Removal of muscle artifacts from EEG recordings of spoken language production

M. De Vos[1], S. Riès[2,3], K. Vanderperren[1], B. Vanrumste[1], F.-X. Alario[2], S. Van Huffel[1], and B. Burle[3]

[1] Department of Electrical Engineering (ESAT), Katholieke Universiteit Leuven, Leuven, Belgium
[2] Laboratoire de Psychologie Cognitive, Aix-Marseille Universit, CNRS, Marseille, France
[3] Laboratoire de Neurobiologie de la Cognition, Aix-Marseille Universit, CNRS, Marseille, France

Abstract. Research on the neural basis of language processing has often avoided the issue of spoken language production by fear of the electromyographic (EMG) artifacts that articulation induces on electroencephalogram (EEG) signal. Indeed, such articulation artifacts are typically much larger than the brain signal of interest. Recently, a Blind Source Separation technique based on Canonical Correlation Analysis was proposed to separate tonic muscle artifacts from continuous EEG recordings in epilepsy. In this paper, we show how the same algorithm can be adapted to remove the short EMG bursts due to articulation on every trial. Several analyses indicate that this method accurately cleans the EEG recordings providing to the neurolinguistic community a powerful tool to investigate the brain processes at play during overt language production.

1 Introduction

Psychologists and neuroscientists have made extensive use of brain activity measures to construct models of language processing (Stemmer and Whitaker, 2008). In most of these studies, human participants are requested to understand or decode language (e.g. reading or listening to utterances). By contrast, the research on how the brain produces spoken language is considerably under-developed (for a review see Indefrey and Levelt (2004)). Such an underdevelopment is, partly, related to the artifacts induced by overt speaking on the signals measured by imagery techniques. For example, articulatory movements tend to reduce the signal to noise ratio in functional magnetic resonance imaging. Magneto- and Electroencephalography suffer from the contamination of the brain activity (evoked fields or potentials) by the Electromyographic (EMG) activity of the face muscles involved in overt speaking.

Previous research has shown how these artifacts may render the data uninterpretable. For example, McAdam and Whitaker (1971) reported larger slow potentials in an EEG experiment during both the production of polysyllabic

words, and in analogous non-speech articulatory movements (e.g. puffing). The critical finding was that these slow potentials were left-lateralized during word production but not in the control condition. This effect was interpreted as a reflection of linguistic processing tied to left inferior frontal cortical activity. However, shortly after publication, this interpretation was questioned based on methodological concerns (Morrell et al., 1971). Recordings of EMG activity of various articulators suggested that the reported lateralization effect may not have occurred during the preparation of the verbal response, but instead during articulation itself. Consequently, it may be a mere consequence of EMG activity. Brooker and Donald (1980) investigated this issue in close detail (see also references therein). These authors also compared speech to non-speech articulatory responses, while they measured both EEG scalp activity as well as the EMG of several articulators. Results showed strong correlations between the activity recorded by electrodes placed over the articulators and most of the one picked by the electrodes placed over the scalp (with the notable exception of the vertex scalp electrode). Further analyses were conducted in which the muscle–related activity was included as a co-variate factor. They did not reveal any lateralization of speech related cortical potentials. Besides questioning previous findings, the general conclusion was that pre-vocalization potentials are severely confounded by muscle artifacts.
In light of these difficulties, scholars investigating speech production have resorted to a variety of strategies. The most radical one has been to investigate cognitive stages of speech production without speech being overtly produced during experimental trials (speech was produced during interleaved filler trials). In an influential article (van Turennout et al., 1998), participants were asked to press buttons in response to visually presented pictures; critically, the responses were guided by linguistic properties of the pictures' names. For example, participants pressed one of two buttons depending on the first phoneme of the picture's name. In this way, participants had to process linguistic information without having to overtly speak. A somewhat similar strategy was adopted by Jescheniak et al. (2002). In their experimental paradigm, participants named pictures while their EEG activity was recorded. However, they did not name the pictures directly after their presentation but when a visual cue was presented to them. In the critical experimental trials, participants heard a distractor word just after the presentation of the picture. A semantic manipulation was performed on this word. The EEG activity of interest was recorded between this distractor word and the visual cue. In this way, the data were free of EMG activity. The downside is that they presumably reflected a combination of brain activities elicited by speaking and listening.

These experimental protocols have provided information about the neural basis of linguistic processing. However, what is being investigated in these studies may not reflect processes of natural speech production. This is because the complex instructions that participants are requested to follow presumably induce considerable non-speech brain activity (e.g. attention focused on decision making in button press or on withholding verbal responses). Another strategy has

been to elicit speech in a spontaneous manner (e.g. immediate picture naming) and thus obtain heavily muscle-contaminated EEG recording. Scientists following this approach try to remove the muscle artifact with heavy low-pass filters. For example, Masaki et al. (2001) or Ganushchak and Schiller (2008) used 10 or 12 Hz low pass filters. This approach is not without problems, however. The frequency spectrum of muscle artifacts has been shown to largely overlap with that of brain signals of theoretical interest (Goncharova et al., 2003). Friedman and Thayer (1991) showed that EMG activity is present both in the alpha (8-13 Hz) and in the beta bands (13-20Hz). Goncharova et al. (2003) reported that facial EMG has a broad frequency distribution, from almost DC to more than 200Hz, even with weak muscle contraction (see especially their Figure 3). Moreover, heavy low pass filtering prevents investigations of EEG activity present above 10 Hz (e.g. beta band), and is known to significantly reduce the amplitude of phasic activities . Visually evoked potentials, for example, often last around 100 ms and will therefore be clearly suppressed after 10 Hz low-pass filtering (Luck, 2005).

The approach we followed was to use Blind Source Separation (BSS) techniques to separate EMG from EEG activity in speech production tasks and remove the sources considered as EMG. We will discuss why this method can be appropriate for this purpose. We will first present the general principle of the Blind Source Separation approach. We will briefly review the use of Independent Component Analysis (ICA), a commonly used BSS algorithm to remove artifacts. ~~and motivate its limitations for EMG removal.~~ We will then present ~~an other~~ the BSS algorithm based on the Canonical Correlation Analysis, that will be used in this study.

This paper only considers removal of EMG contamination on speech recordings. These contamination mainly originates from muscles that are involved in the speech production. The paper does not deal with ongoing EMG activity unrelated to articulation. E.g. paralysis studies reveal that sustained EMG activity is present in the absence of overt movement (Whitham et al., 2007). Also we do not consider EMG contamination due to ongoing mental tasks.

1.1 Blind Source Separation

Applied to EEG, BSS assumes that the electrical signals recorded on the electrodes ($\mathbf{D}$) result from a weighted sum ($\mathbf{M}$) of elementary physiological sources ($\mathbf{S}$), defining the basic linear statistical model

$$\mathbf{D} = \mathbf{M} \cdot \mathbf{S} \quad (+\mathbf{N}) \tag{1}$$

where the EEG $\mathbf{D} \in \mathbb{R}^{I \times T}$ is called the observation matrix, $\mathbf{S} \in \mathbb{R}^{J \times T}$ the source matrix and $\mathbf{N} \in \mathbb{R}^{I \times T}$ the additive noise. $\mathbf{N}$ will not be modelled explicitly. $\mathbf{M} \in \mathbb{R}^{I \times J}$ is the mixing matrix, providing the contribution of each source to each sensor. I is the number of electrodes, J is the number of sources, and T the number of samples in each observation (duration $\times$ sampling rate) .

