out by Thompson et al. (2012) these factors can affect the reliability of BCI systems. The authors reported evidences about the variability of the P300 potential morphology across different BCI sessions. The tuning of parameters exploited to control the system should be frequently updated in order to ensure the highest performance. However, a frequent explicit recalibration of the system (i.e. the supervised acquisition of data to train the classifier) would be time consuming and frustrating for the users. For this reason, classification methods for partial/complete unsupervised learning in P300 based BCIs have been proposed in order to reduce/avoid the calibration process (Kindermans et al., 2012; Lu et al., 2009; Panicker et al., 2010) as well as user friendly solutions to simplify the configuration and calibration procedures (Kaufmann et al., 2012). However the proposed methods were tested on brief controlled BCI sessions (1-2 hours) with no assessment of the inter-sessions variability. Moreover, the proposed methods did not address two important issues in order to fill the gap between

BCI and assistive technology (AT) input devices: (i) BCIs should implicitly withhold control when the user is not attending to the interface, even without an explicit mechanism to enter a pause mode;(ii) BCIs should dynamically adapt the speed of selection to the subject's skills (Dynamic Stopping) and provide an appropriate tradeoff between recognition accuracy and speed, allowing the system to maintain high communication efficiency level. The asynchronous classifier described in chapter 4 (Aloise et al. 2011) addressed these issues increasing the communication efficiency of P300-based BCI systems, both for communication and environmental control applications (Schettini et al., 2012).

This section aims at (i) investigating whether a repeated (automatic) update of the classifier's parameters across several BCI sessions increases the system's performance in terms of accuracy and communication efficiency; (ii) proposing and evaluating a self-calibration algorithm to label data acquired in unsupervised modality. The latter will be used to update the classifier parameters with no need for an explicit calibration session.

5.1 Experimental Protocol

Ten healthy subjects were involved in this study (5 male and 5 female, mean age 25 ± 3). All subjects had previous experience with P300-based BCI and had normal or corrected to normal vision. Scalp EEG signals were recorded (g.USBamp, gTec, Austria, 256 Hz) from 8 scalp positions (Fz, Cz, Pz, Oz, P3, P4, PO7, PO8 – Krusienski et al., 2008) referenced to the right earlobe and grounded to the left mastoid. The stimulation interface consisted in a 6 by 6 matrix containing alphanumeric characters (P300 Speller - Farwell and Donchin, 1988). Stimulation and data recording were managed by the BCI2000 framework (Schalk et al., 2004). Visual stimulation consisted in the pseudo-random intensification of rows and columns on the interface: target stimuli consisted in the intensification of the row or the column containing the character attended by the subject whilst non-targets were the intensifications of any other row or column. Each row and column was intensified for 125ms, and 125ms elapsed between the end of one stimulus and the onset of the following one (inter-stimulus interval). A stimulation *sequence* consisted of the consecutive intensification of every row and column on the interface, for a total of 12 stimuli (2 targets and 10 non-targets). With the term *trial*, we will refer to a set of 8 successive repetitions of the stimulation sequence, relating to the same target character. A *run* is an uninterrupted series of 6 trials, followed by a pause in the EEG acquisition. A *session* consisted of 6 control runs and 2 no-control runs. During each *control* run, 6 different characters were prompted as target; within each session all 36 characters of the interfaces were prompted exactly once. During the two *no-control* runs EEG data was acquired while the subject was in a voluntary no-control state: subjects were required to gaze at a fixation cross at the middle of the interface and to ignore the surrounding stimulation. In one of the two no-control runs subjects were also required to solve simple arithmetical problems posed by the experimenter (Aloise et al., 2011). Each subject underwent 5 recording sessions in the same day at well-defined times: 10:00 AM, 12:00 PM, 2:00 PM, 4:00 PM and 6:00 PM. A total of 180 control trials and 60 no-control trials per subject were collected. The subjects were required to wear the EEG cap for the whole day. Before each session, the experimenter checked the correct position of the cap on the subject's scalp, the electrode-scalp impedance value (which was kept below 10KΩ) and the quality of the EEG signal. Each session lasted about 1 hour, and in the time between two consecutive sessions, the subject could perform daily activities such as working, studying, talking with friends or eating.

5.2 EEG pre-processing

The 8-channels EEG signal was segmented into 800ms epochs starting at the onset of each stimulus. The epochs were then downsampled, replacing each segment of 12 samples with their mean value and then reducing an epoch to 17 samples. The resulting 8x17 data arrays were concatenated creating a 136-elements features vector v_f for each stimulus. The classifier was trained on the resulting set of feature vectors, each associated to the label of a *target* or *non-target* stimulus. In addition, epochs relating to no-control periods were included in the training set by labelling them as non-target epochs.

A stepwise linear discriminant analysis (SWLDA) was applied to identify the most significant features and to estimate the weights *w* of the linear classifier (non-significant features will be given a weight $w_i = 0$). The number of maximum iterations of the algorithm was set to 60. For each iteration, features with p-value < 0.1 were added to the model while features with p-value > 0.15 were removed from it (Krusienski et al., 2006). Score values *y* for each stimulus were computed as the linear combination of the features vectors with the classifier's weights ($y = w^T v_f$)..

5.3 Self-calibration algorithm

The self-calibration algorithm performs an (on-line) unsupervised labeling of data, and processes them to automatically update the classifier weights. The proposed method relies on the introduction of two different threshold sets in the classifier. The first will be denoted as classification threshold (CT) and it will allow for the dynamical stopping and the control suspending features (Section 5.3.1). The second set of thresholds will be denoted as labeling threshold (LT) and it will be used to decide which sequences can be reliably labeled for the continuous updating of the classifier's weights (Section5.3.2).

Figure 5.1 shows a flowchart of the self-calibration algorithm. At the beginning, the classifier's parameters and the thresholds are defined using data from the previous session (or from an explicit calibration session for the very first use of the system). Every time a new stimulation sequence is delivered, the score values for each stimulation class are computed and compared with CTs. When the CT is exceeded for both rows and columns classes (i.e. a character is tentatively selected), the difference score values are estimated (see section5.3.2).). If they exceed also the LT the epochs relating to the current trial, they are labeled according to the classification result and stored for further recalibration. Every time a predefined number of new epochs ($N_{recalibration}$) from the online session are stored, the same amount of the oldest epochs is removed from the training dataset and the classifier weights and the thresholds values are updated. The $N_{recalibration}$ value is set as the 5% of the number of epochs in the recalibration database. Potentially, recalibration might be performed every time a new epoch is added to the recalibration database ($N_{recalibration}$ = 1), if the computational power of the system would allow for it. On the other extreme, $N_{recalibration}$ might be set to 100% of the recalibration database's size, i.e. a recalibration is only performed when a totally fresh dataset is available. In fact, an offline simulation showed that the 5% value is an effective compromise between update frequency and computational requirements.

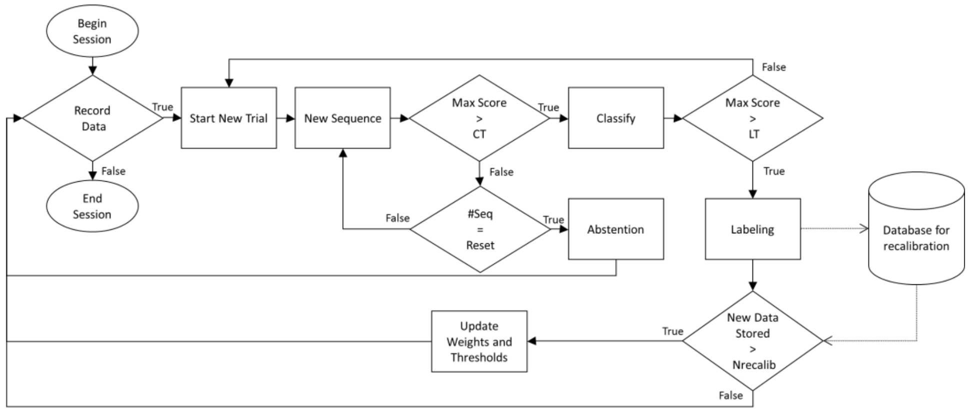

Figure 5.1 Flowchart of the self-calibration algorithm

5.3.1 Classification thresholds

The classification thresholds were defined according to the method described chapter 4. Their value was recomputed every time a new set of $N_{recalibration}$ epochs is available.

For each stimulation sequence available in the training set, 12 scores y_{stim} were computed, 6 for the rows and 6 for the columns on the interface. Within each trial, the stimulus scores of the first, second, … eighth sequence seq were accumulated (summed up), yielding $Y_{stim}[seq] = \sum_{i=1}^{seq} y_{stim}[i]\,, (seq = 1, .., 8)$. For each stimulation sequence the maximum score value $M[seq] = \max_{stim}\{Y_{stim}[seq]\}$ was extracted and a label equal to 1 (target) or 0 (non-target or no-control) was assigned to it. Therefore, we defined a distribution of the maximum score values for each stimulation sequence.

Each distribution was used to plot a receiver-operating characteristic (ROC - - Zweig and Campbell, 1993). In order to reduce the false positive rate (FPR) and ensure a high true positive rate (TPR), the threshold was selected at the intersection between the ROC curve and the segment joining the point (0.05, 0.5) with the point (0,1), since the former point provided a tradeoff between FPR (maximum 5%) and TPR (minimum

50%) and the second point represents the optimum classification (FPR = 0% and TPR = 100%).

5.3.2 Labeling thresholds

With regard to the labeling thresholds definition, a different procedure was applied on training data in order to ensure a high level of robustness to false positive.
Starting from the values of the score accumulated according to the number of stimulation sequences delivered in each trial, scores relating to rows and columns stimuli were sorted and the difference between the highest score value and the second highest was computed for rows and columns separately. The difference scores were labeled as 1 or 0 if the highest score was a target or a non-target/no-control score respectively. Therefore, it was possible to define a distribution of difference scores for each stimulation sequence (8 distributions for rows and 8 distributions for columns in this specific case).

The distribution of difference scores were used to plot 16 ROC curves (8 for rows and 8 for columns). The threshold values corresponded to the point on the ROC curve ensuring a 0% FPR with the maximum value of TPR possible on the training data. Then, considering the testing data, only the trials in which the maximum score value exhibited a high difference with respect to the second one can exceeded the threshold ensuring a high level of robustness to false positives.

5.4 Performance assessment

5.4.1 Intra-session and inter-session performance

We first assessed if accuracy and the communication efficiency of an asynchronous P300-based BCI benefits from a continuous updating of the classifier's parameters. Two conditions were investigated through offline cross-validation performed with the asynchronous classifier (i.e. dynamic stopping and abstention features were enabled): intra-session and inter-session conditions (Figure 5.2).

In the intra-session condition, the training dataset and the testing data set belonged to the same session. For each cross-validation iteration the training data set consisted of 5 control runs and 2 no-control runs while the remaining control run of the same session was used as testing dataset. Every control run was used as testing dataset once, thus 6 cross-validation iterations for each session were performed.
In the inter-session condition the training dataset and the testing dataset belonged to different sessions. The cross-validation design matched as close as possible the intra-session's cross validation design. For each iteration the training data set consisted of 5 control runs and 2 no-control runs extracted from session *i* while the testing dataset consisted of one control run from session *j*, namely the run with index corresponding to the session *i*'s run was not included in the training set. Each pair (*i*,*j*) of sessions participated in the cross-validation. Performances were assessed in terms of correct classifications, errors and abstentions (Figure 5.2).

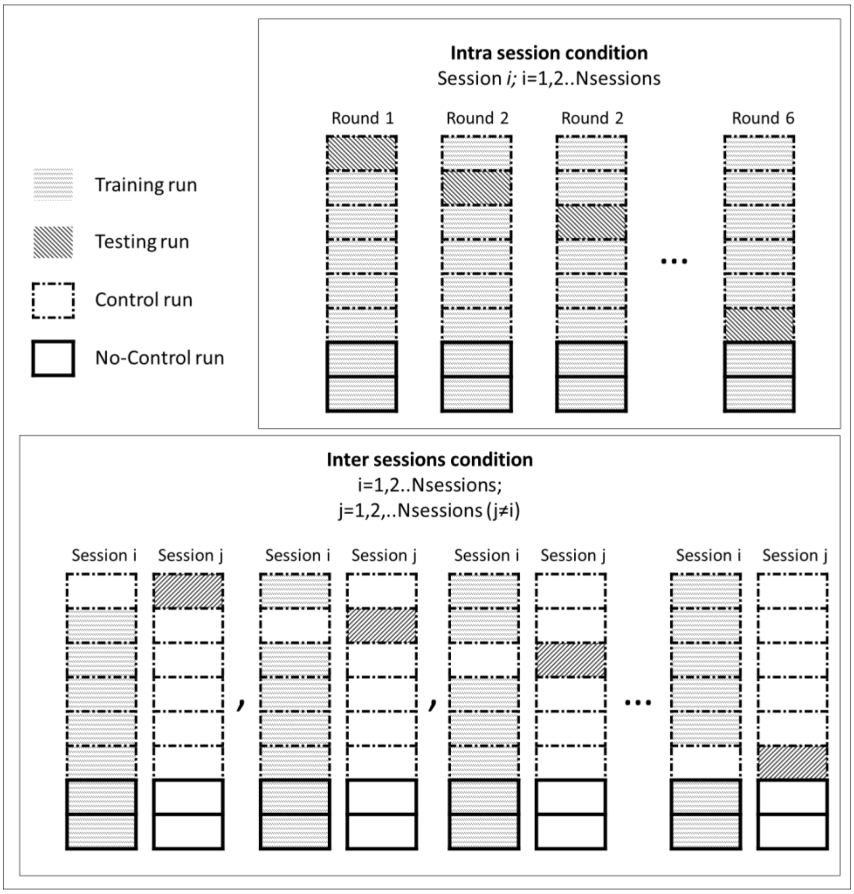

Figure 5.2 Intra-session and Inter-sessions crossvalidation. Each block represents a run relating to the spelling of 6 different characters, except for the two

no-control runs in which the subject was required to divert his attention from the stimulation interface.

5.4.2 Self-calibration algorithm evaluation

The performances of three conditions were compared: no-recalibration, intra-session, and self-calibration. In the no-recalibration condition, the classifier’s weights and classification thresholds values were extracted using data from the first session, and then applied to the other sessions simulating a plausible usage of the system during the day. The intra-session condition was the same of section 5.4.1. Even if this is not a realistic condition, we considered the intra-session condition as a reference level for the best performance achievable by a continuous supervised calibration of the classifier’s parameters.
To assess the self-calibration algorithm performance the following procedure was applied: at the beginning the classifier’s weights, the classification thresholds and the labeling thresholds were extracted using data from the first session, then the self-calibration algorithm was ran on data from the 4 sessions acquired later. Particularly, starting from second session data, performance were sequentially assessed by runs for all the available sessions updating the database for recalibration accordingly. The classifier’s parameters, as well the thresholds values, were updated when 5% of new data (respect to the dimension of the starting calibration dataset) was stored.

5.4.3 Evaluation of communication efficiency

In order to summarize the system performance in the different conditions we adopted a metric to quantify the efficiency of the system from three points of view: accuracy, error rate and speed. An asynchronous classifier has 3 different possible classification outputs: *i)* correct classification, when the target item is correctly recognized by the system; *ii)* error, if the item selected is different from the one attended by the subject; *iii)* abstention, when no item reaches the classification threshold at the end of the trial.
For this reason metrics that only take into account the classification accuracy, such as the Written Symbol Rate (WSR - Furdea et al., 2009) or the Wolpaw’s bit-rate (Wolpaw et al., 2002) produce an incomplete characterization of asynchronous

systems, since they cannot distinguish wrong selections from abstentions. In previous section, the metric of Bianchi et al. (2007) was successful applied to assess the communication efficiency of the proposed asynchronous system. This metric predicts the extent to which the classification accuracy supports communication—i.e., whether the time that is spent in correcting mistakes is shorter than the time needed to generate a correct selection. The efficiency of a system, with regard to the time that is needed to achieve a classification, is expressed as a function of the number of stimulation sequences (NumSeq):

$$Eff = \frac{1}{NumSeq * \overline{\overline{ESC}}}$$

Where ESC denotes the Expected Selection Cost (ESC), which is the mean number of selections that is needed to generate a correct symbol, taking into account the recovery from errors and abstentions. When accuracy is lower than 50%, the time needed to correct errors is longer than the time spent for effective communication; thus ESC approaches infinity and Eff is 0.

In this study, we made the following assumptions about the costs: we associated a cost of 1 with abstentions (the user needs only to repeat the trial, trying again to select the desired character) and a cost of 2 with misclassifications (the user must first delete the incorrect character and then reselect the desired one).

5.5 Results

5.5.1 Inter-sessions and intra session performance

Results about average performance in terms of correct classifications, errors and abstentions for each test session for both intra-session and inter-sessions conditions are reported in Figure 5.3. Three one way repeated measures ANOVAs were applied using the cross-validation conditions as factor (intra-session versus inter-sessions) and correct classifications per testing session, errors per testing session and abstentions per testing session as dependent variables, respectively. The intra-session condition allowed for a significantly higher correct classification rate with respect to the inter-sessions condition ($F(4, 1192)=17.232$, $p<.01$) and this difference was compensated by a significantly lower error rate in the former condition with respect to the latter (F(4, 1192)=15.85, p<.01). No significant differences were detected in the percentage of abstentions within the two conditions (F(4, 1192)=1.49, p=.20).

