

Removal of Muscle Artifacts from EEG Recordings of Spoken Language Production

De Maarten Vos · Stephanie Riès · Katrien Vanderperren · Bart Vanrumste · Francois-Xavier Alario · Van Sabine Huffel · Boris Burle

Published online: 18 May 2010
© Springer Science+Business Media, LLC 2010

Abstract Research on the neural basis of language processing has often avoided investigating spoken language production by fear of the electromyographic (EMG) artifacts that articulation induces on the electro-encephalogram (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 attenuates the muscle contamination on the EEG recordings, providing to the neurolinguistic community a powerful tool to investigate the brain processes at play during overt language production.

Keywords EEG · ERP · EMG · Artifact · BSS · Speech production

Introduction

Psychologists and neuroscientists have made extensive use of brain activity measures to construct models of language processing (e.g. 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 (e.g. Indefrey and Levelt 2004). Such an underdevelopment is, partly, related to the artifacts induced by overt speaking on the signals measured by imaging techniques. For example, articulatory movements tend to reduce the signal-to-noise ratio in functional magnetic resonance imaging. Magneto- and Electro-encephalography 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 large 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.

D. M. Vos (✉) · K. Vanderperren · B. Vanrumste · V. S. Huffel
Department of Electrical Engineering (ESAT),
Katholieke Universiteit Leuven, Leuven, Belgium
e-mail: maarten.devos@gmail.com

S. Riès · F.-X. Alario
Laboratoire de Psychologie Cognitive,
Aix-Marseille Université, CNRS, Marseille, France

S. Riès · B. Burle
Laboratoire de Neurobiologie de la Cognition,
Aix-Marseille Université, CNRS, Marseille, France

B. Vanrumste
MOBILAB, Katholieke Hogeschool Kempen,
Geel, Belgium

Shortly after publication, however, this interpretation was questioned by Morrell et al. (1971) on the basis of methodological concerns. 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) by comparing speech to non-speech articulatory responses; they measured both EEG scalp activity and the EMG of several articulators. The results showed strong correlations between the activity recorded by electrodes placed over the articulators and 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.

The general conclusion from this thread of research was that pre-vocalization potentials are severely confounded by muscle artifacts. In light of these difficulties, scholars investigating speech production have developed strategies of various kinds, which are discussed in the following subsections.

Avoiding Articulation Altogether

The most radical strategy has been to investigate cognitive stages of speech production without speech being overtly produced during experimental trials (in this case, speech is generally 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 speak overtly.

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. 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, yet they presumably reflected a com-

bination of brain activities elicited by both 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 only 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). Furthermore, these protocols do not allow investigating neural processes involved in articulatory control.

Heavy Signal Filtering

One current strategy, adopted by various authors, is to elicit speech in a spontaneous manner (e.g. immediate picture naming). The recorded EEG signal, which is heavily muscle contaminated, is tentatively distinguished from muscle artifacts with heavy low-pass filters. For example, Masaki et al. (2001) and Ganushchak and Schiller (2008) used respectively 10 and 12 Hz low pass filters.

This method 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. Friedman and Thayer (1991) showed that EMG activity is present both in the alpha (8–13 Hz) and in the beta bands (13–20 Hz). Goncharova et al. (2003) reported that facial EMG has a broad frequency distribution, from almost DC to more than 200 Hz, even with weak muscle contraction (see especially their Fig. 3). Moreover, heavy low pass filtering has two major drawbacks: first, it prevents investigations of EEG activity present above 10 Hz (e.g. beta band), and, second, it may reduce dramatically the amplitude of phasic activities. An example of the latter problem comes from visually evoked potentials, whose components often last less than 100 ms and is therefore largely affected by 10 Hz filters (Luck 2005, see also below for a more thorough consideration of this problem).

Blind Source Separation

A possibly more promising approach to disentangle EEG and EMG signals comes from Blind Source Separation (BSS) techniques. These allow decomposing the signal in elementary sources; explicit hypotheses are then made to disentangle EEG and EMG signals among these sources. Since the current study is also based on BSS, we describe its principles in some more detail.