The goal of BSS is to estimate the mixing matrix $\mathbf{M}$, and/or the source vector $\mathbf{S}$, given only the observation matrix $\mathbf{D}$ (Lathauwer et al., 2000; Hyvärinen,

J. Karhunen and E. Oja, 2001). In other words, BSS tries to estimate the underlying, generating sources from the observed mixture. This is carried out by introducing the de-mixing matrix $\mathbf{W} \in \mathbb{R}^{J \times I}$ such that the estimated sources

$$\tilde{\mathbf{S}} = \mathbf{W} \cdot \mathbf{D}, \tag{2}$$

approximate the unknown physiological source signals in $\mathbf{S}$, by a scaling factor. Ideally, $\mathbf{W}$ is the inverse of the unknown mixing matrix $\mathbf{M}$, up to scaling and permutation.
Different $\mathbf{W}$'s will provide different source estimations $\tilde{\mathbf{S}}$. In order to break the EEG down into more elementary sources in a unique way, a decomposition criterion has to be defined. The appropriateness of such assumptions will determine how well the estimated sources $\tilde{\mathbf{S}}$ approximate the real contributing sources $\mathbf{S}$.
In this study, the goal was to separate the brain activity, related to speech production and the muscle activity, used for pronouncing the words.
We first introduce a decomposition based on the assumption of statistical independence (§1.2) and then one based on a difference in autocorrelation (§1.3).

1.2 Independent Component Analysis (ICA)

The most commonly used assumption for estimating $\mathbf{W}$ is that the sources are mutually statistically independent, as well as independent from the noise components. This leads to the concept of Independent Component Analysis (ICA) (Comon, 1994), based on higher order statistics.

ICA has been proposed to clean ictal EEG from the EMG contamination (Weidong and Gotman, 2004; Urrestarazu et al., 2004). Although ICA might be sometimes able to remove parts of tonic, continuous EMG activity in epileptic context, it has been shown to be far from optimal (De Clercq et al., 2006). Recently in (M. Crespo-Garcia and Cantero, 2008), ICA was used to remove muscle artifacts from sleep recordings. It is hard to judge if ICA was appropriate for their goal, as the authors did not show cleaned EEG data. We will show that ICA is not able to remove EMG contamination from speech recordings.

In any case, the use of ICA might be problematic for speech-induced EMG. Indeed, cleaning EEG data with ICA relies on the idea that unwanted activities (EMG in our case) will be captured by a relative low number of components. However, as the muscular pattern induced by articulation depends on the word to articulate, the EMG sources contaminating the EEG will be largely word-dependent. With many different words, the number of components necessary to capture all the possible EMG sources will likely be high, and the risk of mixing brain activity with EMG becomes problematic in this context. Alternatively, one may compute ICA on a word by word basis, but another disadvantage of ICA is that it requires lots of samples in order to estimate accurately the maximally statistical independence. As every trial is very limited in time, there are not enough samples for ICA to decompose the EEG into brain and muscle activity.
We thus opt for another BSS technique already used with success that allows to denoise each individual EEG segment (i.e. obtained for each word), and thus to better adapt to each articulation pattern.

1.3 Blind Source Separation by Canonical Correlation Analysis (BSS-CCA)

Assuming that the EEG and EMG sources are, (a) mutually uncorrelated, and (b) individually correlated in time, a BSS method based on Canonical Correlation Analysis (BSS-CCA) can be defined. Canonical Correlation Analysis is a statistical method originally devised to measure the linear relationship between two multidimensional variables $\mathbf{A}$ and $\mathbf{B}$ (Hotelling, 1936). That is, it estimates the similarity between the two datasets. As such, BSS-CCA has been used in different applications, see e.g. (Hardoon et al., 2004; De Vos et al., 2007). A multidimensional dataset is represented in a predefined basis. However, such a basis can be changed, e.g. by rotation. CCA rotates the two data sets independently, searching for two bases that are maximally correlated. If two datasets are well-described in similar bases, they can be considered as similar.
Besides its use in measuring linear relationship between variables, CCA can also be used to solve the BSS problem and recover source signals. This is how the method will be used for muscle artifact removal. As in equation 4, let us define $\mathbf{S}$ as the matrix of source signals and $\mathbf{D}$ as the observed mixture. From $\mathbf{D}$, we construct two new matrices $\mathbf{A}$ and $\mathbf{B}$, defined as $\mathbf{A} = \mathbf{D}(:, \mathbf{1} : (\mathbf{T} - \mathbf{1}))$ and $\mathbf{B} = \mathbf{D}(:, \mathbf{2} : \mathbf{T})$. Note that we use matlab notation (:) to define that all channels are involved. $\mathbf{A}$ is thus the observation matrix and $\mathbf{B}$ is the observation matrix shifted by one time sample. CCA will look for the linear combination of the mixed signals $\mathbf{D}(:, \mathbf{1} : (\mathbf{T} - \mathbf{1}))$ that correlates best with the linear combination of the mixed signals shifted by one sample ($\mathbf{D}(:, \mathbf{2} : \mathbf{T})$) in order to extract the source signals $\mathbf{S}$.
Mathematically, CCA is computed by solving matrix equations, see e.g. (Borga and Knutsson, 2001; Friman et al., 2001; Golub and Van Loan, 1996) and appendix A. In this paper, we focus only on how the source signals are extracted. Looking for a similar basis between a dataset ($\mathbf{A}$) and its shifted variant ($\mathbf{B}$) means explaining variance common to both. All sources are simultaneously estimated, but we explain the extraction of sources sequentially. The first extracted source will be a linear combination of the mixed signals with maximal (auto)correlation (using time lag 1, as defined above), thus explaining most of the variance between $\mathbf{A}$ and $\mathbf{B}$. This first extracted source, also called canonical correlation component, defines a first basis vector. The regression weights defining the linear combination, will provide the first row of the demixing matrix $\mathbf{W}$. The data is then projected away from this basis vector (this implies the mutual orthogonality between different sources), and a similar procedure is used to find the second source signal. Again the second source signal will be a linear combination with maximal autocorrelation, this time under the constraint that it is uncorrelated with the first component. The coefficients defining this linear combination will define the second row of $\mathbf{W}$. This procedure of finding autocorrelated source signals is repeated until the data is fully decomposed. In contrast to ICA, where the source signals do not have a fixed order, BSS-CCA decomposes the observation matrix $\mathbf{D}$ into sources that are sorted in decreasing

order of their autocorrelation (highly autocorrelated sources ranked first, weakly autocorrelated ones ranked last).