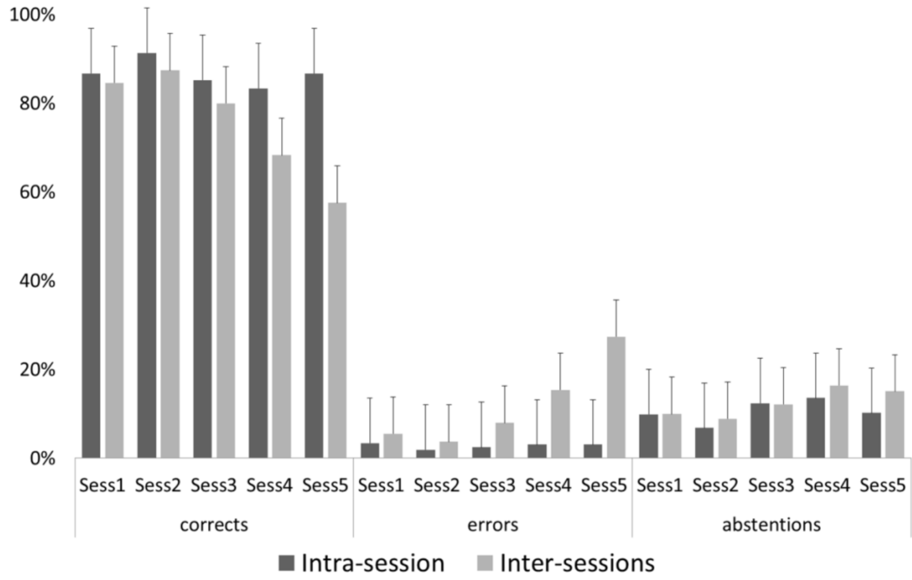

Figure 5.3 Intra-session and Inter-sessions performance as a function of sessions. Bars denote the standard error.

Communication efficiency was significantly higher in the intra-session condition than in the inter-sessions condition, as assessed by a repeated measures ANOVA using cross-validation condition as factors (intra-session and inter-sessions) and efficiency per testing session as dependent variables (F(4, 1192)=10.62, p<.01). Figure 5.4 shows the efficiency mean values for each session over the day in both intra-session and inter-sessions conditions. Table 5.1 reports the average efficiency values for each cross-validation condition. Particularly, values on the main diagonal of the matrix correspond to the intra-session condition while the entries outside the main diagonal refer to the inter-session conditions.

Table 5.1 Communication Efficiency for each crossvalidation condition

		Testing session				
		Sess1	Sess2	Sess3	Sess4	Sess5
Training Session	Sess1	**0.33±0.15**	0.32±0.14	0.26±0.18	0.14±0.14	0.10±0.11
	Sess2	0.29±0.15	**0.35±0.12**	0.27±0.12	0.18±0.14	0.14±0.13
	Sess3	0.29±0.13	0.30±0.13	**0.30±0.15**	0.22±0.12	0.15±0.13
	Sess4	0.28±0.13	0.28±0.14	0.26±0.13	**0.27±0.14**	0.16±0.13
	Sess5	0.26±0.12	0.27±0.08	0.24±0.11	0.21±0.11	**0.26±0.11**

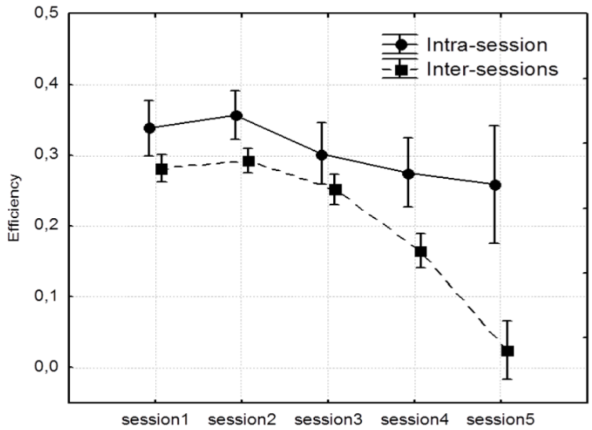

Figure 5.4 Trend of communication efficiency across sessions for intra-session and inter-session crossvalidation conditions

5.5.2 Self-calibration algorithm assessment

Figure 5.5 shows the average performance of the system for all ten subjects in the three different conditions: intra-session, self-calibration and no-recalibration. Since for all the conditions data acquired during the first session was used as training dataset, performances for session 1 were not available. As it can be seen accuracy decreased in the self-calibration and no-recalibration condition over sessions while error rate increased. However, the self-calibration algorithm exhibited a lower decrement in performance with respect to the no-recalibration condition. Three repeated measures ANOVAs have been applied using the cross-validation conditions as factors (intra-session, self-calibration and no-recalibration) and correct classifications per testing session, errors per testing session and abstentions per testing session as dependent variables respectively. Considering the correct classification rate the test pointed out significant differences between conditions ($F(6, 477)=6.62$, $p<.01$) and a Duncan's post-hoc tests showed that correct classification in the intra-session cross-validation condition were significantly higher (p<.05) with respect to the no-recalibration condition (all the sessions) and to the self-calibration condition (sessions 3, 4 and 5). Moreover the latter condition exhibited a significantly higher (p<.01) correct classification rate with respect to the no-recalibration condition for session 3, 4, 5. All other comparisons were not significant. Also for error rate the repeated measures ANOVA showed significant differences between conditions (F(6, 477)=5.78, p<.01). Duncan's post-hoc tests revealed that the no-recalibration condition exhibited significantly higher (p<.05) error rate with respect to the intra-session condition (sessions 3, 4 and 5) and the self-calibration condition (sessions 4 and 5). The error rate in the self-calibration condition was significantly higher (p<.05) than in the intra-session condition only for session 5. All other comparisons were not significant and no significant differences were detected between the three conditions in terms of abstentions (F(6, 477)=1.17, p=.31). With regard to the robustness to false positives during no-control trials, in the self-calibration condition the classification threshold was erroneously exceeded by 4.1% of the no-control trials while in the no-recalibration condition we detected on average 19.4% of unwanted classification. More specifically, in the self-calibration condition we detected 0.04, 0.09, 0.11 and 0.16 false positives/minute in session 2, 3, 4, 5 respectively. In the no-recalibration condition we found 0.09, 0.44, 0.53, 0.88 false positives/minute in session 2, 3, 4, 5 respectively.

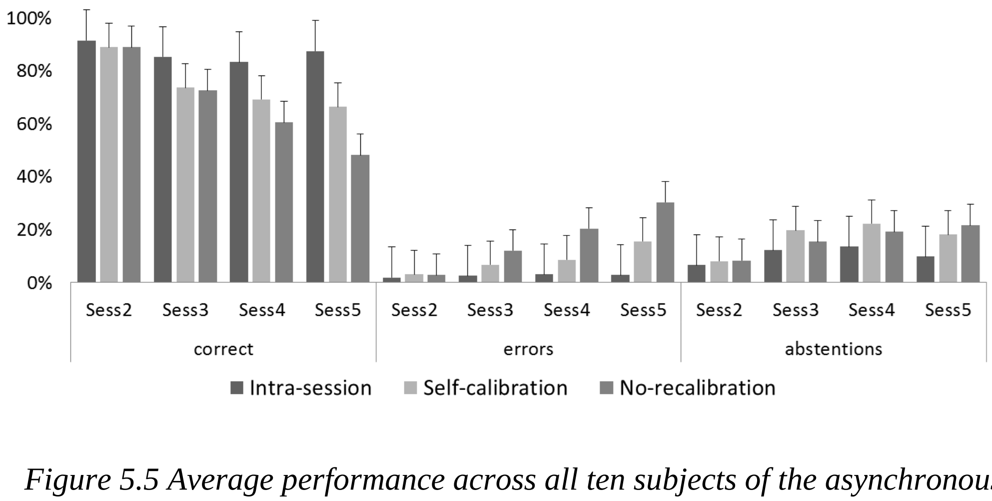

Figure 5.5 Average performance across all ten subjects of the asynchronous system in the intra-session, self-calibration and no-recalibration crossvalidation conditions. Bars denote standard error value.

Figure 5.6 reports the efficiency values for the three conditions over sessions. The self-calibration algorithm allowed for achieving efficiency values close to the intra session condition while the no-recalibration condition exhibited a significant decrease of efficiency with respect to the other two conditions. A repeated measures ANOVA with the cross-validation conditions as factors (intra-session, self-calibration and no-recalibration) and efficiency as dependent factor did not point out statistical differences between the three groups (F(6, 477)=1,8049, p=,09631). A Duncan's post-hoc test pointed out significantly higher efficiency values for the intra-session condition (sessions 4 and 5) and the self-calibration condition (sessions 3, 4 and 5) with respect to the no-recalibration condition. No significant differences were found between the intra-session and self-calibration conditions.

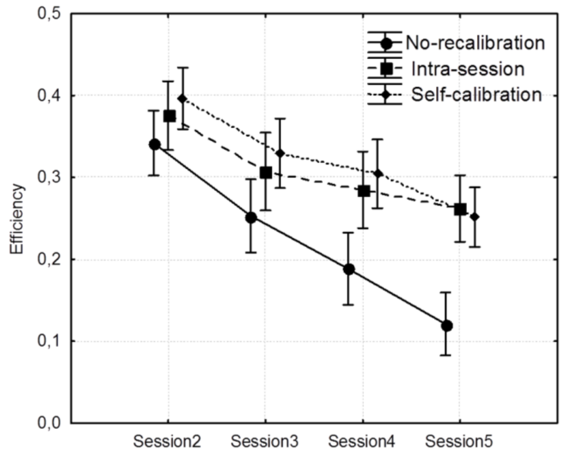

Figure 5.6. efficiency mean values for the no-calibration, intra-session and self-calibration conditions.

5.5.3 Data labeling

On average 36±10% of online data exceeded the LT and thus was stored for recalibration; 95.5±3.8% of stored data was correctly labeled as target or non-target. For two subjects out of ten all data stored was labeled with 100% accuracy, while the highest percentage of wrongly labeled data detected over subjects was 12.5%. However this did not affect performance of the considered subject since significant differences (p<.05) between the self-calibration and the intra-session condition in terms of efficiency were found only for session 4 as assessed by mean of a Duncan's post-hoc test on a repeated measures ANOVA using cross-validation conditions (intra-session, self-calibration and no-recalibration) as factors and efficiency as dependent variable ($F(6, 45)=4.67$, $p<.01$).

5.6 Discussion

This section describes and validates an algorithm for automatic adaptation of the classifier's parameters, designed to be employed during on-line sessions. First we investigated if parameters' updating increases system accuracy. Contrariwise to what was reported in (McFarland et al., 2011) results showed that the re-calibration of the system with data acquired within the same session can ensure a higher reliability and efficiency with respect to the recalibration performed with data acquired during a different session. Moreover, it should be stressed that in the current study the phenomenon of performance variability was investigated only on BCI sessions acquired within the same day and with young and healthy subjects. With regard to the latter aspect, several factors should be considered: i) a decrement of overall performance in the afternoon sessions was detected with respect to morning sessions even in the intra-session condition and this can be due to the subject's fatigue and decrease of motivation after repeated sessions; ii) this study reports results about data acquired in controlled experimental conditions while a test involving end-users in real life context would provide more realistic results; iii) in section 4.2, authors demonstrated how potential end-users exhibited an higher variability of performance in terms of stimulation sequences needed to select the desired item with respect to young healthy subjects, so the inter-sessions/intra-session difference might be greater involving end-users in the evaluation; iv) it could be interesting to assess performance variability over repeated sessions in different days in order to investigate if the PM decrement is always present or if the latter was mainly due to the very strict pace of the experimental protocol.

The self-calibration algorithm performances were compared with two different conditions, the intra session condition that represents a reference condition and the no-recalibration condition in which it was assumed that the user calibrated the system once and continued to use the same parameters for the whole day. Although the correct classification rate of the self-calibration algorithm was significantly lower than the intra-session condition, it significantly overcome the correct classification rate of the no recalibration condition, moreover the latter exhibited a higher error rate with respect to the other two conditions. All these factors as well for the lower number of stimulation sequences needed to exceed the classification in the self-calibration algorithm reflected in the efficiency values. In fact, the communication efficiency observed with the self-calibration algorithms did not exhibit significant differences with respect to the intra-session condition. Finally, the self-calibration algorithm demonstrated to be very reliable in data labeling since less than the 5% of data stored for re-calibration was wrongly labeled. However further tests involving end users in ecological condition with an on-line implementation of the proposed algorithm are needed to confirm the promising results obtained with healthy subjects by mean of off-line speculation.

6 EEG QUALITY INDEX

During experimental BCI sessions performed in research laboratories, a specialized technical figure takes care of all the procedures regarding the configuration of the system such as place the EEG headset, verify that the impedance value between each electrode and the scalp is low enough, and check that there are no environmental electromagnetic noises that might contaminate the acquired data. Moreover, even when the recording session is ongoing, a good experimenter always has a look to the EEG traces in order to immediately recognize acquisition problems. However, these conditions cannot be replicated for a home independent use of BCI systems when specialized technical staff is not available, and then algorithm to automatically detect problems in EEG acquisition should be provided. Barachant et al. (2013) proposed a multivariate automatic and adaptive method for identifying artifacts in continuous EEG data. This method uses covariance matrices as descriptors of EEG signals and employs a Riemannian metric to compare these covariance matrices with an average covariance matrix estimated on the signal baseline. The results of a preliminary study demonstrated that the proposed method allows to rejecting blinks, electrodes movements and eye movements online. In this section, starting from the above mentioned results, the problem of what kind of artifacts can affect BCI performance has been addressed by mean of a simulation study and a solution for the online monitoring of the quality of the EEG signal acquired was provided.

6.1 The Riemannian geometry and the artifact rejection method

The artifact rejection algorithm aims at determining if a segment of EEG signal $X \in \Re^{NxT}$ referring to a time window of T samples over N electrodes contains artifacts. In order project data into the Riemann space, a trial X will be represented by its spatial covariance matrix $C = \frac{1}{T-1} XX^T$ and the criterion for the detection will be defined according to the Riemannian distance computation method. The first step consist of the estimation of a reference covariance matrix $\bar{C}$ and reject every trial which is too far, in term of Riemannian distance, from this reference matrix. The Riemannian distance between C and $\bar{C}$ is defined by Förstner and Moonen, (1999):

$$d_R(C, \bar{C}) = \sqrt{\sum_{n=1}^{N} \log^2(\lambda_R)}$$

With λ_R the eigenvalues of $C^{-1/2}\bar{C}C^{-1/2}$. The trial corresponding to C will be considered as an artifacts if d_R is greater than a threshold th. Thus, the detection algorithm requires two parameters: $\bar{C}$, the reference point in the Riemannian manifold and the threshold th for the detection. The estimation of those two parameters is important part of the algorithm. this algorithm is equivalent to defining a region of interest (ROI) within the manifold of SPD matrices where every point outside the ROI would be considered as an outlier. The reference point is the center of the ROI. This center should be estimated in order to be the expected value of a reference brain activity. The reference point can be estimated in an unsupervised way by using all the EEG signals available. In the case of artifact detection, the reference brain activity is the baseline activity in which there are no artifacts. Once the reference signal has been chosen, the reference covariance matrix can be estimated. Basically, the reference covariance matrix is obtained by averaging all the covariance matrices of the I trials extracted from the reference signal. Here the Riemannian geometric mean was used. Similarly to the arithmetic mean, the Riemannian mean is the point that minimize the sum of the squared distances:

$$\bar{C} = \underset{C}{argmin} \sum_{i=0}^{I} d_R(C, C_i)^2$$

There is no closed form for the solution of this equation. Thus, the estimation of the geometric mean is an iterative process which attends to find a solution to the optimization problem and the easiest way to estimate this mean is to rely on the tangent space mapping.

The threshold is estimated in order to have (almost) all the points used to estimate the reference point in the ROI. This could be done by a statistical estimation.

Particularly in this work a basic estimation has been used: once the reference point has been defined, all the distances d_i between the C_i and the reference point $\bar{C}$ can be assessed. Then the mean and standard deviation can be used to estimate the threshold in order that almost all the points C_i lie in the ROI:

$$th = mean(d_i) + 2.5\ x\ std(d_i)$$

6.2 Preliminary study on the influence of artifacts on P300-based BCI

6.2.1 Dataset description

The dataset used in this simulation study was collected involving 5 healthy subjects in P300-based BCI sessions. The stimulation interface was the P300 Speller (Farwell and Donchin, 1988), containing 36 alphanumerical characters arranged on a 6 by 6 matrix. Each subject was required to spell all the 36 characters on the interface, so 36 trials per subject were collected. During one trial each row and column on the interface was intensified 8 times (8 stimulation sequences per trial). Stimulus duration and Inter Stimulus Interval were constant (125ms). Data from each subject was split in training data set (16 trials) and testing dataset (16 trials).