General Principle

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

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

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

The goal of BSS is to estimate the mixing matrix \mathbf{M} , and/or the source vector \mathbf{S} , given the observation matrix \mathbf{D} only (Comon and Jutten 2010). In other words, BSS tries to estimate the underlying, generating sources from the observed mixture. This is done by introducing the de-mixing matrix $\mathbf{W} \in \mathbb{R}^{J \times I}$ such that the estimated sources

$$\tilde{\mathbf{S}} = \mathbf{W} \cdot \mathbf{D}, \quad (2)$$

approximate the unknown physiological source signals in \mathbf{S} , up to a scaling factor. Ideally, \mathbf{W} is the inverse of the unknown mixing matrix \mathbf{M} , up to scaling and permutation.

Separating Sources

Different \mathbf{W} 's will provide different source estimations $\tilde{\mathbf{S}}$. In order to break down the EEG into elementary sources in a unique way, explicit assumptions about the sources have to be made and a decomposition algorithm has to be defined. The appropriateness of the assumptions will determine how well the estimated sources $\tilde{\mathbf{S}}$ approximate the real contributing sources \mathbf{S} .

We first discuss a decomposition based on the assumption of statistical independence, “Independent Component Analysis”, that has been previously used to discriminate EEG and EMG signals. We then present the decomposition method central to this article, “Canonical Correlation Analysis” (BSS-CCA), where sources are separated on the basis of autocorrelation properties.

Independent Component Analysis (ICA) A commonly used assumption for estimating \mathbf{W} is that the sources are mutually statistically independent, as well as independent from the noise components. In this context, ICA (Comon 1994) has been proposed to clean ictal

EEG from the EMG contamination (Weidong and Gotman 2004; Urrestarazu et al. 2004).

ICA has been shown to have some success in removing continuous EMG activity in the context of epileptic seizures. However, in several cases, the separation between EEG and EMG sources was not optimal (see e.g. De Clercq et al. 2006). In Crespo-Garcia et al. (2008), ICA was used to remove muscle artifacts from sleep recordings. However, since the authors did not show cleaned EEG data, it is hard to judge how appropriately ICA reached this goal. With respect to separating EEG from EMG signal, an interesting validation approach was proposed by McMenemy et al. (2009, 2010). The authors assessed in an objective way (based on bio-equivalence tests) the sensitivity and specificity of ICA as a tool to remove muscle artifacts. It was shown that ICA-based techniques outperformed regression-based correction techniques, but also highlighted that no perfect artifact removal was obtained with ICA.

One common aspect of these studies is that the categorization of the sources as EEG signal vs EMG noise was made manually by expert raters, on the basis of selected properties of the sources (e.g. topography, spectrum, time-course, etc...). Note also that this approach imposes constraints on the size of the data matrix (a rule of thumb is that for a data matrix $\mathbf{D} \in \mathbb{R}^{I \times T}$, T is in the order of $I^2 \times 20$).

Canonical Correlation Analysis (CCA) Canonical Correlation Analysis is a statistical method originally developed to measure the linear relationship between two multidimensional variables \mathbf{A} and \mathbf{B} (Hotelling 1936). A multidimensional dataset is represented in a predefined basis, that can be changed by rotation for example. 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. When applied to a time-series and its shifted version, the method provides an estimate of autocorrelation in the signal.

EMG activity is weakly autocorrelated over time: given its broad spectrum it tends to have white noise properties (Goncharova et al. 2003). In contrast, brain activity is considerably coherent over time, and tends to be more autocorrelated. Under the assumptions that (a) the shapes of EEG and EMG sources are uncorrelated, and (b) EEG sources are individually autocorrelated in time (see below for more details on these hypotheses), a BSS method based on Canonical Correlation Analysis (BSS-CCA) can be defined.¹

¹Similar applications of BSS-CCA to other signals can be found in Hardoon et al. (2004), De Vos et al. (2007).

In De Clercq et al. (2006), the BSS-CCA method was introduced for decomposing the EEG observation matrix \mathbf{D} into sources. The method yielded sources sorted in decreasing order of autocorrelation (highly autocorrelated sources ranked first, weakly autocorrelated ones ranked last). The sources with the highest autocorrelation should correspond to EEG, while the sources with the lowest autocorrelation should correspond to EMG. Expert neurologists inspected the reconstructed signal visually. They removed presumed muscle components one by one until ictal activity, otherwise completely masked in the background tonic EMG activity, became visible. 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 (Vergult et al. 2007).

One highly relevant aspect of this method for our current purpose is that the BSS method can be applied on relatively short stretches of signal, for example on single-trials of a psycho-neuro-linguistic experiment, which usually last between 1 and 3 sec.