1.4 Application of BSS to muscle removal

EMG activity has a broad spectrum and tends to have white noise properties (Goncharova et al., 2003). It is thus weakly autocorrelated over time. In contrast, brain activity is considerably coherent over time, and will be more autocorrelated. In (De Clercq et al., 2006), a BSS-CCA method was introduced for decomposing the EEG, as it decomposes the observation matrix $\mathbf{D}$ into sources that are sorted in decreasing order of their autocorrelation (highly autocorrelated sources ranked first, weakly autocorrelated ones ranked last). After BSS-CCA, the sources with the highest autocorrelation should correspond to EEG, while the sources with the lowest autocorrelation should correspond to EMG. With this method, the authors could reveal the ictal activity otherwise completely masked in the background tonic EMG activity. The method has been further validated in detail on continuous EEG data from 37 patients with refractory partial epilepsy and is now used in clinical practice.
In speech production research, EMG contamination from the muscles involved in articulating words is not a tonic, continuous activity, but instead appears as short EMG bursts ~~produced by articulation~~, that are localised in time around the window of interest. A second difference with De Clercq et al. (2006) is that in epilepsy, a semi-automatic method was used. Neurologists visually removed muscle components one by one until ictal activity became visible. In the current application, we select the EMG sources automatically based on spectral properties (see below).

The goal of this study was to investigate whether the BSS-CCA method could be used in practice to distinguish cortical and EMG signals in electrophysiological recordings made during spoken language production. We did so on a dataset recorded using the picture naming task (a popular method for eliciting speech in psycho- and neuro-linguistic experiments).

2 Materials and Methods

This experiment was conducted for other (psycholinguistic) purposes that will be described elsewhere. Only the aspects relevant for our current purposes will be detailed here.

2.1 Participants and task

12 right-handed native french-speakers (3 females) with normal or corrected to normal vision, participated in the experiment (mean age: 24.5). They all gave their informed consent. Participants were presented with line drawings pictures representing objects that they had to name as fast and accurately as possible (Alario and Ferrand, 1999). Forty-five line drawings (11 $\times$11 cm) of common

objects were used as stimuli. Each drawing was randomly repeated twice during the experiment. Vocal responses were recorded with a voice-key (Eprime 2.0 Professional, Pittsburgh, PA: Psychology Software Tools).

2.2 Electrophysiological recordings

The EEG was recorded from 64 Ag/AgCl pre-amplified electrodes (BIOSEMI, Amsterdam) (10-20 system positions). The sampling rate was 512 Hz (anti-aliasing filters: DC to 104 Hz, 3 db/octave). The passive reference was obtained by offline averaging the signal recorded over the left and right mastoids. The vertical electro-oculogram (EOG) was recorded by means of two electrodes (same type as EEG) just above and below the left eye, respectively, and the horizontal EOG was recorded with two electrodes positioned over the two outer canthi.

2.3 Electrophysiological processing

After acquisition, the EEG data was filtered (high pass = 0.16 Hz). Eye movement artifacts were corrected using the statistical method of Gratton et al. (1983). The continuous EEG was epoched off-line, time-locked to stimulus presentation, starting from -0.2 s up to 2 s after stimulus onset. BSS-CCA decomposition was computed on each epoch separately, obtaining a $\mathbf{W}$ matrix and source signals for each trial (of 2,2 seconds). This was motivated by the fact that different muscular patterns are associated with different words used in this experiment (Marchal, 2007), and hence that the decomposition should be word-dependent. The sources obtained with BSS-CCA are ordered according to their autocorrelation. Because the autocorrelation of a source is an abstract value, another criterion has to be used to select the sources that are considered to be EMG. This criterion can be based on the Power Spectral Density (PSD), because EMG sources can be assumed to also have more high-frequency content than brain sources. ~~The EMG related components were automatically selected according to their Power Spectral Density (PSD).~~ Components were considered to be EMG activity if their average power in the EMG band (approximated by 15-30 Hz) was at least one seventh of the average power in the EEG band (approximated by 0-15 Hz). These values were empirically determined based on visual interpretation of the recordings. We will show that the chosen values (15 and 7) are not critical at all for the proposed method. In fact, the used criterion ~~This procedure~~ removed the least autocorrelated sources. ~~, which confirms the appropriateness of CCA decomposition.~~ This confirms that decomposing EEG for separating brain and muscles sources with BSS-CCA is appropriate. The contributions of the source signals identified as EMG were removed from the surface EEG by setting to zero the corresponding columns in the $\mathbf{M}$ matrix. This $\mathbf{M}$ is estimated as the inverse of $\mathbf{W}$. The new mixing matrix $\mathbf{M_{clean}}$ is then used to reconstruct the denoised EEG signal matrix $\mathbf{D_{clean}}$:

$$\mathbf{D_{clean}} = \mathbf{M_{clean}} \cdot \mathbf{S}. \tag{3}$$

$\mathbf{D_{clean}}$ ~~will only contain contributions from brain activity~~ will contain sources not identified as muscle activity. After removing muscle artifacts, all other artifacts (e.g. related to eye movements, recording problems, persistent muscle activity) were rejected after a trial-by-trial visual inspection of monopolar recordings and the clean EEG segments were re-epoched, centered on the verbal response onset.

For comparison, instead of removing EMG sources with BSS-CCA, we applied a low-pass filter (10 Hz, 24db/oct), and we also applied an ICA algorithm based on higher order statistics , namely RobustICA (Zarzoso and Comon, 2008).

2.4 Statistical evaluation.

In order to quantify the removal of EMG in a more objective way, we computed the amplitude differences for all the subjects between the first negative peak and the following positive peak on one electrode (Fp2). We computed this for 3 datasets: the raw data, the data cleaned with the low-pass filter, and the data cleaned with BSS-CCA. We compared the obtained values with a standard t-test.

3 Results

~~**Behavioural** results~~

The average reaction time (RT) across subjects was 651 $\pm$ 72 ms and average error rate was 1,31% $\pm$ 0,96 %.
We will illustrate the effect of BSS-CCA on one representative subject with considerable EMG contamination. Although we only show results for this subject, very similar figures can be obtained with the other subjects.

3.1 Impact of BSS-CCA on speech-related EEG activity denoising

De Clercq et al. (2006) have already demonstrated the validity of BSS-CCA to remove EMG artifacts by means of simulations. Evaluating the validity of the method on real data is a much more difficult matter, since, by definition, the unaffected signal is unknown. We will ~~this~~ first present several analyses comparing directly the raw with the cleaned data.

Figure 1 presents an example of five central channels (FPz, AFz, Fz, FCz and Cz). The top panel presents the raw data, the middle one plots the identified EMG components, and the bottom panel presents the EEG data after EMG removal. It is to be noted that the removed components clearly represent EMG burst around response time, and no low frequency modulation. In contrast, the cleaned traces present slow modulations, typical of EEG activity with, apparently, no visible EMG activity.

The previous figure shows the cleaning on only one epoch (but several electrodes). We now visualize the impact of denoising on all the trials, for one electrode: the ERP-image technique (Jung et al., 2001). All the trials are represented as parallel color lines, where color codes are used for the polarity and the strength of the activity. The trials are sorted as a function of RT, represented as the black S-Shaped line (trials for shortest RT are at the bottom, trials with the longest RT are at the top). Figure 2 presents the ERP-Images of the original epochs, and of the epochs after cleaning with BSS-CCA for channels Fp2 and T7. We clearly see a reduction of the high frequency noise (visible as a pixelisation of the plot), especially around the response time.