6.1.1.2 *Saturation of an acquisition channel*

One common artifact detectable in EEG traces is related to high fluctuations on one or more channels, this kind of artifact might have a variable duration depending on its cause (subject movements, environmental noise, loss of the contact between the scalp and one or more electrodes). In order to simulate this effect a 3 seconds period sinusoid with amplitude ranging from – 1mV to a 1 mV was injected in the EEG signal. EEG data was band pass filtered between 1 and 20 Hz. *N* artifacts (where *N=0, 1, 2, …7*) were injected on PO7 on each trial of the testing dataset. The position of the artifacts along the trial was randomized each time. Data were then divided in 800ms long epochs. Covariance matrices were computed both for training and testing dataset and then the region of interest and the threshold were estimated using the Riemannian geometry on training data set. A Stepwise Linear Discriminant analysis was applied on training dataset in order to extract most significant features and score values were estimated for each epoch of the testing dataset. Mean accuracy was assessed by averaging the accuracy achieved by each subject for each stimulation sequence.

6.1.2.2 *Eye blinks*

Eye blinks are common artifacts, occurring especially with BCI paradigms relying on visual stimuli. They have a duration of the order of hundreds milliseconds and appears mainly in the frontal and central area. In order to add eye blink artifact in controlled manner to EEG signal, epochs containing eye blinks were manually selected, and then the specific eye blink shape was extracted for each subject using the Blind Source Separation method (BSS - Congedo et al., 2008). The eye blinks were added to the EEG signal as in the previous case, but in this case they affected all the channels with different amplitudes.

6.2.2 Preliminary results

Figure 6.1a and b show the mean classification accuracy achieved in the case in which electrode saturations were added to the testing dataset without and with artifact rejection respectively. Accuracy values were reported both as a function of the number of stimulation sequences accumulated in a trial and as a function of the number of artifacts injected per trial. As it can be seen electrode saturation significantly affect system accuracy and artifact rejection can ensure higher accuracy with respect to the no artifact rejection condition, however if the number of artifacts injected in a trial is higher than 4, on average the 80% of available trials are removed and the system accuracy strongly decreases. In fact, a repeated measures ANOVA with *number of artifacts injected per trial* as independent factor and *stimulation sequence* and *condition (Artifact rejection vs No artifact rejection)* pointed out a significant difference ($F(8, 72)=18{,}781$, $p<.001$).

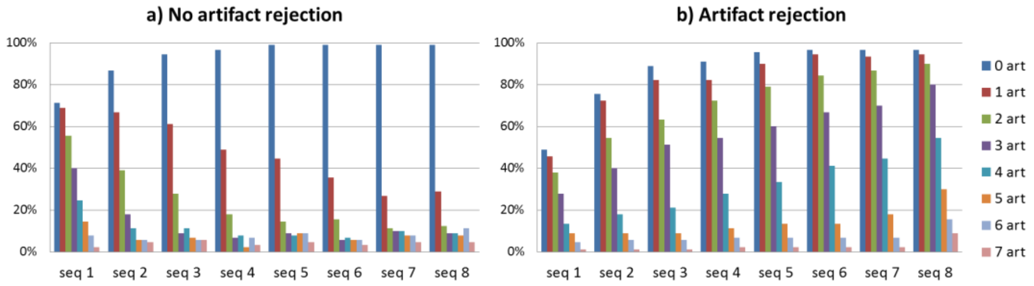

Figure 6.1 Electrode saturation: a)mean classification accuracy achieved without artifact rejection; b) mean classification accuracy with artifact rejection

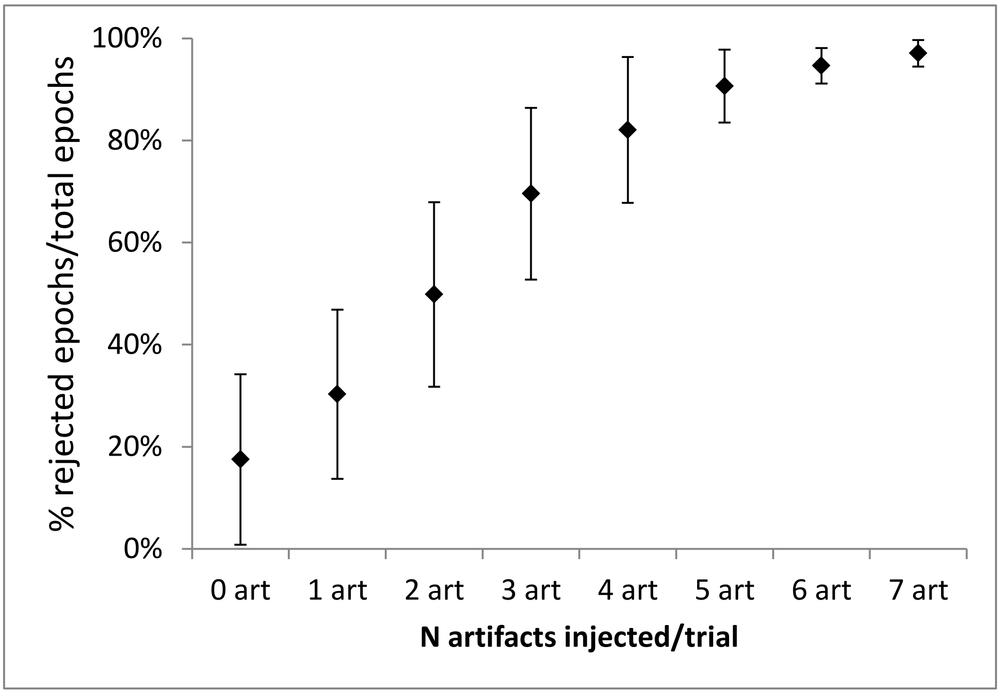

Figure 6.2 Rejected epochs percentage as a function of the number of artifact injected per trial. Values refer to subjects mean.

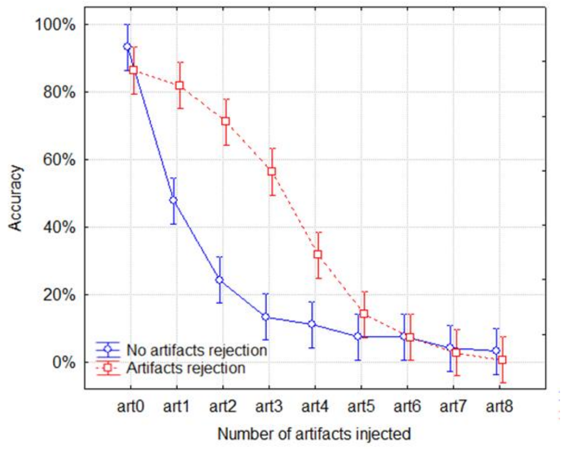

Figure 6.3 Direct comparison of Artifacts rejection and no Artifacts rejection conditions

Figure 6.4 a) and b) illustrate the mean classification accuracy achieved in the case in which eye blinks extracted by mean of BBS were added to the testing dataset without and with artifact rejection respectively. Accuracy values were reported both as a function of the number of stimulation sequences accumulated in a trial and as a function of the number of artifacts injected per trial. As it can be seen from Figure 6.4a blinks injection does not affect system accuracy, on the contrary a decrease in performance can be detected when the artifact rejection is applied (Figure 6.5), and this is due to the reduced number of epochs available for classification after the artifacts rejection.

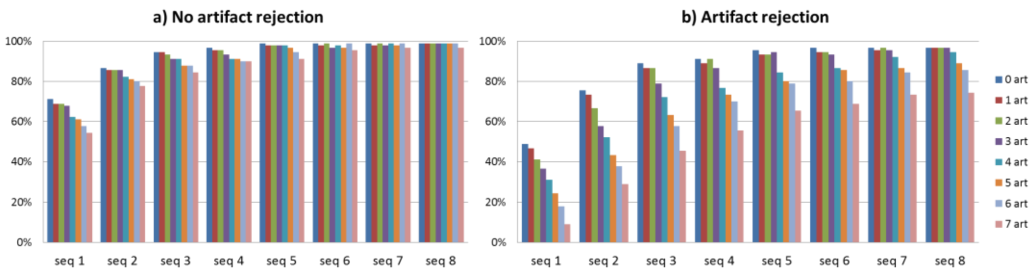

Figure 6.4 Eye blinks: a)mean classification accuracy achieved without artifact rejection; b) mean classification accuracy with artifact rejection

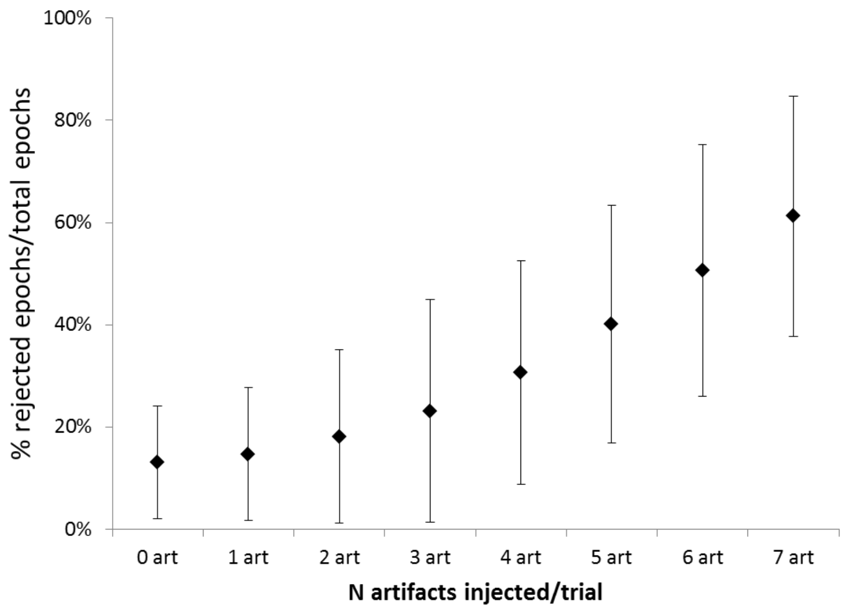

Figure 6.5 Rejected epochs percentage as a function of the number of artifact injected per trial. Values refer to subjects mean

In fact, a repeated measures ANOVA with *number of artifacts injected per trial* as independent factor and *stimulation sequence* and *condition (Artifact rejection vs No artifact rejection)* pointed out a significant difference ($F(8, 72)=3.13$, $p<.001$).

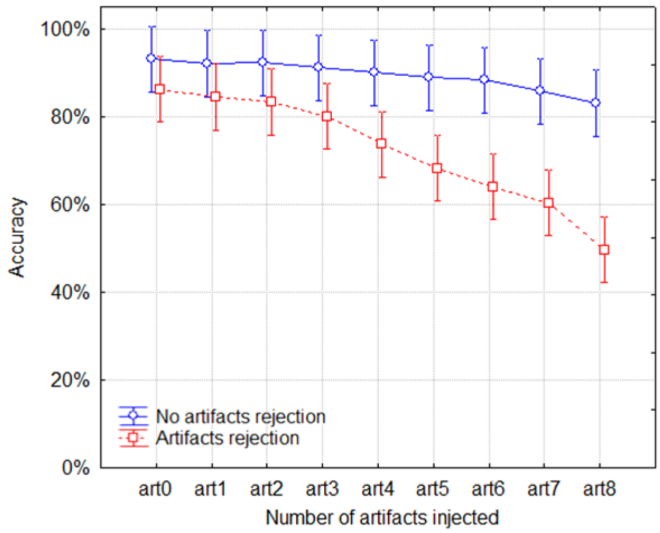

Figure 6.6 Direct comparison of Artifacts rejection and no Artifacts rejection conditions

These preliminary results led to the conclusion that the introduction of an algorithm for on-line artifact rejection does not imply an improvement in overall usability and reliability of P300-based BCI because small entity artifacts not have significant consequences on the accuracy of the system. The results of the first experiment, however, highlighted the need to promptly detect system failures due to problems in the acquisition of the EEG signal in an easy and intuitive manner, even for people without specific technical skills.

6.3 EEG quality index

Considering these preliminary results, instead of provide an algorithm for online detection of artifacts, we decided to define an index for the online assessment of the quality of the EEG signal acquired. The proposed index is based on the algorithm for the detention of the artifacts developed by Barachant et al. (2013). The algorithm has as main objective the detection of artifacts of the EEG signal due to the loss of contact of one and more electrodes, to the deterioration/breakage of a cable, or to the influence of considerable environmental noise. For this purpose, the EEG signal acquired is segmented into 1s epochs, overlapping each other at intervals of 250ms . Every epoch $X_i$can then be represented by its spatial covariance matrix C_i, and the first step is to estimate a covariance matrix $\bar{C}$ using reference data on which has already been done a rejection of artifacts. For each new epoch X_i the Riemannian distance $d_i = \delta_R(\bar{C}, C_i)$ is calculated, this distance is assumed to be the quality index of the EEG signal and if its value exceeds a given threshold for a time longer than a trial the system failure is reported to the operator, who in an ecological usage condition is represented by the user's caregiver.

6.3.1 Performance assessment

For the validation of the quality index the same dataset described in section 6.3.1 was used. 16 trials out of 36 per subject were randomly selected and contaminated adding a 3 seconds period sinusoid with amplitude ranging from –1mV to a 1 mV on channel PO7. Such sinusoid had the same duration of the trial plus 2 seconds corresponding to the result presentation phase. In this way the artifact was longer than the duration of the trial. Then the algorithm for the assessment of the quality index was trained for each subject using 60s of his EEG signal in which were not instrument artifacts, whilst blinks and short muscular artifacts were not removed. Once the center of the ROI ($\bar{C}$) was defined by computing the Riemannian geometric mean of the spatial covariance matrix of all the 1s long epochs, overlapped every 250ms belonging to the training data (C_i), all the distances d_i between the C_i and the reference point $\bar{C}$ can be assessed. Then the mean and standard deviation can be used to estimate the threshold in order that almost all the points C_i lie in the ROI:

$$th = mean(d_i) + 2.5\ x\ std(d_i)$$

Results will be reported in terms of:

- False positives: when a system failure is reported but the current trial is not corrupted by artifacts;
- False negatives: when the current trial is corrupted by an artifact but the system failure is not reported

6.4 Results

The results of the offline evaluation of the reliability of the EEG quality index pointed out on average 1.25% false positives, particularly only for 1 subject a system failure was wrongly reported, whilst for the other 4 subjects no system failures were reported during trials not corrupted by artifacts. With regard to trials in which instrumental artifact was artificially added, on average the algorithm showed 2.5% false negatives since for 2 subjects out of 5 a corrupted trial was not reported.

6.5 Discussion

In this section the problem of the reliability of the EEG signal acquired, and as a consequence the reliability of the BCI performance, was addressed. The algorithm for artifact rejection proposed by Barachant et al. (2013) and based on the Riemannian geometry seems to be a valuable solution to detect and remove artifacts from EEG signals during online BCI sessions. However remove corrupted epochs produces a decrement in performance, or at least a slowdown of the selection speed in order to acquire new epochs to replace the rejected ones. For this reason first a simulation study was performed in order to quantify the advantages of using an artifacts rejection algorithm, and two types of common artifacts were investigated. Preliminary results showed that small entity artifacts such as eye blinks do not affect system performance, and on the contrary remove corrupted epochs causes a decrease in of system accuracy due to the lower number of epochs available for classification. On the contrary, significant artifacts, such as the saturation of one or more EEG acquisition channel, strongly affect performance. These results stressed the importance of provide a tool for the online and continuous monitoring of EEG quality. Such a tool will detect possible failures of the system and will report them to the caregiver, ensuring the correct operation of the system even in ecological situation and without the need for specific technical competences.

7 P300-BASED BCI AS ASSISTIVE TECHNOLOGY FOR USERS WITH ALS

Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disease affecting upper and lower motor neurons with an incidence in European population of 2-3 people per year per 100.000 of the general population over 15 years (Al-Chalabi and Hardiman, 2013). Persons with ALS experience an increasing muscles weakness and atrophy which impairs independence and communication in their daily life. In each phase of the disease, this condition can be temporarily compensated by adopting an assistive device (AD), tailored to the current functional deficit. In fact, augmentative and alternative communication (AAC) can extend/replace means of communication for people with severe physical impairments by adopting solutions ranging from simple technology (e.g. eye-transfer alphabetic tables) to computer-based system (e.g. eye-tracker - Cipresso et al., 2012)).

When muscular contractions become impossible, brain-computer interfaces (BCIs) may represent a solution exploiting neurophysiological signals as input commands to control external devices(Millán et al., 2010; Wolpaw et al., 2002). At the current state of the art, studies with BCI applications have grown exponentially(Wolpaw and Wolpaw, 2012), yet those with targeted end-users with severe motor impairment are still few (Riccio et al., 2013; Sellers et al., 2010; Silvoni et al., 2009; Zickler et al., 2011).

Recently, the “user-centered design” (UCD; ISO 9241-210- 2010), according to which the end user must play as an active role in the device design and development iterative processes, has been introduced in the BCI field of research (Kübler et al., 2013; Kubler et al., 2013; Riccio et al., 2011; Thompson et al., 2013). The adoption of the UCD principles has provided the initial bases to bridge the gap still existing in translating the BCI technology from the laboratory to the real life usage scenario (Kubler et al., 2013). In this regard, several previous studies (Riccio et al., 2011; Thompson et al., 2013; Zickler et al., 2011) have shown the feasibility for the BCI technology to serve as an additional channel to access commercial available AT applications, thus paving the way to a wider applicability of the BCI technology.