The Current Study

Goal We investigated whether the BSS-CCA method could be used in practice to distinguish between cortical and EMG signals in electrophysiological recordings performed 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 (Alario et al. 2004; Glaser 1992). In this task, EMG contamination from the muscles involved in articulating words is not a tonic, continuous activity, but instead appears as short EMG bursts, localized in time around the window of interest. Here, we focused on the removal of phasic EMG contamination induced by the facial muscles involved in articulation. There are also other possible sources of EMG activity (see e.g. Whitham et al. 2007), but we do not explicitly address the issue of ongoing EMG activity, either related to the task or not, originating from other muscular patterns.

Hypotheses a) The recorded signal is a linear combination of EEG and EMG sources. b) The time courses of these sources are not strongly correlated. Coherence between EEG and (limb) EMG signals has been reported (Mima and Hallett 1999; Conway et al. 1995) but only in a rather narrow frequency range and associated with very different frequency spectra in two signals. The frequency range where this coherence was observed was rather narrow, however, and associated with very different frequency spectra in the two signals.

Such limited coherence is not expected to result in strong temporal correlation. c) The EMG signal has a broad frequency range, approximating white noise, thus its autocorrelation will be very low; on the contrary, the EEG signal has a much larger autocorrelation (De Clercq et al. 2006). Finally, d) the muscular pattern engaged for articulation may depend on the word to be uttered, hence a trial-by-trial identification of the EMG sources may be required (Chan et al. 2002).

Evaluation Ideally, an evaluation of a source separation method should provide an exhaustive evaluation of its sensitivity (i.e. how well it detects signal in noise) and of its specificity (i.e. how much signal is incorrectly classified as noise) (McMenamin et al. 2009). Unfortunately, this is not an easy task for the problem at hand, given that the brain activity (of main interest) is, by construction, in close temporal vicinity with overt articulation. Therefore one cannot construct the exhaustive factorial design devised for example by McMenamin et al. (2009). Eliciting covert speech would not solve the problem, as it would obviously not involve articulation and could thus modify the brain activity of interest. Furthermore without overt speech production, an important temporal marker of response execution (onset of speech) is unavailable.

Despite these limitations, we attempt to provide a careful assessment of the methodology we introduced. De Clercq et al. (2006) have already evaluated 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, both the unaffected signal and the EMG contamination are unknown. However, characteristic features of both phasic EMG signal and “clean” EEG are well known.

We will first present several analyses comparing directly the raw with the cleaned data, providing evidence that the clean data are more similar to characteristic EEG than the raw data. We will show the effect of BSS-CCA on single trial level, the change in frequency content after BSS-CCA and the topography of the muscle components. We will discuss the impact of computing averages with or without cleaning and we will compare in a statistical way the results of BSS-CCA with those from alternative available methods (ICA and frequency filtering).

Materials and Methods

This experiment was originally conducted for other (psycholinguistic) purposes that will be described

elsewhere (e.g. Riès et al. 2010). Only the aspects relevant for our current purposes are detailed here.

Participants and Task

Twelve right-handed native French-speakers (three females) with normal or corrected to normal vision and no diagnosed language pathology participated in the experiment (mean age: 24.5). They all gave their informed consent. Participants were presented with line drawings representing common objects, which they had to name as fast and accurately as possible (Alario and Ferrand 1999). Forty-five different drawings were used. The pictures were presented individually (11 × 11 cm). Participants were presented with a sequence of 20 blocks, each of which comprised the 45 pictures. The order of presentation within blocks was random.

Vocal responses were recorded with a software voice-key (Eprime 2.0 Professional, Pittsburgh, PA: Psychology Software Tools), and individually checked off-line for accuracy (using CheckVocal Protopapas 2007).

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 averaging off-line 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, and the horizontal EOG was recorded with two electrodes positioned over the two outer canthi.

Signal Preprocessing

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 until 2 s after stimulus onset.

Blind Source Separation with Canonical Correlation Analysis

Starting from Eq. 1

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

where \mathbf{D} is the EEG observation matrix, we constructed²

$$\mathbf{A} = \mathbf{D}(:, 1 : (\mathbf{T} - 1))$$

$$\mathbf{B} = \mathbf{D}(:, 2 : \mathbf{T})$$

\mathbf{A} is thus (a truncated version of) the observation matrix and \mathbf{B} is (a truncated version of) the same observation matrix shifted by one time sample. CCA will look for a linear combination of the signals \mathbf{A} that correlates best with a linear combination of \mathbf{B} . In practice all sources are estimated simultaneously by solving matrix equations (see Appendix and, e.g. Borga and Knutsson 2001; Friman et al. 2001; Golub and Van Loan 1996). For illustration purposes, we explain here the extraction of sources as if it was a sequential process.