Since EMG contamination is mainly located in high frequencies (see previous figures), we can evaluate the impact of BSS-CCA cleaning more directly on the power spectrum of the recorded activities. Figure 3 shows the averaged spectral content of all epochs on selected channels represented topographically in the original data (black), after EMG removal with BSS-CCA (blue) and after low-pass filtering (red). While the unfiltered EEG contains a lot of high frequency content, the low-pass filter and the BSS-CCA method dramatically reduces high frequency and reveal the expected $1/f^{\alpha}$ shape in the spectrum. Note that the actual α differs between the two methods. It can not be proven from this figure which spectrum is the best one. However, the spectrum after BSS-CCA is closer to the original one than the spectrum obtained after filtering. We can also see that, prior to applying BSS-CCA, there is more high frequency content on the frontal and temporal compared to the central, parietal and occipital recording sites. After muscle removal~~BSS-CCA~~, however, an $1/f^{\alpha}$ shape of the spectrum is visible on all electrodes.

In order to quantify this improvement, we compared the spectrum plotted on a log-log scale with a straight
Table 2 shows the correlation between a straight line and the log-log transformed
spectra before and after cleaning for some electrodes. For all the electrodes, the power spectrum is closer to a

To compare BSS-CCA and filtering in more detail, we plot the averaged ERP on two channels (Fp2 and P3) in figure 4. It can be seen that the ERP cleaned with BSS-CCA resembles much more the ERP of the raw data then after filtering. Furthermore, the amplitude difference between the first negativity and the following positivity is clearly larger after BSS-CCA than after filtering.

In the next figure (Fig. 5), we also compare the removal with BSS-CCA and the removal with ICA based on higher order statistics. ICA based on higher order statistics was previously proposed to clean EEG from muscle artifacts. The main disadvantage that was mentioned for using ICA is the selection of sources related to EMG. To overcome this problem, we used exactly the same criterion as we proposed for BSS-CCA. The figure clearly illustrates that the problem is not the selection of EMG components, but the decomposition itself. Cleaning with ICA removes less EMG and follows less the raw ERP.

In figure 6, we show that the actual threshold for selecting components is not relevant. We compared 3 different parameter settings on the averaged ERP (border frequency 15, and removing an EMG source when there is more than 1/7 of the power above this frequency; 16 and 1/10; 13 and 1/5). The figure shows that the actual threshold does not influence the final result.

Finally, we evaluated where the removed components were stemming from by plotting the topography of the removed muscle components (Figure 7). To do so, one needs to provide a representative topography. Plotting the topography of an individual removed component is not appropriate since it might not be representative. Instead, we normalise the estimated source signals to a variance of 1, and weighted the mixing vectors by the appropriate variance. The average of all mixing vectors is then a representative estimate of the topography of the EMG sources. Indeed, consider again the estimation

$$\hat{\mathbf{S}} = \mathbf{W} \cdot \mathbf{D} \tag{4}$$

for the decomposition

$$\mathbf{D} = \mathbf{M} \cdot \mathbf{S}. \tag{5}$$

After removal of the EMG sources, we can construct 2 datasets: $\mathbf{D_{clean}}$, containing the cleaned EEG and $\mathbf{D_{EMG}}$ containing muscle artifacts:

$$\mathbf{D_{clean}} = \mathbf{M_{clean}} \cdot \mathbf{S} \text{ and } \mathbf{D_{EMG}} = \mathbf{M_{EMG}} \cdot \mathbf{S} \tag{6}$$

If the sources are ordered by decreasing auto-correlation, $\mathbf{M_{clean}}$ and $\mathbf{M_{EMG}}$ are defined as:

$$\mathbf{M_{clean}} = \mathbf{M} \cdot \begin{pmatrix} 1 & 0 & 0 & \dots & \\ 0 & 1 & 0 & \dots & \\ \vdots & \vdots & \ddots & & \\ & & & 0 & \\ & & & & \ddots \end{pmatrix} \text{ and } \mathbf{M_{EMG}} = \mathbf{M} \cdot \begin{pmatrix} 0 & 0 & \dots & & \\ \vdots & \ddots & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & \ddots \end{pmatrix}. \tag{7}$$

This can be seen, because only the columns from the first sources will be retained in $\mathbf{M_{clean}}$ and only the last columns will be retained in $\mathbf{M_{EMG}}$. The coefficients in $\mathbf{M_{EMG}}$ give an estimate of how much each EMG source loads on the observation signal. These coefficients are constant within trial and they do not reflect the exact EMG contribution at any time instant, as the time courses of the components are ignored while they can have all possible values. A global representation of EMG topography can be obtained nonetheless by averaging all the EMG mixing vectors because an actual mixing topography will be given by a weighted sum of these mixing vectors. A similar reasoning can be applied to the effect of filtering via Fourier transform, and it was already derived in the literature that singular value decomposition (SVD) can be interpreted as a filter (Hansen and Jensen, 1998). We average this estimate over trials and plot it on a 3D head. Most activation is frontal as expected, as many muscles involved in

speech production are facial. Also some EMG contributions of the neck muscles are observed.

In order to validate the applicability of the BSS-CCA method on all the subjects, we compared the amplitude differences between the raw data, the data cleaned with BSS-CCA and the data cleaned with filtering. The averaged peak-ro-peak values are shown in table 1. The amplitude difference was significantly greater after BSS-CCA than after heavy low-pass filtering ($t(11) = 5{,}54$, p ¡ 0,001). There was no significant difference in amplitude between the amplitude difference measured on the raw data compared to after BSS-CCA ($t(11) = 1{,}51$, $p = 0{,}10$). There was a significant difference in amplitude between the peak to peak values measured on the raw data compared to after ICA ($t(11) = 4.44$, p ¡.05).

4 Discussion

Psycholinguists have long avoided electrophysiological (EEG and/or MEG) investigations of spoken language because of the, justified, fear of the artifacts induced by facial EMG. Here we propose a solution to this problem based on a BSS technique that exploits the difference in autocorrelation between brain and muscle signals in order to separate them. This method was originally developed, and validated, for long term epilepsy recordings. In the present study, we adapted and automated it to remove short bursts of myographic activity related to speech production. Validating artifact removal techniques is a challenging task since neither the the EMG artifact nor the EEG signal related to speech production are well known. We thus investigated in detail the effects of the proposed method based on CCA in various ways.

BSS-CCA successfully reduced the EMG artifacts in the EEG signal recorded during speech. The efficiency of BSS-CCA is clearly visible on a trial by trial basis (figures 1 to 2) but also on the power spectra of the grand averages (figure 3). Before muscle artifact removal, the power spectra showed high frequency activity. After cleaning, the power spectra had a clear $1/f^{\alpha}$ shape on all electrodes. This is exactly the shape of the EEG power spectrum in Goncharova et al. (2003) in the condition when subjects were asked not to contract facial muscles (see their figure 3). ~~So we can be quite confident that we are dealing with the power spectrum of clean EEG signal. It is important to point out that heavy filtering such as used in recent EEG studies of overt speech severely alters this $1/f^{\alpha}$ shape of the power spectrum as all frequencies above 10 or 12 Hz are removed.~~ This shape model is not a very strong one, and can also be obtained with filtering, although the actual α value in filtering will be higher, and the spectra after filtering will less follow the original spectrum.