One of the fundamental steps in the UCD cycling is the evaluation of technology design against the user's requirements (ISO 9241-210 - 2010). Following the adoption of the UCD in the BCI technology development, an effort has been made to apply objective metrics derived from the UCD to evaluate BCI controlled applications (Zickler et al., 2011). A preliminary framework of these metrics, adapted to evaluate the BCI technology in terms of usability, has recently been proposed (Kubler et al., 2013) and applied to communication and entertainment applications operated via an electroencephalographic (EEG)-based BCI.

While BCIs are often referred to as a potential solution to improve communication in people with severe motor disabilities, assistive solutions usable for daily life activities are seldom available. In two papers this goal was achieved by using the BCI as input channel to a commercial AT solution. A first study evaluated the usability of a system where a commercial AT software was controlled by a P300-BCI (Zickler et al., 2011) in a group of 4 end users with motor deficits. In the second study, an unmodified commercial AT was functionally operated through a BCI keyboard (Thompson et al., 2013). The authors demonstrated that using a BCI to control an unmodified commercial AT does not affect BCI performance in a group of 11 end users with ALS and 22 control subjects.

To optimize the BCI driven technology in terms of assistive systems, we recently proposed an AD prototype which provided functionalities seamlessly accessible through several conventional/alternative channels devices including a P300-based BCI (for a review about P300-based BCI see Kleih et al., 2011) to enhance/allow basic needs for communication and the environmental interaction (Caruso et al., 2013). The multimodal access to such AD prototype provided the user with an adaptable system to cope with his decreasing muscular abilities, and in the case of progressive loss of muscular function (like in ALS) it can be switched to an exclusive BCI control.

In this study we aimed at test the feasibility and to evaluate the system usability of the implemented AD prototype operated via the P300-BCI channel according to the metrics derived from the UCD approach. To this purpose, we adopted an experimental design in which the use of the AD prototype operated via the P300-BCI was compared to 2 conditions:

i) A widely validated stand-alone P300-BCI. Here, we specifically investigated whether the introduction of a dynamic interface consisting of matrices of different sizes would affect system usability. Such matrices allow the access to a range of different applications (virtual keyboard, domotic control, etc.)

ii) The same AD prototype operated via conventional/alternative channels based on residual muscular abilities. Here, our investigation focused on whether the limits in speed and accuracy of the BCI channel could affect usability with respect to conventional/alternative input devices.

7.1 AD Prototype design

The functionalities to be included in the prototype have been selected according to the results of a preliminary survey and two focus groups. The survey involved three classes of direct or indirect users: 20 professional stakeholders (i.e., experts in assistive technologies), 7 end-users, and 13 caregivers. Participants were asked to rate how useful the inclusion of further functionalities in the domains of interpersonal communication, environmental interaction and personal autonomy would be (Figure 7.1Figure 7.1). The two focus groups involved end-users, caregivers and stakeholders and were carried out in order to discuss the potentialities and the limits of a BCI system as assistive technology. Four main topics emerged from the two focus groups: i) the need of more information on BCIs and their potential applications; ii) the importance of having a modular system customizable to each user's needs, and to be able to follow the patient throughout the progression of the degenerative disease; iii) the relevance of emotional aspects in the relationship with the technology; iv) the importance for the end-users to remain active.

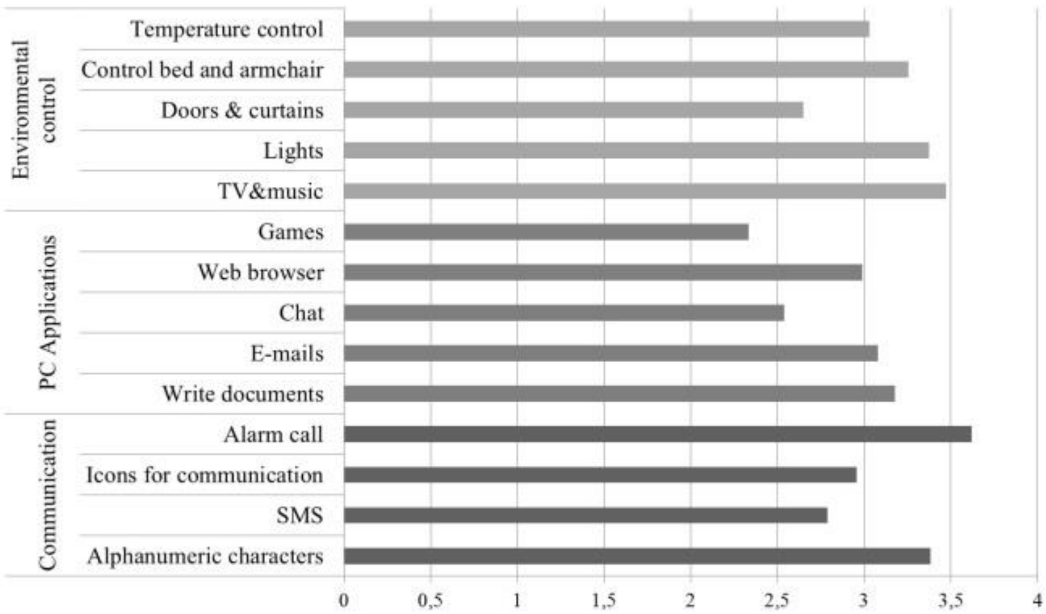

Figure 7.1: Preliminary results of the user survey, 20 professional stakeholders, 13 caregivers and 7 end users were interviewed about the useful (0 useless, 4 very useful) of the inclusion of specific functionalities in the domains of interpersonal communication, environmental interaction and personal autonomy.

7.1.1 Prototype description

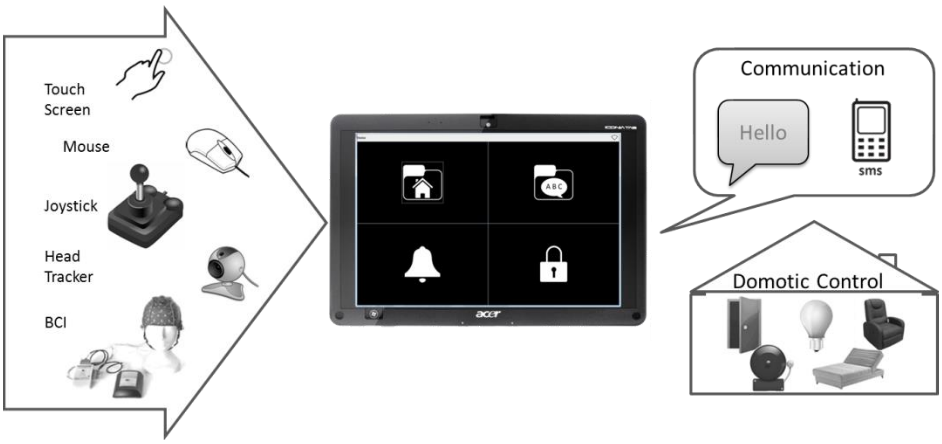

Figure 7.2 The prototype can be accessed using several conventional (touchscreen, mouse, keyboard) and assistive (buttons, switches, head tracker) input devices and a P300-based BCI.

7.1.1.1 *Functionalities*

Concerning interpersonal communication, the AD prototype provided three main applications: *(i)* an alarm bell to draw the attention of the caregiver; *(ii)* a simple text editor for both face-to-face and remote (e-mail, SMS) communication; and *(iii)* an interface to select predefined sentences or keywords for quick communication. For the environmental control, simple functionalities have been required by users such as TV control, movement of motorized armchair/bed, lights switching and doors opening (Caruso et al., 2013). These have been implemented by using the KNX standard to control the devices available at an apartment-like space designed to be fully accessible by people with motor disabilities for occupational therapy.

7.1.2.1 *Hardware and software*

To ensure portability and affordability, the AD prototype was developed on a 10’’ tablet and the software written in Java and C++ running on the Windows operating system. As far as the BCI, a specifically developed software program allowed to overlay visual stimuli (green grids in this case) necessary to generate evoked potentials on the user interface. Stimulation timing and data acquisition were managed by the BCI2000 framework and stimuli were delivered by a proxy application that managed the communication between the BCI2000 and the prototype user interface. All components of the software BCI ran on the tablet, as well as the other software components (including the user interface) of the prototype (Figure 7.2).

7.2 Experimental Design

7.2.1 Participants

Eight end users with ALS (5 male, 3 female; mean age=60 +/- 12 years; time since diagnosis=24+/-26.6 months) were recruited from the ALS Center (Department of Neurology and Psychiatry, Sapienza University of Rome) and gave their written informed consent to participate in the study, which was approved by the local ethical committee of Fondazione Santa Lucia, IRCCS Rome, Italy. We included in the study end-users who i) showed severe impairment in communication or environmental control, as measured with the ALS Functional Rating Scale revised (ALS-FRS-r - Cedarbaum et al., 1999). More specifically, they had to be scored ≤2 at one item of the ALS-FRS-r among item 1 ("word articulation"), item 4 ("writing ability"), item 5 ("ability to cut food/use tools"), and item 6 ("hygiene/personal care"). ii) End users had to show some residual muscular capabilities in order to use a conventional/alternative input device in addition to the P300-BCI. Demographic, clinical and neuropsychological description of the end users is reported in Table 7.1. Five of the end-users showed a deficit of the executive functions (as assessed by means of the Winsconsin Card Sorting Test – WCST - Heaton et al., 1999), two of them showed a deficit of the selective attention (SA) and four of them showed a deficit in a working memory (WM) task (both assessed by means of the computerized Test for attentional Performance – TAP - Leclercq and Zimmermann, 2002).

Table 7.1 Participants: Demographic, clinical and neuropsychological description of the end users. ALSfrs-r (revised ALS functional rating scale, Cedarbaum et al., 1999) is a validated scale monitoring the progression of disability in patients with amyotrophic lateral sclerosis, with scores ranging from 0 to 48. WCST (Winsconsin Card Sorting Test – Heaton et al., 1999) was used to evaluate the executive functions of the end-users. SA (selective attention) and WM (working memory) were evaluated by means of the computerized Test for Attentional Performance (TAP - Leclercq and Zimmermann, 2002). An equal sign indicates performance in the normal range. Down arrow indicates performance in the pathologic range. Up arrow indicates performance above normal.

	Sex	Age	ALSfrs-r	Onset	Conventional/Assistive input device	WCST	SA	WM
Subj1	M	55	13	Spinal	Automatic Scanning – 1 button	=	=	=
Subj2	M	59	37	Spinal	Touch screen	↓↓	↓	↓
Subj3	M	47	34	Bulbar	Mouse	=	=	=
Subj4	F	75	38	Bulbar	Mouse & Keyboard	↓↓	=	↓
Subj5	F	72	34	Bulbar	Mouse	↓↓↓↓	=	↓
Subj6	M	40	31	Spinal	Scanning - 2 buttons	↓↓↓↓	=	=
Subj7	M	61	28	Bulbar	Scanning - 2 buttons	=	=	=
Subj8	F	72	41	Bulbar	Mouse	↑	↓	↓

7.2.2 Experimental protocol

The overall usability of the AD prototype was evaluated by comparing three different experimental conditions performed in 3 experimental sessions (one per week), each lasting about one hour and half. In I and III (see next three sections), scalp electroencephalographic (EEG) signals were recorded (g.MOBIlab, gTec, Austria, 256Hz) from 8 Ag/AgCl electrodes (Fz, Cz, Pz, Oz, P3, P4, PO7 and PO8, referenced to the right earlobe and grounded to the left mastoid - Krusienski et al., 2008) according to the 10-10 standard.

7.2.1.2 *Condition I:P300 Speller*

Condition I concerned the control of a widely validated stand-alone P300-BCI (P300-Speller - Farwell and Donchin, 1988) and aimed at testing the baseline users' ability to control a BCI system. The P300-speller consisted of a 6 by 6 matrix containing 36 alphanumeric characters, which were intensified by rows and columns for 125ms, with 125ms of inter stimulus interval. Users had to spell 7 predefined words of 5 characters (so called copy mode). The selection of a character occurred after a train of stimuli (trial), in which every row and column on the interface was intensified 10 times. Prior to the start of each trial, the system cued the user with the character to spell. No feedback was provided to the users while spelling the first 3 words. This EEG-data set was used to extract the BCI classifier parameters by applying a stepwise linear discriminant analysis (SWLDA - Krusienski et al., 2006). The extracted parameters (features' weights) determined the online feedback (i.e. the selected character) during the spelling of the remaining 4 words.

7.2.2.2 *Condition II: AD Prototype controlled with conventional/alternative input device*

Condition II aimed at introducing the AD prototype to the users, who were asked to operate it via a conventional or an alternative input device (e.g. mouse, buttons, etc.). The experimenter showed the applications integrated in the AD prototype to the end users, who were then encouraged to explore it until they felt "confident enough". Throughout the session the users performed two pre-established tasks which mimic everyday life scenarios. To this aim they used the input device which best matched their residual motor abilities (see Table 7.1- Conventional/Alternative input device). Both tasks required a minimum of 8 selections. Since the experimenter provided the end user with only an indication about the final goal of the task, leaving the user free of developing his own strategy to fix possible mistakes, we will refer to these tasks as "self-managed tasks".

- Self-managed Environmental control task: The user had to perform sleep time actions, i.e., starting from the home menu of the visual interface, he was required to low the backrest of the motorized bed, turn off the light and go back to the home page.

Self-managed Communication task: The experimenter asked the subject "How is the weather today", and the user had to answer "BELLO" ("fine" in Italian) by writing it on the virtual keyboard and vocalizing it via the vocal synthesizer.

7.2.3.2 *Condition III: AD Prototype controlled with the P300-based BCI input*

The third session concerned the control of the AD prototype via the P300-based BCI, namely our main experimental condition. The prototype visual interface consisted of several menus with a minimum of 4 items (2 by 2 matrix) and a maximum of 36 items (6 by 6 matrix). Stimulation timing and repetition was the same as the P300 Speller session, for each item. Each end user carried out a total of 6 calibration (no feedback provided) runs: 2 with a 2 by 2 matrix, 2 with a 4 by 4 matrix and 2 with a 6 by 6 matrix. During each calibration run the users were required to attend four items prompted by the experimenter. Classifier parameters were calibrated as described for condition I on the ensemble of the calibration runs. During the subsequent online sessions, the end users were asked to perform 2 tasks consisting of a well-defined sequence of actions that were cued step by step by the experimenter (in case of a wrong selection the experimenter also suggested how to fix the error). This task allowed a direct comparison with Condition I. We will refer to these tasks as copy tasks:

- Copy - Environmental control task: the end user had to perform wake up actions i.e., starting from the home menu, he/she was required to turn on the light, raise the seatback of the motorized bed and then go back to the starting menu
- Copy - Communication task: The experimenter asked to the subject "How is the weather today", and the user had to answer "PIOVE" ("it rains" in Italian) by writing it on the editor and vocalizing it via the vocal synthesizer.

Finally each subject performed the same two self-managed tasks described in Condition II paragraph.

7.2.3 Prototype usability assessment

As in previous studies (Riccio et al., 2011; Zickler et al., 2011), specific performance metrics were considered for each of the three usability domains: effectiveness, efficiency and satisfaction. Metrics such as the time for correct selection and the Written Symbol Rate (WSR - Furdea et al., 2009) were considered as in between the efficiency and effectiveness domains.
The *effectiveness* was quantified in terms of

i) *BCI online copy accuracy,* expressed in terms of percentage of correct selections for the online copy tasks of condition I and III. It was calculated by dividing the number of correct selections by the number of total selections.

ii) *BCI offline accuracy*: expressed in terms of accuracy for stimuli repetition and assessed by mean of a 7-fold cross-validation on the copy-mode words spelled in Condition I and 6-fold cross-validation on the calibration runs of condition III. Specifically, for each round of cross-validation a different run was used as testing data set. On the results of cross-validation relating to condition III it was also assessed the number of stimuli repetitions allowing for the highest WSR and the corresponding accuracy value for each user.

iii) *AD prototype online accuracy during self-managed tasks*, expressed in terms of percentage of correct selections and calculated dividing the number of correct selections by the number of total selections of the self-managed tasks performed in condition II and condition III. Since the same number of stimuli repetition was applied all end-users during the online task of condition III (i.e. no optimization was performed), we also reported the accuracy value corresponding to the maximum WSR as estimated off-line. In the following, this condition will be referred to as condition III*.

Efficiency was quantified in terms of:

i) *BCI offline WSR:* the WSR was assessed for the copy tasks of conditions I and III as a function of the number of stimulus repetitions delivered in a trial.

ii) *AD prototype time for correct selection,* i.e. the total time (in seconds) to complete the task divided by the number of correct selections. It was calculated for the self-managed tasks of conditions II and III. As a separate analysis (Condition III*) we performed an off-line optimization of the time per correct selection. In fact, the number of stimuli per trial of on-line BCIs are usually calibrated to maximize communication speed. In our study, sequences of 10 stimuli were used for all subjects, to make group analysis more straightforward. In the offline simulation, for each subject the number of stimuli repetitions was set to the value which maximizes the WRS.

iii) *Workload,* measured by means of NASA-tlx (Hart and Staveland, 1988).