The first extracted source (the first row of the \mathbf{S} matrix) will be a linear combination of the EEG observation matrix \mathbf{D} with maximized autocorrelation (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 whose coefficients are the regression weights of the linear combination; they provide the first row of the demixing matrix \mathbf{W} .

The data is then projected away from this basis vector (which ensures the mutual orthogonality between different sources), and a similar procedure is used to find the second source. The second extracted source will also be a linear combination with maximal autocorrelation, but this time under the constraint that it is uncorrelated to 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 are fully decomposed.

Different muscular patterns are associated with different words (Chan et al. 2002), hence the decomposition should be word-dependent. BSS-CCA was computed, as defined above, on each epoch separately, obtaining a \mathbf{W} matrix and source signals for each trial (duration 2.2 sec).

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 autocorrelation (highly autocorrelated sources ranked first, weakly autocorrelated ones ranked last). 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. One can define

²We use $(:)$ to denote that all channels are involved.

explicit quantitative criteria that provide an automatic classification. For example, a criterion can be defined on the basis of Power Spectral Density (PSD), because a differential power density across the sources can be assumed. EEG has lower power at high frequencies, while EMG has higher power at high frequencies. Here, components were considered to be EMG activity if their average power in the EMG band (approximated by 15–30 Hz) was at least $\frac{1}{n}$ (with n being by default set to 7) of the average power in the EEG band (approximated by 0–15 Hz). These values were empirically determined but, as will be shown below, the results do not appear to depend critically on them.

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, estimated as the inverse of \mathbf{W} . The new mixing matrix $\mathbf{M}_{\text{clean}}$ was then used to reconstruct the denoised EEG signal matrix $\mathbf{D}_{\text{clean}}$. More specifically, consider again the corresponding source estimation and signal decomposition

$$\hat{\mathbf{S}} = \mathbf{W} \cdot \mathbf{D} \text{ and } \mathbf{D} = \mathbf{M} \cdot \mathbf{S} \quad (3)$$

In BSS-CCA the sources are ordered by decreasing autocorrelation, thus $\mathbf{M}_{\text{cleanEEG}}$ and \mathbf{M}_{EMG} can be defined as:

$$\mathbf{M}_{\text{cleanEEG}} = \mathbf{M} \cdot \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ \vdots & \vdots & \ddots & \\ & & & 0 \\ & & & & \ddots \end{pmatrix} \text{ and}$$

$$\mathbf{M}_{\text{EMG}} = \mathbf{M} \cdot \begin{pmatrix} 0 & 0 & \dots \\ \vdots & \ddots & \\ & & 1 \\ & & & 1 \\ & & & & \ddots \end{pmatrix} \quad (4)$$

$\mathbf{D}_{\text{cleanEEG}}$, containing the cleaned EEG, and \mathbf{D}_{EMG} , containing muscle artifacts, are then

$$\mathbf{D}_{\text{cleanEEG}} = \mathbf{M}_{\text{cleanEEG}} \cdot \mathbf{S} \text{ and } \mathbf{D}_{\text{EMG}} = \mathbf{M}_{\text{EMG}} \cdot \mathbf{S} \quad (5)$$

Finally, the processed EEG segments with muscle artifacts removed, were re-epoched, this time centered on the verbal response onset.

Topography of the Removed Components

Although the topography was never taken into account for selecting EMG sources, we do have strong expectations about the location of those EMG sources:

they should mainly be close to the facial muscles, hence largely present on frontal electrodes. As an extra validation criterion, a representative topography can be obtained by normalizing the estimated source signals to a variance of one, and weighting 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.

More specifically, as defined in Eq. 4, only the last columns will be retained in \mathbf{M}_{EMG} . These coefficients are constant within trials, and hence do not reflect the exact EMG contribution at any time instant, since this contribution is the product of the topography (given by \mathbf{M}_{EMG}) and the time course of the component, which is ignored in this analysis. However, they give an estimate of how much each EMG source loads on the observation signal on a given trial. Averaging EMG mixing vectors across trials provides a global representation of EMG topography. A similar reasoning can be applied to the effect of filtering via Fourier transform. Indeed, it was already derived in the literature that singular value decomposition (SVD) can be interpreted as a filter (Hansen and Jensen 1998).