Based on our results, we can certainly claim that the removal of EMG contamination with BSS-CCA outperformed the removal with filtering (figure 4) and with ICA (figure 5). That filtering at 10 Hz removes also brain activity is well-known, and again confirmed by the figure.

It cannot said in advance that ICA might or might not remove speech-induced EMG. Our simulations however show that BSS-CCA preserves better the shape of the ERP while removing more disturbing signal. Another advantage of BSS-CCA is that it requires less samples than ICA based on higher order statistics to reliably estimate sources.

In order to further validate the removal, we studied the part of the signal that was removed by the algorithm. The topography (figure 7) of the rejected data indicates that most of the muscle activity was frontal. Also in the raw data, we observed that the impact of facial muscle contraction was related to the location of the recording electrode. These topographic differences are confirmed by the physical position of the articulation muscles, i.e. facial, and thus reflected on frontal electrodes. The locations where most activity was found on these topographies also corresponds to where the power spectra of the raw data contained most high frequencies.

The last validation concerned the peak-to-peak values. Based on the statistical test performed on all the subjects, no significantly difference could be seen between the values on raw data and the values after BSS-CCA.

Altogether, although the validation remaines a difficult problem, our results incline us to believe we successfully removed the most prominent EMG artifacts present in the EEG signal without significantly altering actual brain signal. However, although the global picture is convincing, the question remains whether a removed source contained purely EMG and thus whether or not BSS-CCA removes some brain activity along with the EMG artifacts in EEG. We have started to observe event-related potentials elicited by speech production that had not been described previously, see also (Ries et al., 2009). The future may further clarify possible limitations in the next experiments we will perform and if other neuroscientific research groups also use this algorithm (freely available at www.neurology-kuleuven.be/index.php?id=210) on their event-related potentials and starting thorough speech research. Covering activated muscles with more electrodes can maybe further improve the separation.

~~We believe that t~~ The BSS-CCA algorithm clearly outperforms the existing filtering method used to overcome the problem of EMG artifacts elicited by articulation in the EEG signal without avoiding direct overt speech in the experimental tasks. This method therefore enables neuroscientific investigations of spoken language production beyond current practice. ~~We have indeed already started to observe event-related potentials elicited by speech production that had not been described previously.~~

The results of this method in language processing experiments are also promising and open perspectives towards new applications, both with continuous and event-related EEG. In particular, addressing the problem of EMG in neuromuscular paralysis or EMG during ongoing mental tasks might be an interesting topic for future research.

Appendix A: CCA

Ordinary correlation analysis quantifies the relation between (realizations of) two variables $\mathbf{a}(t)$ and $\mathbf{b}(t)$ by means of a correlation coefficient ρ:

$$\rho = \frac{Cov[\mathbf{a}, \mathbf{b}]}{\sqrt{V[\mathbf{a}]V[\mathbf{b}]}} \tag{8}$$

in which Cov and V indicate respectively the co - and variance. CCA is a multivariate extension of ordinary correlation analysis. Consider 2 multivariate zero-mean vectors
$\mathbf{A}=[\mathbf{a}_1(t), \ldots, \mathbf{a}_m(t)]^T$ and $\mathbf{B}=[\mathbf{b}_1(t), \ldots, \mathbf{b}_n(t)]^T$, $t = 1, .., N$. Two new scalar variables, $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}}$, are generated as a linear combination of the components in $\mathbf{A}$ and $\mathbf{B}$:

$$\begin{aligned} \tilde{\mathbf{a}} &= w_{\mathbf{a}_1}\mathbf{a}_1 + \ldots + w_{\mathbf{a}_m}\mathbf{a}_m = \mathbf{w}_\mathbf{a}^T\mathbf{A} \\ \tilde{\mathbf{b}} &= w_{\mathbf{b}_1}\mathbf{b}_1 + \ldots + w_{\mathbf{b}_m}\mathbf{b}_m = \mathbf{w}_\mathbf{b}^T\mathbf{B} \end{aligned} \tag{9}$$

CCA computes the coefficients $\mathbf{w_a}$ and $\mathbf{w_b}$ that maximize the correlation between $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}}$. These coefficients are the regression weights and $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}}$ are denoted as canonical variates. The resulting correlation coefficient is the canonical correlation coefficient.
It can be shown that finding these regression weights correspond to solving an eigen value problem.
By inserting equation (9) into the definition of the correlation coefficient (8), and assuming the means of $\mathbf{A}$ and $\mathbf{B}$ zero, we obtain:

$$\rho = \frac{\mathbf{w}_\mathbf{a}^T\mathbf{C_{ab}}\mathbf{w_b}}{\sqrt{(\mathbf{w}_\mathbf{a}^T\mathbf{C_{aa}}\mathbf{w_a})(\mathbf{w}_\mathbf{b}^T\mathbf{C_{bb}}\mathbf{w_b})}} \tag{10}$$

with $\mathbf{C_{aa}}$ and $\mathbf{C_{bb}}$ the variance matrices from respectively $\mathbf{A}$ and $\mathbf{B}$ and $\mathbf{C_{ab}}$ the covariance matrix from $\mathbf{A}$ and $\mathbf{B}$. ρ is a function of $\mathbf{w_a}$ and $\mathbf{w_b}$. In order to maximise the correlation coefficients , we impose the partial derivatives with respect to $\mathbf{w_a}$ and $\mathbf{w_b}$ to be zero. This results in following system:

$$\begin{aligned} \mathbf{C}_\mathbf{aa}^{-1}\mathbf{C_{ab}}\mathbf{C}_\mathbf{bb}^{-1}\mathbf{C_{ba}}\mathbf{w}_{\mathbf{a}_i} &= \rho^2\mathbf{w}_{\mathbf{a}_i} \\ \mathbf{C}_\mathbf{bb}^{-1}\mathbf{C_{ba}}\mathbf{C}_\mathbf{aa}^{-1}\mathbf{C_{ab}}\mathbf{w}_{\mathbf{b}_i} &= \rho^2\mathbf{w}_{b_i} \end{aligned} \tag{11}$$

This system is an eigenvalue decomposition. The matrices $\mathbf{C}_\mathbf{aa}^{-1}\mathbf{C_{ab}}\mathbf{C}_\mathbf{bb}^{-1}\mathbf{C_{ba}}$ and $\mathbf{C}_\mathbf{bb}^{-1}\mathbf{C_{ba}}\mathbf{C}_\mathbf{aa}^{-1}\mathbf{C_{ab}}$ have the same eigenvalues. The vectors $\mathbf{w_a}$ and $\mathbf{w_b}$ we are looking for, are the eigenvectors corresponding to the highest eigenvalue. This eigenvalue is the square of the maximal correlation between the canonical variates.