Satisfaction was reported by the end users by means of:

i) *VAS scores:* users were administered a Visual Analogue Scale (VAS: 1-10) to assess the overall satisfaction experienced with the P300 Speller (Condition I) and the AD prototype (Conditions II and III)

ii) *SUS scores:* obtained by administering to the end users the System Usability Scale (SUS: 1-100), which investigates end users' satisfaction in terms of pleasure experienced using the P300 Speller (Condition I) and the AD prototype (Conditions II and III)

The self-reported questionnaires (NASA-tlx, VAS and SUS) were administered by a psychologist at the end of each session.

7.2.4 Neuropsychological factors influencing BCI performance

7.4.1.2 *Rapid Serial Visual Presentation task*

In order to investigate the influence of temporal filters on P300-based BCI performance, temporal attention capabilities of participants were assessed by using the RSVP task as in Kranczioch et al. (2007). In the rapid serial visual presentation (RSVP) task (Figure 7.3), two targets were embedded in a stream of distracter stimuli. Each stream included 16 or 19 items, of which one or two were targets. All stimuli were presented at central fixation on a monitor with a white background at a presentation rate of 10 Hz. Each letter subtended a region on the screen of about 1.5°×1.38° of visual angle. Distracters were black capital consonants (except F, K, Q, X, Z) and the distracter sequence was pseudo-randomly extracted, with the constraint that the same letter could not be presented within three sequential positions. The first target (T1) was a green letter, which could either be a vowel or a consonant (except F, K, Q, X, Z), and the second target (T2) was a black capital "X". Each trial started with the presentation of a fixation cross for 900 to 1100 ms (mean 1000ms). T1 was presented randomly as 4th, 5th, 6th or 7th item in the stream. In 20% of trials T2 was not presented, whereas it followed with no (lag 1), one (lag 2), three (lag 4) or five (lag 6) intervening distracters, in 20% of trials for each condition. After the end of the stimulus stream, two successive screens appeared asking whether the green letter (T1) was a vowel and whether the black X (T2) was contained in the stimulus stream, as in Kranczioch et al. (2007). Participants completed 20 practice trials before completing 160 experimental trials, presented in a randomized order (32 trials for each of the five conditions). Due to the motor disabilities of the participants, they were asked to give a binary response (yes or no) to the operator with the residual communication channel. The detection accuracy of T1 (T1%) in the RSVP task was considered as an index of the temporal attentional filtering capacity of the participants. Since the detection accuracy of T2 (T2%) was considered as an index of the capability to adequately update the attentive filter, only trials in which T1 had been correctly identified were selected in order to determine T2%.

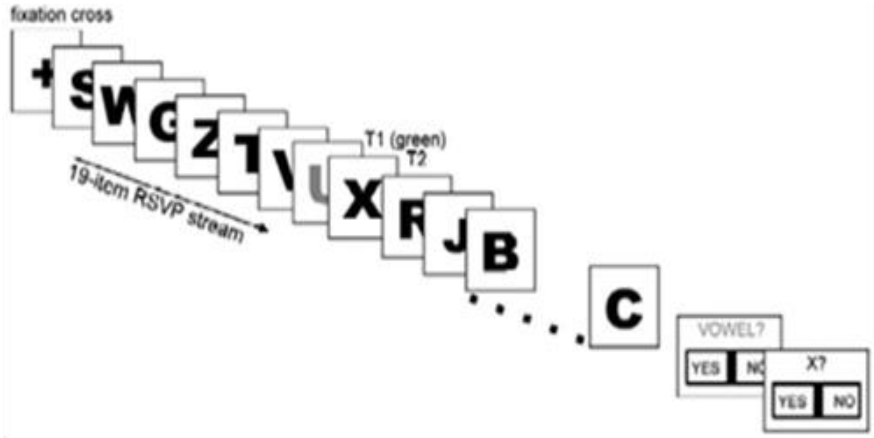

Figure 7.3: Rapid Serial Visual Presentation Task

7.4.2.2 *P300 morphology*

EEG data from the P300 Speller session was high pass and low pass filtered with cut off frequencies of 0.1 Hz and 10 Hz respectively using a 4th order Butterworth filter. In addition, a notch filter was used to remove 50 Hz contamination due to the AC interference. Data was divided into 1000ms long epochs starting with the onset of each stimulus. Epochs in which peak amplitude was higher than 70μV or lower than−70μV were identified as artifacts and removed. A baseline correction was done based on the average EEG activity within 200ms immediately preceding each epoch. The average waveform for both target and non-target epochs was computed for each trial in order to assess P300 peak amplitude. Particularly, amplitude of the P300 potential in Cz was defined as the highest value of the difference between target and non-target average waveforms in the time interval 250–700ms (P300 amp, Figure 7.4).

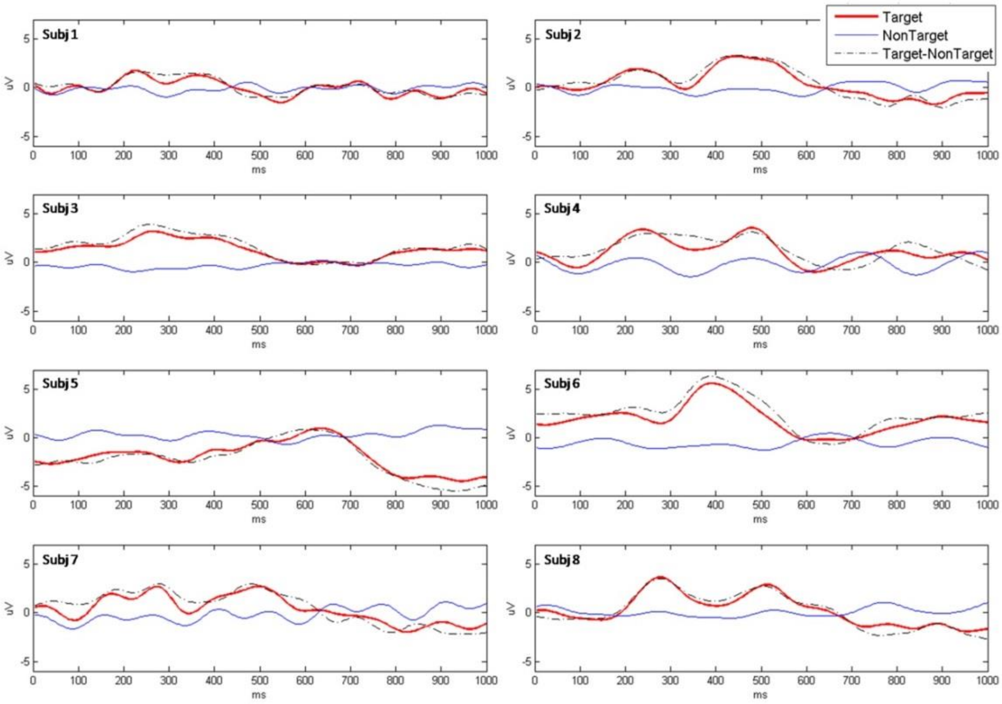

Figure 7.4: P300 amplitude in Cz for each subject involved in the experimentation

7.4.3.2 *Single trial classification*

To provide an estimate of the classifier accuracy the binary classification problem target vs. non-target (Blankertz et al., 2011) that takes into account the correct classification of a target or of a non-target was considered. Frequency filtering, data segmentation and artifact rejection were conducted as in P300 morphology section. EEG data were then resampled by replacing each sequence of 12 samples with their mean value, yielding 17×8samples per epoch (eight being the number of channels), which were concatenated in a feature vector (Krusienski et al., 2006). A sevenfold cross-validation was used to evaluate the binary accuracy (BA) of the classifier on each participant’s dataset. For each iteration a SWLDA was applied on the testing dataset (consisting of six words) to extract the 60 most significant control features (Draper and Smith, 1998) and the BA was assessed on the training dataset (the remaining word).

7.3 Results

7.3.1 Effectiveness

i) *BCI online copy accuracy*. We compared BCI online copy accuracy between condition I and III (Figure 7.5a). Since the distribution of these values for condition I violated the assumption of normality, we applied a Wilcoxon's matched pairs test. The analysis did not reveal statistical differences ($Z=1.18$, $p=.23$). However the distribution of the differences (accuracy condition I – accuracy condition III) was normal (Shapiro-Wilk's test, $W=.95$, $p=.78$, mean value = 6.6% ± 12), thus we can conclude that accuracy in condition I is on average less than 7% lower than in condition III. This difference was not significantly different from 0 as assessed by a one-sample t-test ($p=.19$).

ii) *BCI offline accuracy* of condition I and condition III was compared by means of a repeated measure ANOVA with conditions (I, III) as factors and the BCI offline accuracy per number of stimuli repetitions as dependent variables. Despite condition I exhibited a higher accuracy with respect to condition III, no significant differences were revealed ($F(9,792)=1.053$, $p=.35$; Figure 7.5b).

AD prototype online accuracy during self-managed tasks values of conditions II, III, and III* were also compared (Figure 7.5c). No statistical differences were detected between the three conditions as assessed by a one-way ANOVA ($F(2,21)=1.26$, $p=.30$).

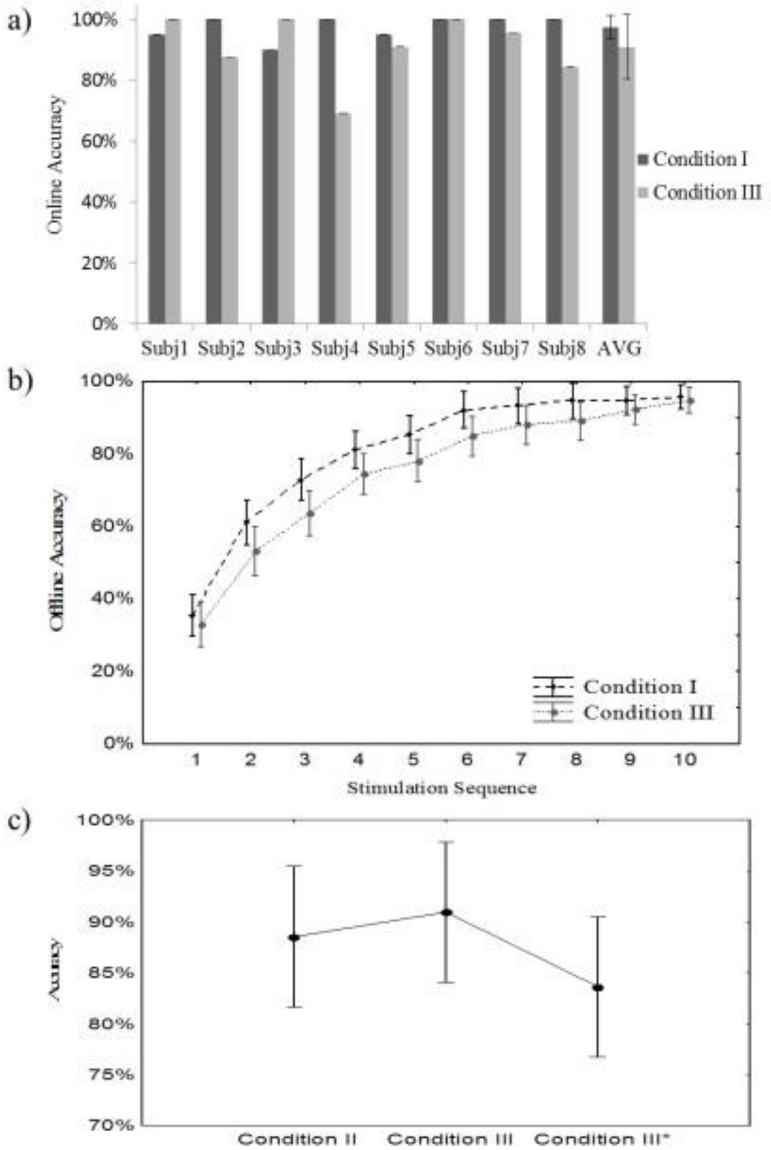

Figure 7.5 a) Online accuracy values achieved during the copy tasks of Condition I and III; b) mean accuracy trend condition I and III as a function of the number of stimulus repetitions delivered and assessed by mean of offline cross-validations; c) Online accuracy achieved on average during the self-managed tasks, under condition II and III. Condition III denoted values corresponding to the maximum end-user's WSR optimized for condition III.*

7.3.2 Efficiency

i) *BCI offline WSR* scores in conditions I and III were compared. No significant differences were revealed, as assessed by a repeated measure ANOVA with conditions as factors and WSR values per stimulation sequence as dependent variables (Figure 7.6a $F(9, 792)=1.33$, $p=.21$).

ii) *AD prototype time for correct selection* for conditions II, III and III* were compared by means of a one-way ANOVA. Condition II exhibited a significantly lower time per correct selection (8.31±6.81s on average) with respect to Condition III (31.69±7.59s on average) and Condition III* (19.43±9.3s on average; Figure 7.6b; $F(2, 21)=17.2$, $p<.01$)

Workload: On average, the workload was perceived as highest in condition I (see Table 2 for workload values). The total workload scores obtained from the 3 conditions were compared by means of a non-parametric Friedman ANOVA. No statistically significant differences between the three conditions were found (Friedman $\chi^2=3.2$, $p=0.19$).

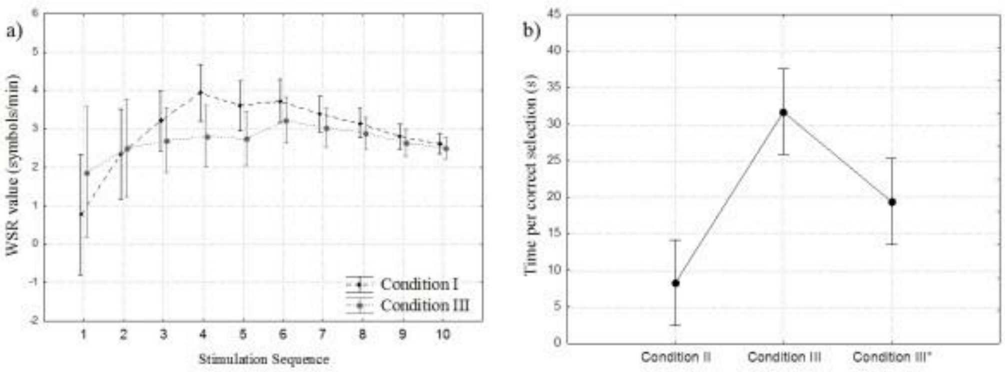

Figure 7.6 a) the WSR mean value for condition I and III assessed by mean of off-line cross-validations ; b) Time per correct selection detected on-line during self-managed tasks in condition II and III. Condition III refers to as the time per correct selection achievable using the number of stimuli repetitions corresponding to the maximum value of WSR.*

Table 7.2 Self-reported metrics

	Satisfaction visual analog scale (VAS - 0-10)			System usability scale (SUS – 0-100)			Wokload Nasa-tlx		
	Prototype			Prototype			Prototype		
	P300 Speller	Muscular input	BCI	P300 Speller	Muscular input	BCI	P300 Speller	Muscular input	BCI
Subj1	9.8	10	10	77.5	57.5	85	44	30.33	37.66
Subj2	10	9	8.3	82.5	84.28	55	14,33	10.66	27.66
Subj3	8.6	7.8	9	80	92,5	90	27.33	22.66	29.66
Subj4	9.7	10	10	42.5	70	70	60.33	56	32.28
Subj5	9.7	7	9	100	85	77.5	62.33	55.33	39.66
Subj6	10	10	10	95	100	92.5	17	29	13.33
Subj7	9	10	9	85	100	90	39	20.3	62.33
Subj8	9.6	9	9,3	82.5	85	82.5	28,33	27	15.66
Mean	**9.5**	**9.1**	**9.3**	**80.6**	**84.3**	**80.3**	**36.6**	**31.4**	**32.3**
SD	**0.5**	**1.1**	**0.6**	**17.2**	**14.5**	**12.6**	**18.2**	**16.2**	**15.3**

7.3.3 User satisfaction

VAS score was higher in condition III with respect to condition II. SUS score was higher in condition II with respect to conditions I and III. None of the differences between conditions reached significance, as determined by means of 2 non-parametric Friedman ANOVAs performed for the VAS scores (Friedman $\chi 2=0.24$, $p=0.88$) and SUS scores (Friedman $\chi 2=4.06$, $p=0.13$).

7.3.4 Selective attention and P300-based BCI performance

One participant (participant 8) was excluded from the analysis regarding the RSVP task, due to technical problems encountered during data recording. In brief, analysis on the scores collected by means of the RSVP task (T1% and T2%) were performed on seven participants. The offline BA in performing the BCI task was on average of 87.4% (SD=2.4%, range=84.5–92.3%, N=8). The mean amplitude for P300 amp in Cz, was 3.3μV (SD=1.6, range=1.1–6.5μV, N=8). In the RSVP task, mean accuracy of detection for T1 was 77.2% (SD=10.4%, range=65–96.25%, N=7) and for T2 67.7% (SD=14.1%, range=50.3–87.1%, N=7). A significant positive correlation was observed between T1% and the offline BA ($r=0.79, p<0.05$), showing that participants with higher T1% had a higher accuracy in the offline binary classification. To estimate the predictive value of T1% on the BA a regression analysis was performed which resulted in an $F=8.34$ with a $p<0.05$, indicating that the variance of the binary performance is predictable by the participant temporal filtering capacity, with$\beta=0.79$. A significant positive correlation was found between T1% and P300 amp in Cz (in$r=0.84, p<0.05$) showing that participants with higher T1% had a larger P300 amp in Cz. As a result of the linear regression, T1 accuracy was significantly predictive of P300 amp in Cz ($F=16,23$ with a$p<0.05$) with $\beta=0.87$. No significant correlation was found between T2%, the offline binary performance and P300 amp in Cz (Riccio et al., 2013).