Comparison with Other Methods

We implemented two previously used procedures to provide a comparison with BSS-CCA. In the first method, we applied a low-pass filter (10 Hz, 24db/oct) to the EEG signal. In the second method, we applied an ICA algorithm based on higher order statistics, namely RobustICA (Zarzoso and Comon 2008) to the EEG signal. RobustICA is an improved version of the popular fastICA algorithm. We decomposed every trial with RobustICA and used exactly the same criterion for selecting the EMG-related sources as we proposed for BSS-CCA. This allows for a direct comparison on the quality of the decomposition itself.

Statistical Evaluation

In order to compare the impact of EMG removal with BSS-CCA, low-pass filtering, and ICA we computed the peak-to-peak amplitude of the visually evoked potentials. This was done for two electrodes: Fp2, a frontal electrode close to the facial EMG sources, and P3, a more posterior electrode where the visual evoked potentials are larger. For Fp2, the peak-to-peak value is defined as the difference between the first clear negativity and the following positivity. The values are computed in the time interval 60–190 ms. For P3, the value was the difference between the first negativity and the preceding positivity (see Fig. 6). This value is

computed in the time interval 40–160 ms. The peak-to-peak value was estimated for all participants and its variation across BSS methods was assessed by means of Student-*t* tests.

Furthermore, in order to evaluate the influence of the parameters chosen above for the source selection criterion, we compared three different parameter settings (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). We then computed the relative amount of removed variance for the different settings. We also compared the impact of the different parameter settings on the peak-to-peak values of the averaged ERPs, as defined above. The influence of the parameter settings on these measures was assessed with a multivariate repeated measures ANOVA.

Results

We illustrate in different ways the effect of BSS-CCA. First, we derive qualitative descriptions on one representative participant (mean response time (RT) 686 ± 108 ms, mean error rate $1.8\% \pm 0.96\%$) with considerable EMG contamination. Although we focus on this subject, similar results are obtained on all other participants, as the statistical evaluation will show. Afterwards, we also present some quantitative descriptions, to confirm that our results are globally valid.

Qualitative Description

Following clinical practice (e.g. De Clercq et al. 2006), we first provide a qualitative description of the findings. Figure 1 presents the data for a single epoch over five anterior-central channels (FPz, AFz, Fz, FCz and Cz). The removed components present a sizeable EMG burst around response time. Such burst is absent in the reconstructed EEG signal, which shows a typical pattern of EEG traces.

The same pattern of results is apparent for two representative electrodes in Fig. 2, where all the trials of a given participant are represented with the ERP-image technique (Jung et al. 2001). The comparison of the original epochs (panels a and c) and the reconstructed signal after removal of the identified EMG components (panels b and d) shows a reduction of the high frequency noise (visible as a pixelisation of the plot), especially around response time.

An overview of these observations is given by the power spectra computed on a sample of 19 electrodes (Fig. 3). The comparison of the original data and the