Acknowledgements

This research is funded by a PhD grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen)and a doctoral grant for the french ministry of research; Research supported by ANR-07-JCJC-0074; Research Council KUL: GOA-AMBioRICS, CoE EF/05/006 Optimization in Engineering (OPTEC), IDO 05/010 EEG-fMRI, IOF-KP06/11 FunCopt, several PhD/postdoc & fellow grants; Flemish Government: FWO: PhD/postdoc grants, projects, G.0407.02 (support vector machines), G.0360.05 (EEG, Epileptic), G.0519.06 (Noninvasive brain oxygenation), FWO-G.0321.06 (Tensors/Spectral Analysis), G.0302.07 (SVM), G.0341.07 (Data fusion), research communities (ICCoS, ANMMM); IWT: TBM070713-Accelero, TBM-IOTA3; Belgian Federal Science Policy Office IUAP P6/04 (DYSCO, 'Dynamical systems, control and optimization', 2007-2011); EU: BIOPATTERN (FP6-2002-IST 508803), ETUMOUR (FP6-2002-LIFESCIHEALTH 503094), Healthagents (IST200427214), FAST (FP6-MC-RTN-035801), Neuromath (COST-BM0601) ESA: Cardiovascular Control (Prodex-8 C90242)

Table 1. The average amplitude difference between the first negativity and the following positivity between the 12 subjects. It can be seen that filtering largely reduces this amplitude difference.

Raw data	BSS-CCA	ICA	filtering
8,36 ($\sigma = 3,25$)	7,96 ($\sigma = 3,05$)	7,67 ($\sigma = 2,70$)	6,11 ($\sigma = 2,54$)

Table 2. ~~The correlation between the log-log transformed spectrum and a straight line before and after EMG removal. Increased correlation means better $1/f^{\alpha}$ shape of spectrum.~~

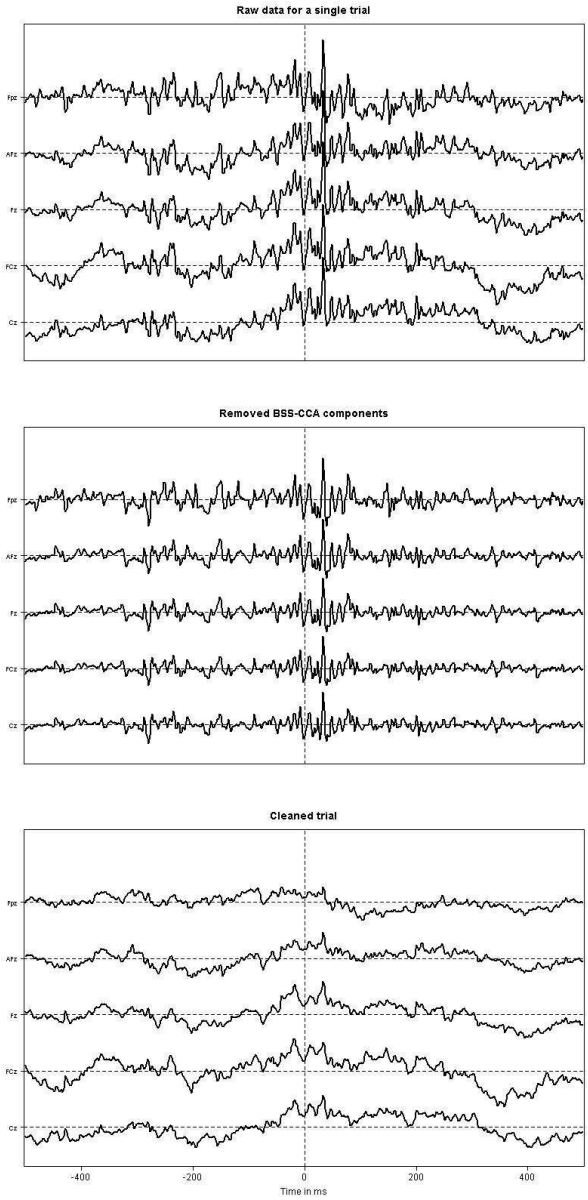


Fig. 1. Single trial of EEG data on 5 channels around voice onset (0 ms). We show the original evoked potential, the components that BSS-CCA removed and the evoked potential after cleaning with BSS-CCA. The removed components correspond to high frequency activity, and the cleaned trial still contains the low-frequency fluctuation.

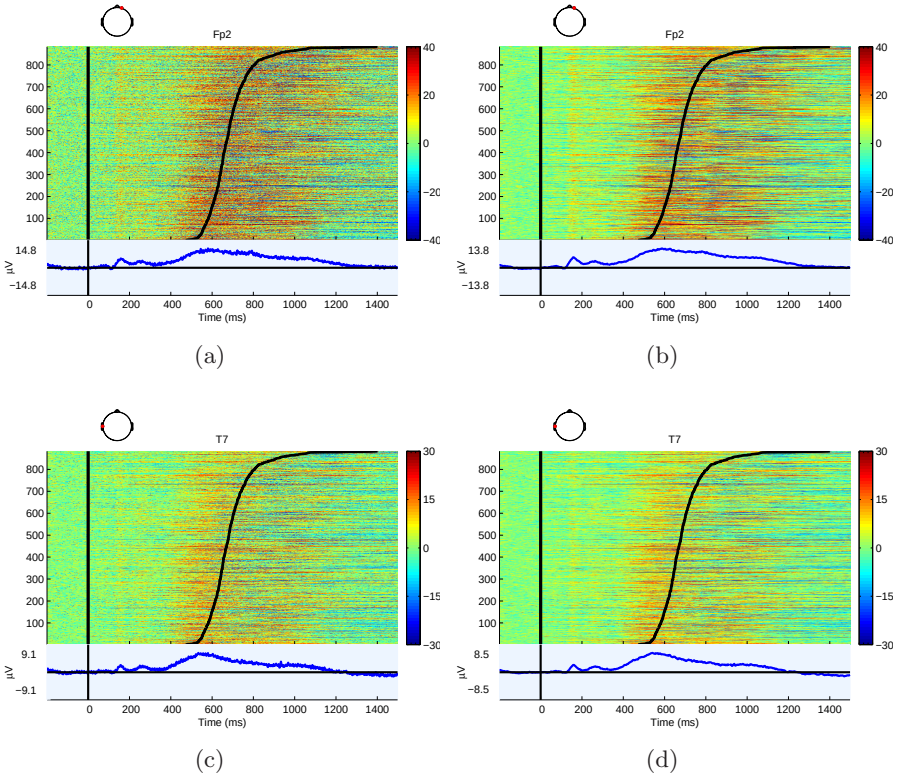


Fig. 2. Raster plot of the single trials at channel Fp2 and T7 before (a,c) and after (b,d) muscle removal, ordered according to reaction time. The black line represents reaction time and the blue line the ERP. After BSS-CCA, less high-frequency activity, supposed to be related to EMG contamination, is visible.