7.4 Discussion

The aim of this study was to test the feasibility and to evaluate the usability of an AD prototype which was meant to provide people with severe motor impairment due to ALS with several applications to support communication and environmental control. Such AD prototype was endowed with a wide accessibility options, ranging from conventional/alternative input devices to a BCI device and it met the users' requirements mainly targeting everyday communication, such drawing the caregiver's attention to themselves and basic environmental interactions (such as turning on/off lights and control the television, see Figure 7.1).

We demonstrated the feasibility and the usability of the system, designed and developed to provide a multimodal access for communication and environmental control applications. Usability was assessed applying metrics derived from the UCD and adapted to evaluate BCI technology. No differences were found in terms of effectiveness between the three considered conditions. This result is in line with those of a previous work (Thompson et al., 2013): using a P300-BCI to control a complex user interface (condition III) does not affect system accuracy and WSR with respect to using a stand-alone BCI (P300-Speller, condition I) with a static menu interface. With respect to efficiency, the conventional/alternative input devices (condition II) resulted significant faster than the BCI (condition III) when used to access the proposed prototype. However this aspect did not affect the workload perceived by end users, which was not different in the three conditions. Finally users reported a high level of satisfaction with each of the conditions with BCI input (conditions I and III) exceeding in conventional/alternative input. This might be due to greater involvement and commitment required to the user in the use of BCI, which in case of success can be more satisfying. On the contrary SUS scores obtained when controlling the AD prototype with conventional/alternative input devices were higher than the other two considered conditions. This is likely due to the lower speed of BCI with respect to conventional/alternative devices and to the need for calibration of the former input modality.
In addition, despite some of the end-users had cognitive deficits, they were able to control the prototype by means of the BCI.

A BCI endowed as input channel in an AD system, is a step forward for the process of translation from the laboratory to daily life. Bringing such system in everyday life could potentially positively influence the perception of quality of life of end users. However such improvement can be expected by presenting to the user a usable aid, projected according to UCD principles. Finally we identified in the temporal filtering capacity, investigated by mean of the RSVP task, a predictor of both the P300-based BCI accuracy and of the amplitude of the P300 elicited performing the BCI task.

This feasibility study was conducted in controlled conditions involving 8 end-users with moderate motor impairment, still able to use conventional inputs. This limitation prevented us to evaluate the proposed system with standardized instruments (Dumont et al., 2002), which must be addressed in longitudinal studies involving a larger cohort of end-users using the system along the course of the disease in an ecological environment.

8 CONCLUSION

The transfer of BCI systems from research prototypes to assistive technologies accessible by end users (people with severe motor disabilities) in their own homes requires several improvements and technological solutions that cannot be addressed in a PhD thesis. At the beginning of the PhD course I identified some of the weak points of the existing P300-based BCI systems stressing their differences with the conventional and assistive input devices, and then I designed and evaluated solutions to reduce this gap.

In an effective and efficient interaction and communication the user should be able to decide the speed of the information delivery. On the contrary a system forcing the user to perform a choice or a selection every N seconds could result frustrating and wearisome, giving rise to unwanted selections and communication problems. Let us imagine a person, interacting with a PC by using a mouse, forced to choose an item (a character or an icon) and click on it every 20 seconds (neither faster or slower), otherwise the system would perform a random choice. This is the way of operation of synchronous P300-based BCIs and was the first problem addressed in this PhD thesis by defining and evaluating an asynchronous classifier. The latter exhibit considerable advantages as compared with synchronous modality, both in terms of system usability and communication efficiency. First, it is able to automatically suspend the control when the user does not attend to the stimulation, therefore, avoiding the need for an explicit “pause” command that the user should otherwise issue. Second, the system feature of “abstention” could avoid misclassification when the EEG feature is not sufficiently reliable. Finally, it can continuously adapt the time required for each classification to the changes of user state, finding an optimal trade-off between speed and accuracy. These results reflect in higher communication efficiency (in terms of speed of communication and errors recovery) of the asynchronous classifier with respect to a synchronous one.

Afterwards, the reliability of P300-based BCI accuracy has been investigated over repeated sessions in the same day. In order to simplify the calibration procedures of the system and ensure high reliability over time, an algorithm for the automatic and continuous adaptation of the classifier parameters was designed and validated involving 10 healthy subjects. Results showed that a continuous recalibration of the classifier parameters boosts the system performance among several sessions in the same day and that the proposed algorithm can perform a recalibration of the system using unlabeled data from on-line sessions and ensuring the stability of the performances. The proposed algorithm represents a step forward for increasing the usability of BCI systems as assistive technology since after an initial supervised calibration session, the whole recalibration procedure is hidden to the user.

Another important point for the home independent use of BCI systems is the possibility of detect and report possible failures in the acquisition of the EEG signal without the need for specific technical knowledge. For this reason the influence of some common artifacts on BCI performance were investigated and an algorithm for the online assessment of the EEG "quality" was designed and validated over 5 healthy subjects.

The last section of this thesis presents a work that was carried out in the context of the Brindisys project (*Brain-Computer interface devices to support individual autonomy in locked-in individuals*, an Italian project funded by the Italian Agency for the research on the Amyotrophic lateral sclerosis – ARISLA). The latter aimed at designing and developing and assistive device for communication and environmental control accessible by means of several input devices, including a self-contained P300-based BCI, to support and/or restore interaction with the external world in people with ALS during the different stages of the disease. The preliminary validation of such device indicated the potential effectiveness and usability of the proposed system. In fact, no differences in effectiveness were found between the proposed prototype accessed both using muscular inputs and BCI with respect to a widely validated P300 based BCI interface. As expected, the efficiency of the prototype when operated by mean of the BCI was significantly lower in terms of time to achieve a correct selection. However, this aspect did not affect negatively the workload perceived by subjects. Finally the end users stated that they were highly satisfied of the proposed system used both with BCI and muscular input device.

In conclusion, this thesis provides some solutions to improve the overall usability of P300-based systems as assistive devices for people with severe motor impairments, but this is just one piece of a larger mosaic: further improvement are required in different fields. First, the design and the development of BCI-based assistive technologies should be focused on the user-centered design approach. This aspect requires the collaboration of different professional figures to deal with the problem from different points of view, such as bioengineers, HCI experts, assistive technology stakeholders, medical doctors and psychologists. Finally, the active participation of companies from biomedical and assistive technology fields will give an important contribution for the development of more reliable and user friendly devices, translating the current existing prototypes in commercial devices.

9 References

[1] ISO 9241-210:2010, Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems. Multiple. Distributed through American National Standards Institute.

[2] Al-Chalabi, A., Hardiman, O., 2013. The epidemiology of ALS: a conspiracy of genes, environment and time. Nat. Rev. Neurol. 9, 617–628. doi:10.1038/nrneurol.2013.203

[3] Allison, B.Z., Pineda, J.A., 2006. Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: implications for a BCI system. Int. J. Psychophysiol. Off. J. Int. Organ. Psychophysiol. 59, 127–140. doi:10.1016/j.ijpsycho.2005.02.007

[4] Aloise, F., Aricò, P., Schettini, F., Riccio, A., Salinari, S., Mattia, D., Babiloni, F., Cincotti, F., 2012. A covert attention P300-based brain-computer interface: Geospell. Ergonomics 55, 538–551. doi:10.1080/00140139.2012.661084

[5] Aloise, F., Schettini, F., Aricò, P., Leotta, F., Salinari, S., Mattia, D., Babiloni, F., Cincotti, F., 2011. P300-based brain-computer interface for environmental control: an asynchronous approach. J. Neural Eng. 8, 025025. doi:10.1088/1741-2560/8/2/025025

[6] Andersen, P.M., Borasio, G.D., Dengler, R., Hardiman, O., Kollewe, K., Leigh, P.N., Pradat, P.-F., Silani, V., Tomik, B., EFNS Task Force on Diagnosis and Management of Amyotrophic Lateral Sclerosis, 2005. EFNS task force on management of amyotrophic lateral sclerosis: guidelines for diagnosing and clinical care of patients and relatives. Eur. J. Neurol. Off. J. Eur. Fed. Neurol. Soc. 12, 921–938. doi:10.1111/j.1468-1331.2005.01351.x

[7] Baldoni, R., Di Ciccio, C., Mecella, M., Patrizi, F., Querzoni, L., Santucci, G., Dustdar, S., Li, F., Truong, H.-L., Albornos, L., Milagro, F., Antolin Rafael, P., Ayani, R., Rasch, K., Garcia Lozano, M., Aiello, M., Lazovik, A., Denaro, A., Lasala, G., Pucci, P., Holzner, C., Cincotti, F., Aloise, F., 2009. An Embedded Middleware Platform for Pervasive and Immersive Environments for-All, in: 2009 6th IEEE Annual Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks Workshops. Presented at the 2009 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks Workshops, Rome, Italy, pp. 1–3. doi:10.1109/SAHCNW.2009.5172921

[8] Barachant, A., Andreev, A., Congedo, M., 2013. The Riemannian Potato: an automatic and adaptive artifact detection method for online experiments using Riemannian geometry, in: Proceedings of TOBI Workshop IV. Sion, Suisse, pp. 19–20.

[9] Bensch, M., Karim, A.A., Mellinger, J., Hinterberger, T., Tangermann, M., Bogdan, M., Rosenstiel, W., Birbaumer, N., 2007. Nessi: An EEG-Controlled

Web Browser for Severely Paralyzed Patients. Comput. Intell. Neurosci. 2007. doi:10.1155/2007/71863

[10] Bianchi, L., Quitadamo, L.R., Garreffa, G., Cardarilli, G.C., Marciani, M.G., 2007. Performances evaluation and optimization of brain computer interface systems in a copy spelling task. IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc. 15, 207–216. doi:10.1109/TNSRE.2007.897024

[11] Birbaumer, N., Ghanayim, N., Hinterberger, T., Iversen, I., Kotchoubey, B., Kübler, A., Perelmouter, J., Taub, E., Flor, H., 1999. A spelling device for the paralysed. Nature 398, 297–298. doi:10.1038/18581

[12] Blain-Moraes, S., Schaff, R., Gruis, K.L., Huggins, J.E., Wren, P.A., 2012. Barriers to and mediators of brain-computer interface user acceptance: focus group findings. Ergonomics 55, 516–525. doi:10.1080/00140139.2012.661082

[13] Blanchard, S., 2005. 3 - ANATOMY AND PHYSIOLOGY, in: Enderle, J.D., Blanchard, S.M., Bronzino, J.D. (Eds.), Introduction to Biomedical Engineering (Second Edition), Biomedical Engineering. Academic Press, Boston, pp. 73–125.

[14] Blankertz, B., Lemm, S., Treder, M., Haufe, S., Müller, K.-R., 2011. Single-trial analysis and classification of ERP components--a tutorial. NeuroImage 56, 814–825. doi:10.1016/j.neuroimage.2010.06.048

[15] Brouwer, A.-M., van Erp, J.B.F., 2010. A tactile P300 brain-computer interface. Front. Neurosci. 4, 19. doi:10.3389/fnins.2010.00019

[16] Cano, M.E., Class, Q.A., Polich, J., 2009. Affective valence, stimulus attributes, and P300: color vs. black/white and normal vs. scrambled images. Int. J. Psychophysiol. Off. J. Int. Organ. Psychophysiol. 71, 17–24. doi:10.1016/j.ijpsycho.2008.07.016

[17] Caruso, M., Cincotti, F., Leotta, F., Mecella, M., Riccio, A., Schettini, F., Simione, L., Catarci, T., 2013. My-World-in-My-Tablet: An Architecture for People with Physical Impairment, in: Kurosu, M. (Ed.), Human-Computer Interaction. Interaction Modalities and Techniques. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 637–647.

[18] Cedarbaum, J.M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond, B., Nakanishi, A., 1999. The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. J. Neurol. Sci. 169, 13–21. doi:10.1016/S0022-510X(99)00210-5

[19] Cipresso, P., Carelli, L., Solca, F., Meazzi, D., Meriggi, P., Poletti, B., Lulé, D., Ludolph, A.C., Silani, V., Riva, G., 2012. The use of P300-based BCIs in amyotrophic lateral sclerosis: from augmentative and alternative communication to cognitive assessment. Brain Behav. 2, 479–498. doi:10.1002/brb3.57

[20] Congedo, M., Gouy-Pailler, C., Jutten, C., 2008. On the blind source separation of human electroencephalogram by approximate joint diagonalization of second order statistics (arXiv e-print No. 0812.0494).

[21] Del Millan, J.R., Hauser, A., Renkens, F., 2002. Adaptive brain interfaces - ABI: simple features, simple neural network, complex brain-actuated devices, in: 2002 14th International Conference on Digital Signal Processing, 2002. DSP 2002. Presented at the 2002 14th International Conference on Digital Signal Processing, 2002. DSP 2002, pp. 297–300 vol.1. doi:10.1109/ICDSP.2002.1027891

[22] Del R Millan, J.J., Galan, F., Vanhooydonck, D., Lew, E., Philips, J., Nuttin, M., 2009. Asynchronous non-invasive brain-actuated control of an intelligent wheelchair. Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Conf. 2009, 3361–3364. doi:10.1109/IEMBS.2009.5332828

[23] Diez, P.F., Mut, V.A., Avila Perona, E.M., Laciar Leber, E., 2011. Asynchronous BCI control using high-frequency SSVEP. J. NeuroEngineering Rehabil. 8, 39. doi:10.1186/1743-0003-8-39

[24] Draper, N.R., Smith, H., 1998. Applied Regression Analysis, Third. ed. Wiley-Interscience.

[25] Dumont, C., Vincent, C., Mazer, B., 2002. Development of a standardized instrument to assess computer task performance. Am. J. Occup. Ther. Off. Publ. Am. Occup. Ther. Assoc. 56, 60–68.

[26] Duncan, C.C., Mirsky, A.F., Lovelace, C.T., Theodore, W.H., 2009. Assessment of the attention impairment in absence epilepsy: comparison of visual and auditory P300. Int. J. Psychophysiol. Off. J. Int. Organ. Psychophysiol. 73, 118–122. doi:10.1016/j.ijpsycho.2009.03.005

[27] Duncan-Johnson, C.C., Donchin, E., 1977. On quantifying surprise: the variation of event-related potentials with subjective probability. Psychophysiology 14, 456–467.

[28] F I Mahoney, Barthel, D.W., 1965. Functional evaluation: the Barthel index. Md. State Med. J. 14, 61–65.

[29] Fabiani, M., Gratton, G., Coles, M., Cacioppo, J., Tassinary, L., Bertson, G., 2000. Event-related brain potentials - methods, theory, and applications, in: Handbook of Psychophysiology. Cambridge: University Press, pp. 55–83.

[30] Farwell, L.A., Donchin, E., 1988. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroencephalogr. Clin. Neurophysiol. 70, 510–523.

[31] Folstein, M.F., Folstein, S.E., McHugh, P.R., 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res. 12, 189–198.

[32] Förstner, W., Moonen, B., 1999. A Metric for Covariance Matrices.

[33] Furdea, A., Halder, S., Krusienski, D.J., Bross, D., Nijboer, F., Birbaumer, N., Kübler, A., 2009. An auditory oddball (P300) spelling system for brain-

computer interfaces. Psychophysiology 46, 617–625. doi:10.1111/j.1469-8986.2008.00783.x

[34] Gao, X., Xu, D., Cheng, M., Gao, S., 2003. A BCI-based environmental controller for the motion-disabled. IEEE Trans. Neural Syst. Rehabil. Eng. 11, 137–140. doi:10.1109/TNSRE.2003.814449

[35] Geisler, M.W., Polich, J., 1992. P300 and Individual Differences: Morning/Evening Activity Preference, Food, and Time-of-Day. Psychophysiology 29, 86–94. doi:10.1111/j.1469-8986.1992.tb02019.x

[36] Gonsalvez, C.L., Polich, J., 2002. P300 amplitude is determined by target-to-target interval. Psychophysiology 39, 388–396.

[37] Goodin, D.S., Squires, K.C., Starr, A., 1978. Long latency event-related components of the auditory evoked potential in dementia. Brain J. Neurol. 101, 635–648.

[38] Guger, C., Daban, S., Sellers, E., Holzner, C., Krausz, G., Carabalona, R., Gramatica, F., Edlinger, G., 2009. How many people are able to control a P300-based brain-computer interface (BCI)? Neurosci. Lett. 462, 94–98. doi:10.1016/j.neulet.2009.06.045

[39] Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research, in: Human Mental Workload. North-Holland, pp. 139–183.