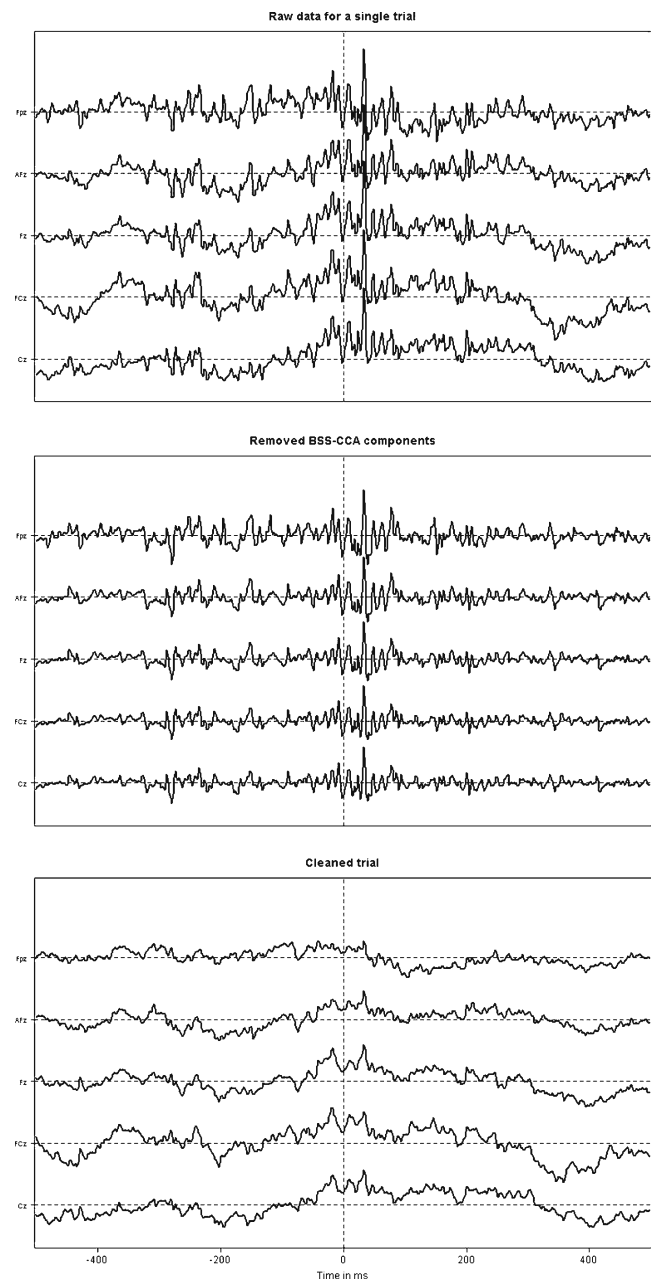


Fig. 1 Single trial of EEG data on five channels around voice onset (0 ms). The *top panel* presents the original recorded signal, the *middle one* shows the components that BSS-CCA removed and the *bottom panel* presents the reconstructed EEG data after EMG removal. The *removed components* correspond to high frequency activity, and the *cleaned trial* contains the low-frequency fluctuation

signal reconstructed after BSS-CCA shows a reduction of power in high frequencies, and reveals the expected $1/f^\alpha$ shape.

Finally, we plotted the topography of the removed muscle components in Fig. 4. Most of the “activation” is frontal, as expected given that muscles involved in