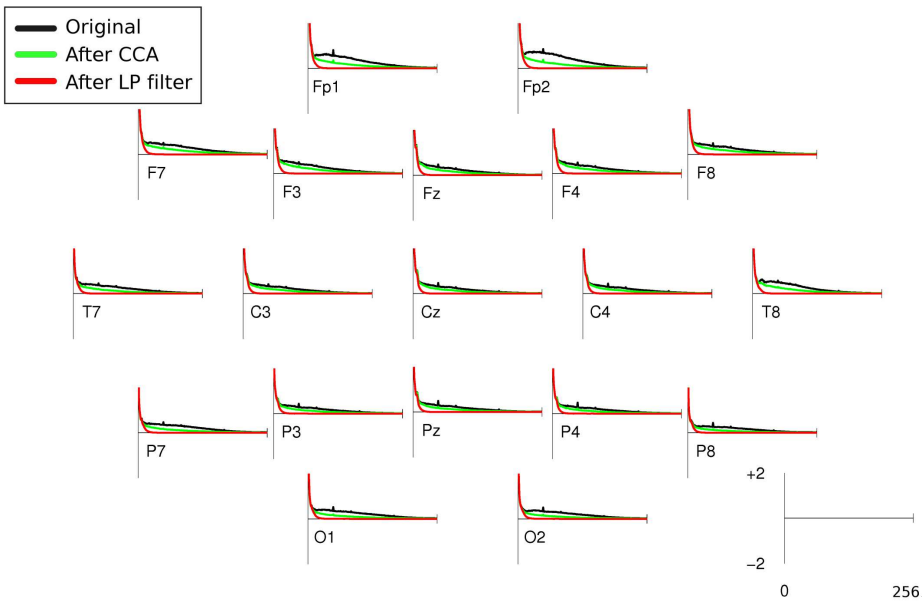


Fig. 3. Averaged Fourier transform over the trials of selected channels, equally distributed over the head. Black is the raw spectrum, blue represents the spectrum after BSS-CCA and the spectra after filtering is shown in red. The blue and red spectra have a $\frac{1}{f^{\alpha}}$ shape. However, the blue spectrum follows the original spectrum more closely.

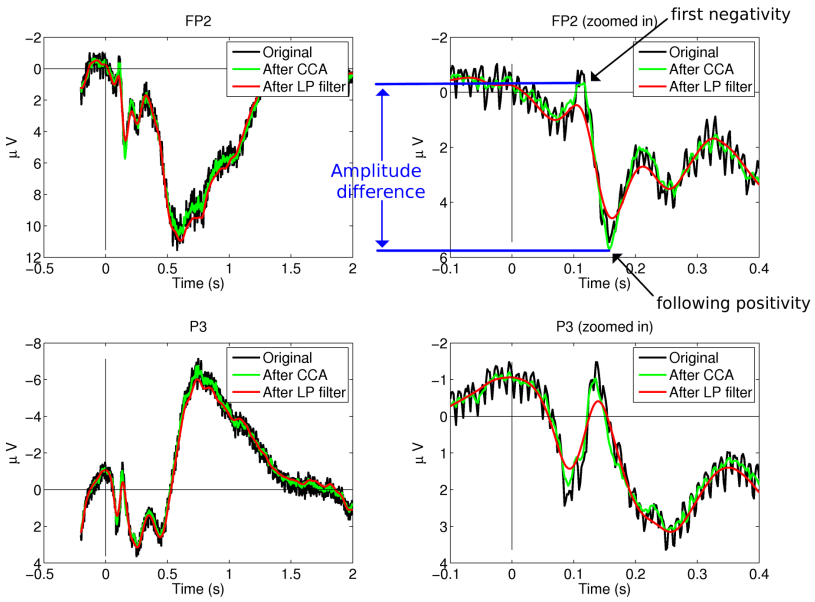


Fig. 4. Averaged ERP on two channels. We can see that BSS-CCA follows much more closely the ERP from the raw data than filtering.

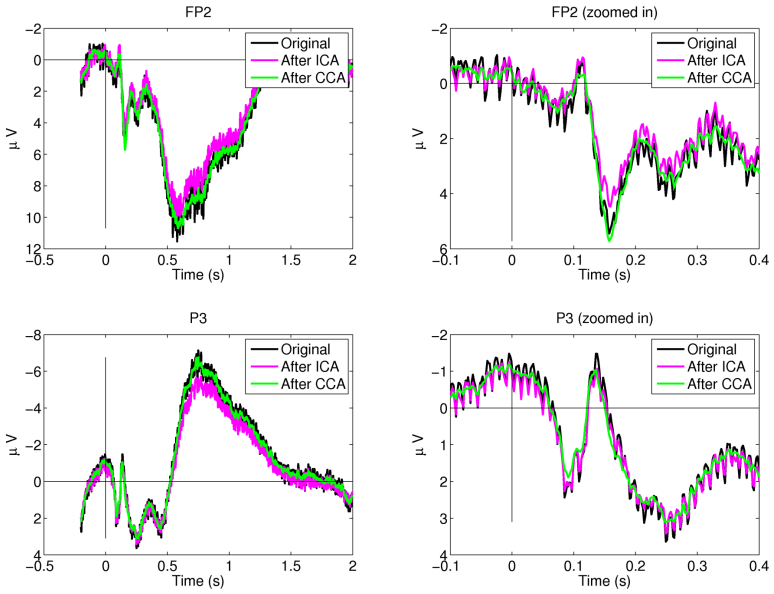


Fig. 5. Averaged ERP on two channels. We can see that BSS-CCA outperforms ICA.

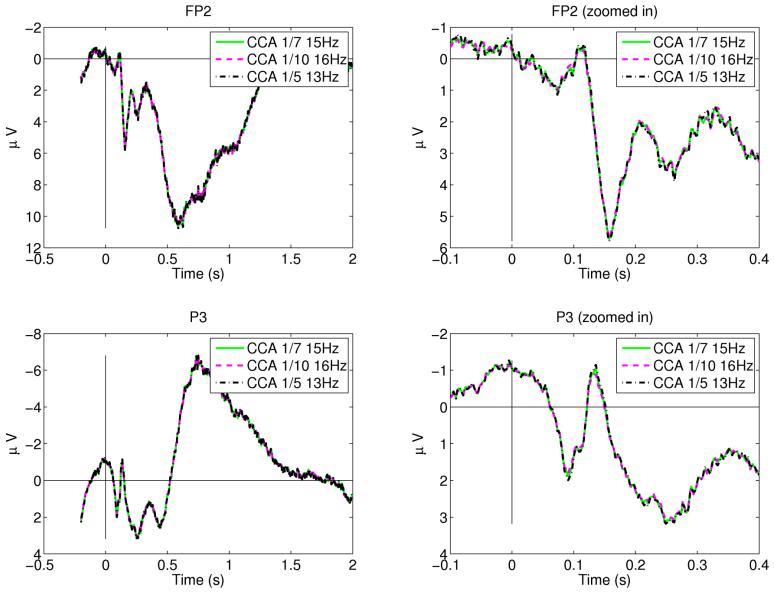
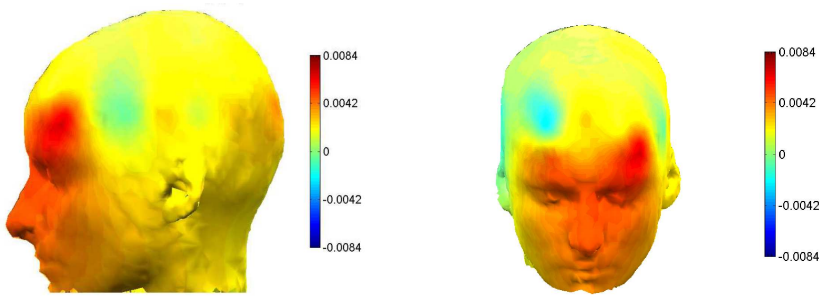


Fig. 6. Averaged ERP on two channels. We can see that the parameters for source selection do not influence the final ERP much.