[40] Heaton, R.K., Avitable, N., Grant, I., Matthews, C.G., 1999. Further crossvalidation of regression-based neuropsychological norms with an update for the Boston Naming Test. J. Clin. Exp. Neuropsychol. 21, 572–582. doi:10.1076/jcen.21.4.572.882

[41] Horne, J.A., Ostberg, O., 1976. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. Int. J. Chronobiol. 4, 97–110.

[42] Huang, J., Katsuura, T., Shimomura, Y., Iwanaga, K., 2006. Diurnal changes of ERP response to sound stimuli of varying frequency in morning-type and evening-type subjects. J. Physiol. Anthropol. 25, 49–54.

[43] Huggins, J.E., Wren, P.A., Gruis, K.L., 2011. What would brain-computer interface users want? Opinions and priorities of potential users with amyotrophic lateral sclerosis. Amyotroph. Lateral Scler. Off. Publ. World Fed. Neurol. Res. Group Mot. Neuron Dis. 12, 318–324. doi:10.3109/17482968.2011.572978

[44] Jin, J., Allison, B.Z., Sellers, E.W., Brunner, C., Horki, P., Wang, X., Neuper, C., 2011. An adaptive P300-based control system. J. Neural Eng. 8, 036006. doi:10.1088/1741-2560/8/3/036006

[45] Jurcak, V., Tsuzuki, D., Dan, I., 2007. 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems. NeuroImage 34, 1600–1611. doi:10.1016/j.neuroimage.2006.09.024

[46] Kaldeli, E., Warriach, E., Bresser, J., Lazovik, A., Aiello, M., 2010. Interoperation, Composition and Simulation of Services at Home. Serv.-Oriented Comput. 167–181.

[47] Karim, A.A., Hinterberger, T., Richter, J., Mellinger, J., Neumann, N., Flor, H., Kübler, A., Birbaumer, N., 2006. Neural internet: Web surfing with brain potentials for the completely paralyzed. Neurorehabil. Neural Repair 20, 508–515. doi:10.1177/1545968306290661

[48] Kaufmann, T., Völker, S., Gunesch, L., Kübler, A., 2012. Spelling is Just a Click Away - A User-Centered Brain-Computer Interface Including Auto-Calibration and Predictive Text Entry. Front. Neurosci. 6, 72. doi:10.3389/fnins.2012.00072

[49] Kindermans, P.-J., Verstraeten, D., Schrauwen, B., 2012. A bayesian model for exploiting application constraints to enable unsupervised training of a P300-based BCI. PloS One 7, e33758. doi:10.1371/journal.pone.0033758

[50] Kleih, S.C., Kaufmann, T., Zickler, C., Halder, S., Leotta, F., Cincotti, F., Aloise, F., Riccio, A., Herbert, C., Mattia, D., Kübler, A., 2011. Out of the frying pan into the fire--the P300-based BCI faces real-world challenges. Prog. Brain Res. 194, 27–46. doi:10.1016/B978-0-444-53815-4.00019-4

[51] Kranczioch, C., Debener, S., Maye, A., Engel, A.K., 2007. Temporal dynamics of access to consciousness in the attentional blink. NeuroImage 37, 947–955. doi:10.1016/j.neuroimage.2007.05.044

[52] Krusienski, D.J., Sellers, E.W., Cabestaing, F., Bayoudh, S., McFarland, D.J., Vaughan, T.M., Wolpaw, J.R., 2006. A comparison of classification techniques for the P300 Speller. J. Neural Eng. 3, 299–305. doi:10.1088/1741-2560/3/4/007

[53] Krusienski, D.J., Sellers, E.W., McFarland, D.J., Vaughan, T.M., Wolpaw, J.R., 2008. Toward enhanced P300 speller performance. J. Neurosci. Methods 167, 15–21. doi:10.1016/j.jneumeth.2007.07.017

[54] Kubler, A., Holz, E., Kaufmann, T., Zickler, C., 2013. A User Centred Approach for Bringing BCI Controlled Applications to End-Users, in: Fazel-Rezai, R. (Ed.), Brain-Computer Interface Systems - Recent Progress and Future Prospects. InTech.

[55] Kübler, A., Neumann, N., Wilhelm, B., Hinterberger, T., Birbaumer, N., 2004. Predictability of Brain-Computer Communication. J. Psychophysiol. 18, 121–129. doi:10.1027/0269-8803.18.23.121

[56] Kübler, A., Nijboer, F., Mellinger, J., Vaughan, T.M., Pawelzik, H., Schalk, G., McFarland, D.J., Birbaumer, N., Wolpaw, J.R., 2005. Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. Neurology 64, 1775–1777. doi:10.1212/01.WNL.0000158616.43002.6D

[57] Kübler, A., Zickler, C., Holz, E., Kaufmann, T., Riccio, A., Mattia, D., 2013. Applying the user-centred design to evaluation of Brain-Computer Interface controlled applications. Biomed. Tech. (Berl). doi:10.1515/bmt-2013-4438

[58] Kurtzke, J.F., 1983. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 33, 1444–1452.

[59] Leclercq, M., Zimmermann, P., 2002. Applied Neuropsychology of Attention: Theory, Diagnosis and Rehabilitation. Psychology Press.

[60] Leeb, R., Sagha, H., Chavarriaga, R., Millán, J.D.R., 2011. A hybrid brain-computer interface based on the fusion of electroencephalographic and electromyographic activities. J. Neural Eng. 8, 025011. doi:10.1088/1741-2560/8/2/025011

[61] Lenhardt, A., Kaper, M., Ritter, H.J., 2008. An adaptive P300-based online brain-computer interface. IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc. 16, 121–130. doi:10.1109/TNSRE.2007.912816

[62] Linden, D.E.J., 2005. The P300: Where in the Brain Is It Produced and What Does It Tell Us? The Neuroscientist 11, 563 –576. doi:10.1177/1073858405280524

[63] Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., Arnaldi, B., 2007. A review of classification algorithms for EEG-based brain-computer interfaces. J. Neural Eng. 4, R1–R13. doi:10.1088/1741-2560/4/2/R01

[64] Lu, S., Guan, C., Zhang, H., 2009. Unsupervised brain computer interface based on intersubject information and online adaptation. IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc. 17, 135–145. doi:10.1109/TNSRE.2009.2015197

[65] Lutz, A., Slagter, H.A., Rawlings, N.B., Francis, A.D., Greischar, L.L., Davidson, R.J., 2009. Mental training enhances attentional stability: neural and behavioral evidence. J. Neurosci. Off. J. Soc. Neurosci. 29, 13418–13427. doi:10.1523/JNEUROSCI.1614-09.2009

[66] Mason, S.G., Birch, G.E., 2000. A brain-controlled switch for asynchronous control applications. IEEE Trans. Biomed. Eng. 47, 1297–1307. doi:10.1109/10.871402

[67] McFarland, D.J., Sarnacki, W.A., Wolpaw, J.R., 2011. Should the parameters of a BCI translation algorithm be continually adapted? J. Neurosci. Methods 199, 103–107. doi:10.1016/j.jneumeth.2011.04.037

[68] Millán, J. del R., Renkens, F., Mouriño, J., Gerstner, W., 2004. Noninvasive brain-actuated control of a mobile robot by human EEG. IEEE Trans. Biomed. Eng. 51, 1026–1033. doi:10.1109/TBME.2004.827086

[69] Millán, J.D.R., Rupp, R., Müller-Putz, G.R., Murray-Smith, R., Giugliemma, C., Tangermann, M., Vidaurre, C., Cincotti, F., Kübler, A., Leeb, R., Neuper, C., Müller, K.-R., Mattia, D., 2010. Combining Brain-Computer Interfaces and Assistive Technologies: State-of-the-Art and Challenges. Front. Neurosci. 4. doi:10.3389/fnins.2010.00161

[70] Mugler, E.M., Ruf, C.A., Halder, S., Bensch, M., Kler, A., 2010. Design and Implementation of a P300-Based Brain-Computer Interface for Controlling an

Internet Browser. IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc. doi:10.1109/TNSRE.2010.2068059

[71] Muller, K., Blankertz, B., 2006. Toward noninvasive brain-computer interfaces. IEEE Signal Process. Mag. 23, 128–126. doi:10.1109/MSP.2006.1708426

[72] Niedermeyer, E., Silva, F.H.L. da, 2005. Electroencephalography: Basic Principles, Clinical Applications, and Related Fields. Lippincott Williams & Wilkins.

[73] Nijboer, F., Sellers, E.W., Mellinger, J., Jordan, M.A., Matuz, T., Furdea, A., Halder, S., Mochty, U., Krusienski, D.J., Vaughan, T.M., Wolpaw, J.R., Birbaumer, N., Kübler, A., 2008. A P300-based brain-computer interface for people with amyotrophic lateral sclerosis. Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol. 119, 1909–1916. doi:10.1016/j.clinph.2008.03.034

[74] Panicker, R., Puthusserypady, S., Sun, Y., 2010. Adaptation in P300 Brain-Computer Interfaces: A Two-Classifier Co-Training Approach. IEEE Trans. Biomed. Eng. doi:10.1109/TBME.2010.2058804

[75] Panicker, R.C., Puthusserypady, S., Sun, Y., 2011. An asynchronous P300 BCI with SSVEP-based control state detection. IEEE Trans. Biomed. Eng. 58, 1781–1788. doi:10.1109/TBME.2011.2116018

[76] Parra, L.C., Spence, C.D., Gerson, A.D., Sajda, P., 2003. Response error correction--a demonstration of improved human-machine performance using real-time EEG monitoring. IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc. 11, 173–177. doi:10.1109/TNSRE.2003.814446

[77] Piccione, F., Giorgi, F., Tonin, P., Priftis, K., Giove, S., Silvoni, S., Palmas, G., Beverina, F., 2006. P300-based brain computer interface: reliability and performance in healthy and paralysed participants. Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol. 117, 531–537. doi:10.1016/j.clinph.2005.07.024

[78] Picton, T.W., Bentin, S., Berg, P., Donchin, E., Hillyard, S.A., Johnson, R., Miller, G.A., Ritter, W., Ruchkin, D.S., Rugg, M.D., Taylor, M.J., 2000. Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. Psychophysiology 37, 127–152.

[79] Polich, J., 1997. On the relationship between EEG and P300: individual differences, aging, and ultradian rhythms. Int. J. Psychophysiol. 26, 299–317. doi:10.1016/S0167-8760(97)00772-1

[80] Polich, J., 2004. Clinical application of the P300 event-related brain potential. Phys. Med. Rehabil. Clin. N. Am. 15, 133–161.

[81] Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol. 118, 2128–2148. doi:10.1016/j.clinph.2007.04.019

[82] Polich, J., Bondurant, T., 1997. P300 sequence effects, probability, and interstimulus interval. Physiol. Behav. 61, 843–849.

[83] Polich, J., Criado, J.R., 2006. Neuropsychology and neuropharmacology of P3a and P3b. Int. J. Psychophysiol. Off. J. Int. Organ. Psychophysiol. 60, 172–185. doi:10.1016/j.ijpsycho.2005.12.012

[84] Polich, J., Kok, A., 1995. Cognitive and biological determinants of P300: an integrative review. Biol. Psychol. 41, 103–146.

[85] Quitadamo, L.R., Abbafati, M., Cardarilli, G.C., Mattia, D., Cincotti, F., Babiloni, F., Marciani, M.G., Bianchi, L., 2011. Evaluation of the performances of different P300 based brain-computer interfaces by means of the efficiency metric. J. Neurosci. Methods. doi:10.1016/j.jneumeth.2011.10.010

[86] Rankin, J., 1957. Cerebral vascular accidents in patients over the age of 60. II. Prognosis. Scott. Med. J. 2, 200–215.

[87] Ravden, D., Polich, J., 1999. On P300 measurement stability: habituation, intra-trial block variation, and ultradian rhythms. Biol. Psychol. 51, 59–76. doi:10.1016/S0301-0511(99)00015-0

[88] Riccio, A., Leotta, F., Bianchi, L., Aloise, F., Zickler, C., Hoogerwerf, E.-J., Kübler, A., Mattia, D., Cincotti, F., 2011. Workload measurement in a communication application operated through a P300-based brain-computer interface. J. Neural Eng. 8, 025028. doi:10.1088/1741-2560/8/2/025028

[89] Riccio, A., Simione, L., Schettini, F., Pizzimenti, A., Inghilleri, M., Belardinelli, M.O., Mattia, D., Cincotti, F., 2013. Attention and P300-based BCI performance in people with amyotrophic lateral sclerosis. Front. Hum. Neurosci. 7:, 732. doi:10.3389/fnhum.2013.00732

[90] Schalk, G., McFarland, D.J., Hinterberger, T., Birbaumer, N., Wolpaw, J.R., 2004. BCI2000: a general-purpose brain-computer interface (BCI) system. IEEE Trans. Biomed. Eng. 51, 1034–1043. doi:10.1109/TBME.2004.827072

[91] Scherer, R., Muller, G.R., Neuper, C., Graimann, B., Pfurtscheller, G., 2004. An asynchronously controlled EEG-based virtual keyboard: improvement of the spelling rate. IEEE Trans. Biomed. Eng. 51, 979–984. doi:10.1109/TBME.2004.827062

[92] Schettini, F., Aloise, F., Aricò, P., Salinari, S., Mattia, D., Cincotti, F., 2012. Control or No-Control? Reducing the gap between Brain-Computer Interface and classical input devices, in: Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications - ETRA '10. Presented at the the 2010 Symposium, Austin, Texas, p. 77. doi:10.1145/1743666.1743685

[93] Schreuder, M., Rost, T., Tangermann, M., 2011. Listen, You are Writing! Speeding up Online Spelling with a Dynamic Auditory BCI. Front. Neurosci. 5, 112. doi:10.3389/fnins.2011.00112

[94] Schulzer, M., 1994. Diagnostic tests: a statistical review. Muscle Nerve 17, 815–819. doi:10.1002/mus.880170719

[95] Sellers, E.W., Kübler, A., Donchin, E., 2006. Brain-computer interface research at the University of South Florida Cognitive Psychophysiology Laboratory: the P300 Speller. IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc. 14, 221–224. doi:10.1109/TNSRE.2006.875580

[96] Sellers, E.W., Vaughan, T.M., Wolpaw, J.R., 2010. A brain-computer interface for long-term independent home use. Amyotroph. Lateral Scler. Off. Publ. World Fed. Neurol. Res. Group Mot. Neuron Dis. 11, 449–455. doi:10.3109/17482961003777470

[97] Silvoni, S., Volpato, C., Cavinato, M., Marchetti, M., Priftis, K., Merico, A., Tonin, P., Koutsikos, K., Beverina, F., Piccione, F., 2009. P300-Based Brain-Computer Interface Communication: Evaluation and Follow-up in Amyotrophic Lateral Sclerosis. Front. Neurosci. 3, 60. doi:10.3389/neuro.20.001.2009

[98] Thompson, D.E., Gruis, K.L., Huggins, J.E., 2013. A plug-and-play brain-computer interface to operate commercial assistive technology. Disabil. Rehabil. Assist. Technol. 1–7. doi:10.3109/17483107.2013.785036

[99] Thompson, D.E., Warschausky, S., Huggins, J.E., 2012. Classifier-based latency estimation: a novel way to estimate and predict BCI accuracy. J. Neural Eng. 10, 016006. doi:10.1088/1741-2560/10/1/016006

[100] Townsend, G., Graimann, B., Pfurtscheller, G., 2004. Continuous EEG classification during motor imagery--simulation of an asynchronous BCI. IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc. 12, 258–265. doi:10.1109/TNSRE.2004.827220

[101] Townsend, G., LaPallo, B.K., Boulay, C.B., Krusienski, D.J., Frye, G.E., Hauser, C.K., Schwartz, N.E., Vaughan, T.M., Wolpaw, J.R., Sellers, E.W., 2010. A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns. Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol. 121, 1109–1120. doi:10.1016/j.clinph.2010.01.030

[102] Treder, M.S., Blankertz, B., 2010. (C)overt attention and visual speller design in an ERP-based brain-computer interface. Behav. Brain Funct. BBF 6, 28. doi:10.1186/1744-9081-6-28

[103] Valbuena, D., Volosyak, I., Graser, A., 2010. sBCI: fast detection of steady-state visual evoked potentials. Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Conf. 2010, 3966–3969. doi:10.1109/IEMBS.2010.5627990

[104] Vaughan, T.M., McFarland, D.J., Schalk, G., Sarnacki, W.A., Krusienski, D.J., Sellers, E.W., Wolpaw, J.R., 2006. The Wadsworth BCI Research and Development Program: at home with BCI. IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc. 14, 229–233. doi:10.1109/TNSRE.2006.875577

[105] Warriach, E., Kaldeli, E., Bresser, J., Lazovik, A., Aiello, M., 2010. A Tool for Integrating Pervasive Services and Simulating Their Composition. Serv.-Oriented Comput. 726–727.

[106] Williamson, J., Murray-Smith, R., Blankertz, B., Krauledat, M., Müller, K.-R., 2009. Designing for uncertain, asymmetric control: Interaction design for brain–computer interfaces. Int. J. Hum.-Comput. Stud. 67, 827–841. doi:10.1016/j.ijhcs.2009.05.009

[107] Wolpaw, J., Wolpaw, E.W., 2012. Brain-Computer Interfaces: Principles and Practice. Oxford University Press.