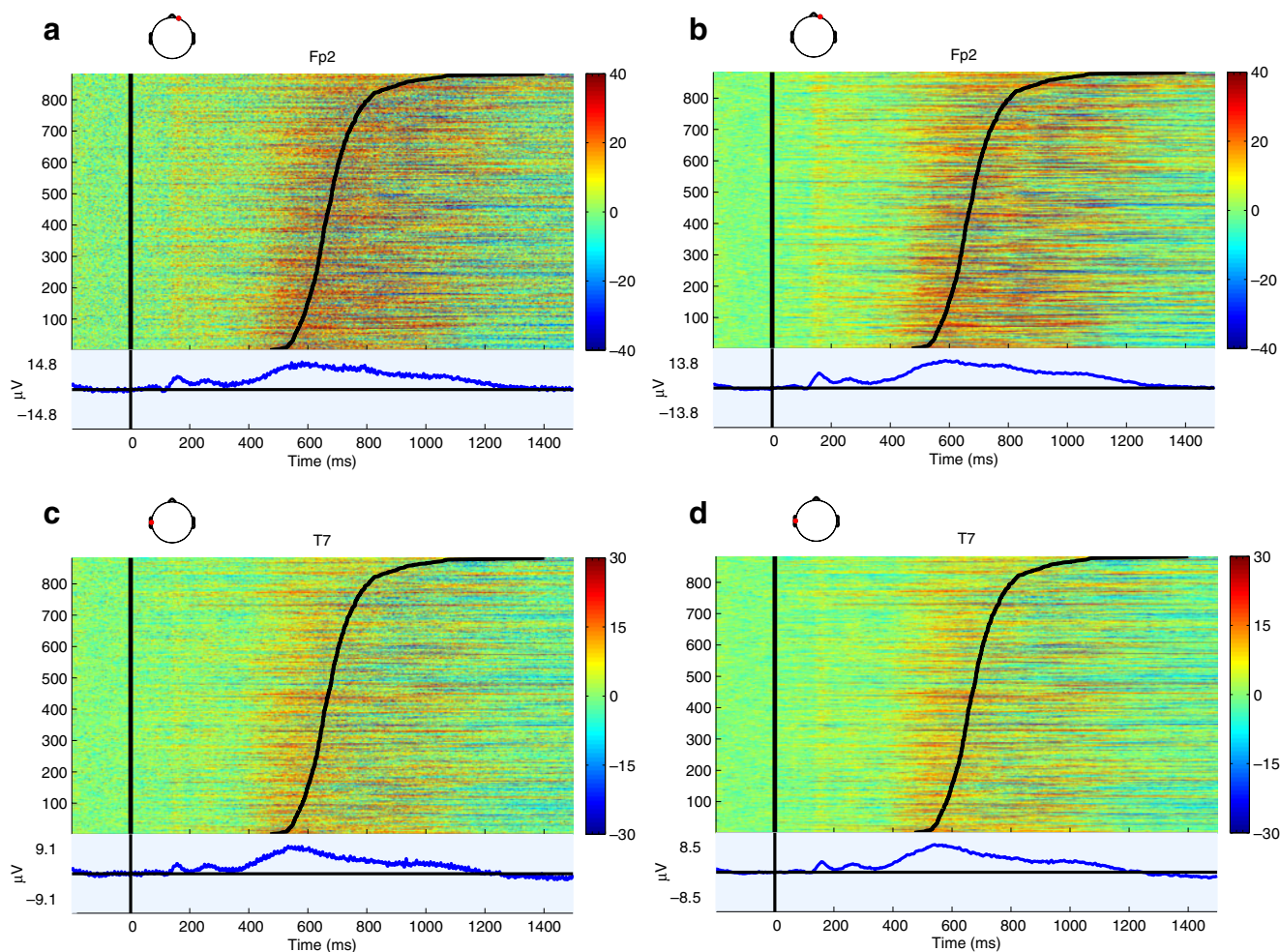


Fig. 2 ERP image (a.k.a. raster plot) for all individual trials of a given participant at channels Fp2 and T7 before (a,c) and after (b,d) removal of the identified EMG components. The trials are represented as parallel color lines, where color codes the polarity and strength of the activity. The trials 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). The blue line at the bottom of each panel represents the average of all trials. On the left panel, the high frequency components appear as a pixelisation of the graph. Such a pixelisation is largely reduced after BSS-CCA (right panel)

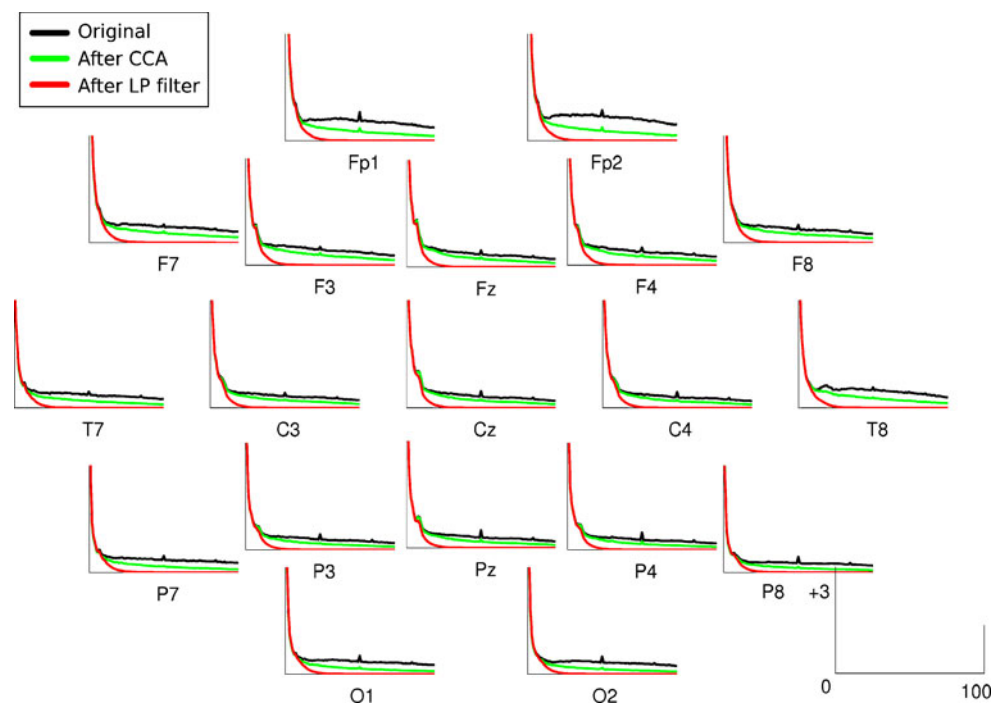
speech production are mainly facial. Some EMG contributions of the neck muscles are also observed, which may correspond to postural activity stabilizing the head during speech.