(a) (b)

Fig. 7. Averaged topography of the removed muscle components. Mainly frontal activity, corresponding to cheek muscles, can be observed. Note that the colors outside the EEG electrodes are not very reliable, due to extrapolation.

Bibliography

Alario, F.-X., Ferrand, L., 1999. A set of 400 pictures standardized for french: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. Behavior Research Methods, Instruments & Computers 31 (3), 531 – 552.

Borga, M., Knutsson, H., 2001. A canonical correlation approach to blind source separation. Tech. Rep. LiU-IMT-EX-0062, Dept. of Biomedical Engineering, Linköping University, Sweden.

Brooker, B. H., Donald, M. W., 1980. Contribution of the speech musculature to apparent human eeg asymmetries prior to vocalization. Brain and Language 9, 226 – 245.

Comon, P., 1994. Independent component analysis, a new concept? Signal Process 36, 287 – 314.

De Clercq, W., Vergult, A., Vanrumste, B., Van Paesschen, W., Van Huffel, S., 2006. Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. IEEE Transactions on Biomedical Engineering 53, 2583 – 2587.

De Vos, M., Laudadio, T., Simonetti, A., Heerschap, A., Van Huffel, S., 2007. Fast nosologic imaging of the brain. Journal of Magnetic Resonance 184, 292 – 301.

Friedman, B. H., Thayer, J. F., 1991. Facial muscle activity and eeg recordings: redundancy analysis. Electroencephalography and clinical Neurophysiology 79, 358 – 360.

Friman, O., Cedefamn, J., Lundberg, P., Borga, M., Knutsson, H., 2001. Detection of neural activity in functional MRI using canonical correlation analysis. Magnetic Resonance in Medicine 45, 323 – 330.

Ganushchak, L. Y., Schiller, N. O., 2008. Motivation and semantic context affect brain error-monitoring activity: An event-related brain potentials study. NeuroImage 39, 395 – 405.

Golub, G., Van Loan, C. F., 1996. Matrix computations, 3rd Edition. John Hopkins University Press, Baltimore.

Goncharova, I. I., McFarland, D. J., Vaughan, T. M., Wolpaw, J. R., 2003. Emg contamination of eeg: Spectral and topographical characteristics. Clinical Neurophysiology 114 (9), 1580 – 1593.

Gratton, G., Coles, M., Donchin, E., 1983. A new method for off-line removal of ocular artifact. Electroencephalography and Clinical Neurophysiology 55, 468 – 484.

Hansen, P. C., Jensen, S. H., 1998. Fir filter representations of reduced-rank noise reduction. IEEE Trans. Signal Proc. 46, 1737–1741.

Hardoon, D., Szedmak, S., Shawe-Taylor, J., 2004. Canonical correlation analysis: An overview with application to learning methods. Neural Computation 16, 2639 – 2664.

Hotelling, H., 1936. Relations between two sets of variates. Biometrika 28, 321 – 377.

Hyvärinen, J. Karhunen and E. Oja, A., 2001. Independent Component Analysis. John Wiley & Sons.

Indefrey, P., Levelt, W. J. M., 2004. The spatial and temporal signatures of word production components. Cognition 92 (1), 101 – 144.

Jescheniak, J. D., Schriefers, H., Garrett, M. F., Friederici, A. D., 2002. Exploring the activation of semantic and phonological codes during speech planning with event-related brain potentials. Journal of Cognitive Neuroscience 14 (6), 951 – 964.

Jung, T., S.Makeig, Westerfield, M., Townsend, J., Courchesne, E., Sejnowski, T., 2001. Analysis and visualization of single-trial event-related potentials. Human Brain Mapping 14, 166–85.

Lathauwer, L. D., Moor, B. D., Vandewalle, J., 2000. An introduction to independent component analysis. Journal of Chemometrics 14, 123–149.

Luck, S. L., 2005. An Introduction to the Event-Related Potential Technique. MIT Press.

M. Crespo-Garcia, M. A., Cantero, J., 2008. Muscle artifact removal from human sleep EEG by using independent component analysis. Journal of Biomedical Engineering 36, 467–475.

Marchal, A., 2007. La production de la parole (Speech production). Lavoisier, Hermes Science.

Masaki, H., Tanaka, H., Takasawa, N., Yamazaki, K., 2001. Error-related brain potentials elicited by vocal errors. Neuroreport: For Rapid Communication of Neuroscience Research 12 (9), 1851 – 1855.

McAdam, D. W., Whitaker, H. A., 1971. Language production: Electroencephalographic localization in the normal human brain. Science 172 (3982), 499 – 502.

Morrell, L. K., Huntington, D. A., McAdam, D. W., Whitaker, H. A., 1971. Electrocortical localization of language production. Science 174 (4016), 1359 – 1360.

Ries, S., Janssen, N., Dufau, S., Alario, F.-X., Burle, B., 2009. Electroencephalographic evidence for general purpose monitoring during speech production. submitted.

Stemmer, B., Whitaker, H., 2008. Handbook of the Neuroscience of Language. Academic Press.

Urrestarazu, E., Iriarte, J., Alegre, M., Valencia, M., Viteri, C., Artieda, J., 2004. Independent component analysis removing artifacts in ictal recordings. Epilepsia 45 (9), 1071 – 1078.

van Turennout, M., Hagoort, P., Brown, C. M., 1998. Brain activity during speaking: From syntax to phonology in 40 milliseconds. Science 280, 572 – 574.

Weidong, Z., Gotman, J., 2004. Removal of emg and ecg artifacts from eeg based on wavelet transform and ica. In: 26th Annual International Conference of the Engineering in Medicine and Biology Society. Vol. 1. pp. 392 – 395.

Whitham, E.M.and Pope, K., Fitzgibbon, S. L. T., Clark, C., Loveless, S., Broberg, M., A., W., DeLosAngeles, D., Lillie, P., Hardy, A., Fronsko, R., A., P., Willoughby, J., 2007. Scalp electrical recording during paralysis: quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. Clinical Neurophysiology 118, 1877–1888.

Zarzoso, V., Comon, P., 2008. Robust independent component analysis for blind source separation and extraction with application in electrocardiography. In: 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE EMBS '08). Vancouver, Canada, pp. 3344–3347.