[108] Wolpaw, J.R., Birbaumer, N., Heetderks, W.J., McFarland, D.J., Peckham, P.H., Schalk, G., Donchin, E., Quatrano, L.A., Robinson, C.J., Vaughan, T.M., 2000. Brain-computer interface technology: a review of the first international meeting. IEEE Trans. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc. 8, 164–173.

[109] Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., Vaughan, T.M., 2002. Brain-computer interfaces for communication and control. Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol. 113, 767–791.

[110] Yagi, Y., Coburn, K.L., Estes, K.M., Arruda, J.E., 1999. Effects of aerobic exercise and gender on visual and auditory P300, reaction time, and accuracy. Eur. J. Appl. Physiol. 80, 402–408.

[111] Zhang, D., Song, H., Xu, H., Wu, W., Gao, S., Hong, B., 2012. An N200 speller integrating the spatial profile for the detection of the non-control state. J. Neural Eng. 9, 026016. doi:10.1088/1741-2560/9/2/026016

[112] Zhang, H., Guan, C., Wang, C., 2008. Asynchronous P300-based brain-computer interfaces: a computational approach with statistical models. IEEE Trans. Biomed. Eng. 55, 1754–1763.

[113] Zickler, C., Riccio, A., Leotta, F., Hillian-Tress, S., Halder, S., Holz, E., Staiger-Sälzer, P., Hoogerwerf, E.-J., Desideri, L., Mattia, D., Kübler, A., 2011. A brain-computer interface as input channel for a standard assistive technology software. Clin. EEG Neurosci. 42, 236–244.

[114] Zweig, M.H., Campbell, G., 1993. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin. Chem. 39, 561–577.

PUBLICATIONS

Full Paper

[J 1] **F. Schettini**, F. Aloise, P.Aricò, S. Salinari, D. Mattia, F. Cincotti "Self-calibration algorithm in an asynchronous P300-based brain computer interface" Accepted for publication on Journal of Neural Engineering

[J 2] **F. Schettini**; A. Riccio; L. Simione; G. Liberati; M. Caruso; V. Frasca; B.Calabrese; M. Mecella; A. Pizzimenti; M. Inghilleri; F. Cincotti. "An assistive device with conventional, alternative and Brain-computer interface inputs to enhance interaction with the environment for people with Amyotrophic Lateral Sclerosis: a feasibility and usability study" Conditionally accepted for publication on Archives of Physical Medicine and Rehabilitation

[J 3] P.Aricò, F.Aloise, **F.Schettini**, S.Salinari, D.Mattia, F.Cincotti. "Influence of P300 latency jitter over ERPs based BCIs performance". Accepted for publication on Journal of Neural Engineering

[J 4] A. Riccio, L. Simione, **F. Schettini**, A. Pizzimenti, M. Inghilleri, M. O. Belardinelli, D. Mattia, e F. Cincotti, "Attention and P300-based BCI performance in people with amyotrophic lateral sclerosis", Front. Hum. Neurosci., vol. 7:, pag. 732, 2013

[J 5] F. Aloise, P. Aricò, F. Schettini, S. Salinari, D. Mattia, and F. Cincotti, "Asynchronous gaze-independent event-related potential-based brain–computer interface," Artificial Intelligence in Medicine, vol. 59, no. 2, pp. 61–69, Oct. 2013.

[J 6] F. Aloise, F. Schettini, P. Aricò, S. Salinari, F. Babiloni, e F. Cincotti, "A comparison of classification techniques for a gaze-independent P300-based brain-computer interface", J Neural Eng, vol. 9, n. 4, pag. 045012, ago. 2012.

[J 7] F. Aloise, P. Aricò, **F. Schettini**, A. Riccio, S. Salinari, D. Mattia, F. Babiloni, F. Cincotti. "A Covert Attention P300-based Brain Computer Interface: GeoSpell". Ergonomics, vol. 55, n°. 5, pp. 538–551, May 2012;

[J 8] **F. Schettini**, F. Aloise, P. Aricò, F. Leotta, S. Salinari, F. Babiloni, D. Mattia, F. Cincotti. "Improving Asynchronous Control for P300-based BCI: towards a completely autoadaptative system". International Journal of Bioelectromagnetism, Vol. 13, No. pp. 150 – 151, 2011; ISSN: 1456-7865

[J 9] F. Aloise, P. Aricò, **F. Schettini**, E. Lucano, S.Salinari, F.Babiloni, D. Mattia, F. Cincotti. "Can the P300-based BCI training affect the ERPs?" International Journal of Bioelectromagnetism, Vol. 13, No. 3, pp. 148 – 149, 2011; ISSN: 1456-7865

[J 10] P. Aricò, F. Aloise, **F. Schettini**, A. Riccio, S.Salinari, F. Babiloni, D. Mattia, F. Cincotti. "GeoSpell: an alternative P300-based speller interface towards no eye gaze required". International Journal of Bioelectromagnetism, Vol. 13, No. 3, pp. 152 – 153, 2011; ISSN: 1456-7865

[J 11] R. Ortner, F. Aloise, R. Prückl, **F. Schettini**, V. Putz, J. Scharinger, E. Opisso, U. Costa, e C. Guger, "Accuracy of a P300 speller for people with motor impairments:

a comparison", Clin EEG Neurosci, vol. 42, n°. 4, pp. 214–218, Oct 2011; doi:10.1177/155005941104200405

[J 12] F.Aloise, **F.Schettini**, P.Aricò, S.Salinari, C.Guger, J.Rinsma, M.Aiello, D.Mattia, F.Cincotti. "Asynchronous P300-based BCI to control a virtual environment: initial tests on end users". Clin EEG Neurosci, vol. 42, n°. 4, pp. 219–224, Oct 2011. doi:10.1177/155005941104200406

[J 13] F. Aloise, **F. Schettini**, P. Arico, F. Leotta, S. Salinari, D. Mattia, F. Babiloni, F. Cincotti. "P300-based brain–computer interface for environmental control: an Asynchronous approach" J. Neural Eng. 8 (2011) 025025; doi: 10.1088/1741-2560/8/2/025025

[J 14] F. Aloise, I. Lasorsa, **F. Schettini**, AM Brouwer, D. Mattia, F. Babiloni, S. Salinari, M.G. Marciani, F. Cincotti. "Multimodal stimulation for a P300-based BCI". International Journal of Bioelectromagnetism 2007, Vol. 9, No. 3 ISSN: 1456-7865.

Conference Proceedings

[C 1] A. Riccio, L. Simione, **F. Schettini**, A. Pizzimenti, M. Inghilleri, M. Olivetti, D. Mattia, F. Cincotti. "Brain Computer interface and amyotrophic lateral sclerosis: attention and P300-based BCI performance". 18th Conference of the European Society for Cognitive Psychology. Budapest, Hungary, 29th August -1st September, 2013.

[C 2] **F. Schettini**, A. Riccio, L. Simione, G. Liberati, M. Caruso, B. Calabrese, N. Ielpo, A. Palumbo, V. Frasca, M. Mecella, F. Amato, A. Pizzimenti, M. Inghilleri, D. Mattia and F. Cincotti."From touch to brain control: augmenting communication in persons with ALS". 12th European AAATE Conference. Vilamoura, Algarve, Portugal; 19-22 September 2013.

[C 3] **F. Schettini**, A. Riccio, L. Simione, G. Liberati, M. Caruso, B. Calabrese, N. Ielpo, A. Palumbo, V. Frasca, M. Mecella, F. Amato, A. Pizzimenti, M. Inghilleri, D. Mattia and F. Cincotti. "The Brindisys project: Brain Computer Interfaces as assistive technology for people with ALS". BCI Meeting 2013 Fifth International Meeting Asilomar, California June 3-7, 2013 - DOI:10.3217/978-3-85125-260-6-186

[C 4] **F. Schettini**, F. Aloise, P. Aricò, S. Salinari, D. Mattia, F. Cincotti. "Self-Calibration in an asynchronous P300-based BCI". BCI Meeting 2013 Fifth International Meeting Asilomar, California June 3-7, 2013. DOI: 10.3217/978-3-85125-260-6-124

[C 5] P. Aricò, F. Aloise, **F. Schettini**, S. Salinari, D. Mattia, F. Cincotti. "Assessment of the P300 evoked potentials latency stability during c(o)vert attention BCI". BCI Meeting 2013 Fifth International Meeting Asilomar, California June 3-7, 2013.DOI:10.3217/978-3-85125-260-6-12

[C 6] M. Caruso, F. Cincotti, F. Leotta, M. Mecella, A. Riccio, **F. Schettini**, L. Simione, e T. Catarci, "My-World-in-My-Tablet: An Architecture for People with Physical Impairment", in Human-Computer Interaction. Interaction Modalities and Techniques, vol. 8007, M. Kurosu, A c. di Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pagg. 637–647..

[C 7] Riccio A., Simione L., **Schettini F**., Pizzimenti A., Inghilleri V., Olivetti M., Mattia D., Cincotti F. (2013) Attention and P300-based BCI performance in people with amyotrophic lateral sclerosis , Proceedings of TOBI Workshop IV, Practical Brain-Computer Interfaces for End -Users: Progress and Challenges; Sion, Switzerland, January 23-25, 2013

[C 8] P. Aricò, F. Aloise, **F. Schettini**, S. Salinari, D. Mattia, F. Cincotti. "Evaluation of the Latency Jitter of P300 Evoked Potentials during C(o)vert Attention BCI". 4rd Workshop of the TOBI Project: Practical Brain-Computer Interfaces for End-Users: Progress and Challenges. Sion, Switzerland, January 23-25, 2013.

[C 9] **F. Schettini**, A. Riccio, L. Simione, G. Liberati, M. Caruso, B. Calabrese, N. Ielpo, A. Palumbo, V. Frasca, M. Mecella, F. Amato, A. Pizzimenti, M. Inghilleri, and F. Cincotti. "The Brindisys Project: Brain-computer interface devices to support individual autonomy in locked-in individuals". 3° Convegno Arisla: Nuove Prospettive di ricerca per un futuro senza SLA. Milano, Italia, 24 ottobre 2012

[C 10] **F. Schettini,** F. Aloise, P. Arico, S. Salinari, D. Mattia, e F. Cincotti. "Control or no-control? Reducing the gap between brain-computer interface and classical input devices", Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Conf., vol. 2012, pagg. 1815–1818, 2012**.**

[C 11] **F. Schettini**, F. Aloise, M. Mecella, M. Caruso, A. Palumbo, A. Pizzimenti, M. Inghilleri and F. Cincotti. "From Keyboard to Brain Computer Interface: the Brindisys Project". GNB2012, June 26th-29th 2012, Rome, Italy. ISBN: 978 88 555 3182-5.

[C 12] **F. Schettini**, F. Aloise, P. Aricò, S. Salinari, D. Mattia, F. Cincotti. "Improving Communication Efficiency for gaze independent P300 based Brain Computer Interface". GNB2012, June 26th-29th 2012, Rome, Italy. ISBN: 978 88 555 3182-5.

[C 13] F. Aloise, S. Clausi, **F. Schettini**, L. Pizzamiglio, M. Molinari and M. G. Leggio. "An Holter for the mood states: the MoMo device". GNB2012, June 26th-29th 2012, Rome, Italy. ISBN: 978 88 555 3182-5.

[C 14] F. Aloise, P. Aricò, **F. Schettini**, M. Iosa, M. Scarnicchia, S. Salinari, D. Morelli, D. Mattia, and F. Cincotti. "The Brain Computer Interface as augmentative and alternative communication aid: the ABC project". GNB2012, June 26th-29th 2012, Rome, Italy. ISBN: 978 88 555 3182-5.

[C 15] P. Aricò, F. Aloise, **F. Schettini**, S. Salinari, D. Mattia, F. Cincotti. "On the correlation between Brain Computer Interface performance and Chronotype. GNB2012, June 26th-29th 2012, Rome, Italy. ISBN: 978 88 555 3182-5.

[C 16] A. Lechner, R. Ortner, F. Aloise, R. Prückl, **F. Schettini**, V. Putz, J. Scharinger, E. Opisso, U. Costa, J. Medina, C. Guger. "Comparing the Accuracy of a P300 Speller for People with Major Physical Disability". Computers Helping People with Special Needs Lecture Notes in Computer Science, 2012, Vol. 7383/2012, 180-183, DOI: 10.1007/978-3-642-31534-3_27.

[C 17] M. Iosa, F. Aloise, **F. Schettini**, P. Aricò, S. Paolucci, D. Morelli, D. Mattia , F. Cincotti. "Uso dei sistemi di Brain-Neural-Computer Interface nella comunicazione aumentata: il progetto ABC". XII National congress SIRN, 3rd-5th May 2012, Milan, Italy.

[C 18] **F. Schettini**, F. Aloise, P. Aricò, J. Taborri, S. Salinari, D. Mattia, F. Cincotti. "Asynchronous Control for a gaze independent P300 based BCI". 3rd Workshop of the TOBI Project: Bringing BCIs to End-Users: Facing the Challenge, Evaluation, User Perspectives, User Needs, and Ethical Questions. Würzburg, GermanyMar20th-22th, 2012.

[C 19] P. Aricò, F. Aloise, **F. Schettini**, V. Soragnese, S. Salinari, D. Mattia, F. Cincotti. "Variability of P300-based Brain Computer Interface performance across Repeated Sessions in a Day". 3rd Workshop of the TOBI Project: Bringing BCIs to End-Users: Facing the Challenge, Evaluation, User Perspectives, User Needs, and Ethical Questions. Würzburg, GermanyMar20th-22th, 2012.

[C 20] **F. Schettini**, F. Aloise, P. Aricò, S. Salinari, S. Petrichella, D. Mattia and F. Cincotti. "Comparing efficiency for Synchronous and Asynchronous P300-based BCIs". 5th International Brain-Computer Interface Conference 2011. September 22th-24th 2011 Graz, Austria

[C 21] F. Aloise, **F. Schettini**, P. Aricò, S. Salinari, C. Guger, J. Rinsma, M. Aiello, D. Mattia and F. Cincotti. "Validation of an asynchronous P300-based BCI with potential end users to control a virtual environment". 5th International Brain-Computer Interface Conference 2011. September 22th-24th2011 Graz, Austria

[C 22] P. Aricò, F. Aloise, **F. Schettini**, S. Salinari, S. Santostasi, D. Mattia and F. Cincotti. "On the effect of ERPs-based BCI practice on user's performances". 5th International Brain-Computer Interface Conference 2011. September 22th-24th2011 Graz, Austria

[C 23] F.Aloise, **F.Schettini**, P.Aricò, F.Leotta, S.Salinari, D.Mattia, F.Babiloni, F.Cincotti. "Towards Domotic appliances control through a self-paced P300-based BCI". International Conference on Bio-Inspired Systems and Signal Processing: BioSignal 2011. January 26th-29th, 2011; doi: 10.5220/0003162002390244

[C 24] F.Aloise, P. Aricò, **F. Schettini**, A. Riccio, M. Risetti, S. Salinari, D. Mattia, F. Babiloni, F. Cincotti. "A new P300 No Eye-gaze based interface: GeoSpell". International Conference on Bio-Inspired Systems and Signal Processing: BioSignal 2011. January 26-29, 2011; doi: 10.5220/0003161202270232

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Il dottorato di ricerca è un percorso lungo fatto di tante soddisfazioni, ma anche di momenti di incertezza e stress. Per questo voglio ringraziare, senza fare nomi, tutte le persone che mi sono state vicino, condividendo le mie gioie e le mie ansie, ed incoraggiandomi a dare sempre (quasi sempre) il meglio di me.

Grazie a chi ha avuto fiducia nelle mie capacità e mi ha permesso di intraprendere questo percorso, e ancora grazie a chi mi ha seguito e indirizzato durante tutti questi anni con pazienza e professionalità.

Grazie al mio inseparabile compagno di avventure… *a verità*… alla fine ce l'abbiamo fatta…

Grazie a chi è stata sempre lì a darmi supporto (psicologico e non)… siamo una coppia fantastica… la bionda e la bruna… i nostri nomi possono diventare uno solo…ti sei riconosciuta?

Grazie a chi è uscita dall'ascensore nel momento giusto (mon ami)… sarebbe stata veramente dura tutta sola al freddo al gelo senza di te… my friend!

Grazie a tutti quelli con cui ogni giorno condivido idee e dubbi (di natura scientifica), ma anche aneddoti divertenti, ricette, cibo, ecc…

Grazie all'amore della mia vita… per essermi stato sempre vicino… per essere quello che è… ed aver condiviso tutto con me…

Grazie alla mia famiglia… a chi mi supporta da lontano… a chi lo fa di presenza… a chi fa il tifo per me da molto lontano… per aver contribuito a rendermi quella che sono… vi sento tutti… sempre… grazie..

Grazie alle mie amiche… quelle preistoriche (tipo ex-compagne di banco del liceo, quelle con strani gradi di parentela – giuggine – quelle che mi vengono a trovare una volta al mese…)… e quelle storiche (tipo ex-colleghe, Commari con famiglia al seguito, ex-coinquiline)… con voi al vostro fianco diventa tutto più facile e niente è impossibile!