Impact of Muscle Removal on the Average ERP

One goal of EMG removal is to improve the quality of the measurement. Since 1) EMG activity is different from trial to trial, 2) EEG activities are induced by the cognitive operation of interest, and 3) EMG activity is assumed to be independent from EEG activity, the averaging procedure across trials reduces the impact of EMG and reveals the EEG activities of interest. If BSS-CCA removes the EMG part of the signal, less

trials should be necessary to reach an equivalent quality of the average. To assess this point, we performed separated averages on a subset N of all the trials, and compared the quality of such averages to the grand average for the same subject of all the trials, before and after cleaning. We consider this ERP as the golden standard. N was varied from 15 to 480, and for each N , 100 random subsamples were used. Figure 5, presents the mean correlation coefficients between these subaveraged ERPs and the ERPs of the grand average for Fp2 and P3. On channel P3, there is very little difference between the averages before and after cleaning. This can be expected, since little muscle activity is present on this channel. On Fp2, to the contrary, the EEG is severely contaminated. The impact of EMG removal is

Fig. 3 Averaged spectral content (between 0 and 100 Hz) of all epochs on a sample of 19 channels represented topographically in the *black* original data, *green* after EMG removal with BSS-CCA, and *red* after low-pass filtering. The green and red spectra have a $\frac{1}{f^\alpha}$ shape. However, the green spectrum follows the original spectrum more closely

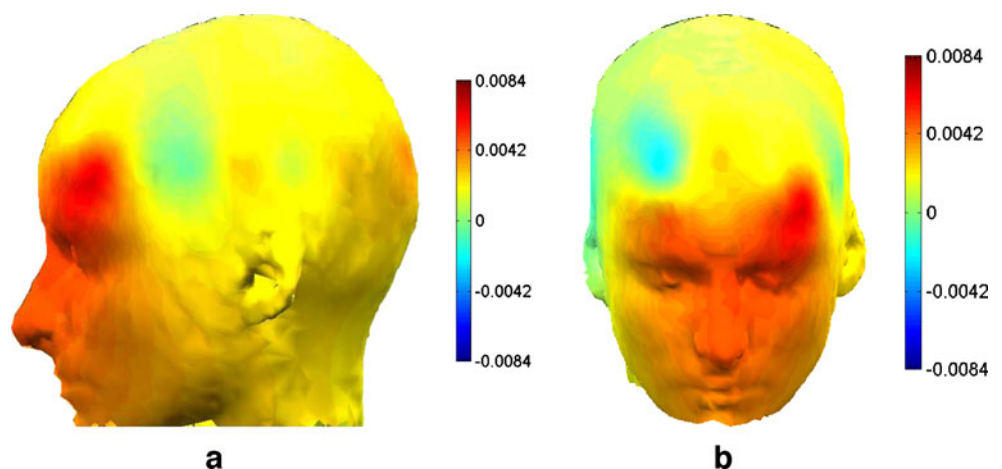


very large on this channel. Indeed, for a given number of averaged trials, the correlation is much higher after BSS-CCA than on raw data. For example, with only 30 trials, the correlation coefficient is already close to .85 after BSS-CCA, while it is less than .70 for raw data. It is also evident that, for a given correlation coefficient, data after BSS-CCA requires consistently fewer than half the trials required using raw data (at least when few trials are averaged). For example, to reach a correlation coefficient of 0.9, only 50 trials are needed with BSS-CCA cleaning, while 200 trials are needed with raw data. This represents an impressive improvement.

Quantitative Assessment: Comparison Between BSS-CCA and Alternative Cleaning Methods

We quantified the impact of the various source separation methods on the evoked potentials. To do so, we first averaged all the trials of each individual participant, thus obtaining average traces per participant. We then measured the amplitude (see Material and Methods section above) of the visual evoked potentials observed for each participant. This was done for the raw data, the BSS-CCA processed data, the low-pass filtered data, as well as the ICA processed data. We

Fig. 4 Averaged topography of the removed muscle components. Mainly frontal activity, corresponding to face muscles, can be observed. Note that the colors outside the EEG electrodes are not very reliable, due to extrapolation. Panel (a) gives a lateral view and (b) a frontal view. The scale is expressed in Volts



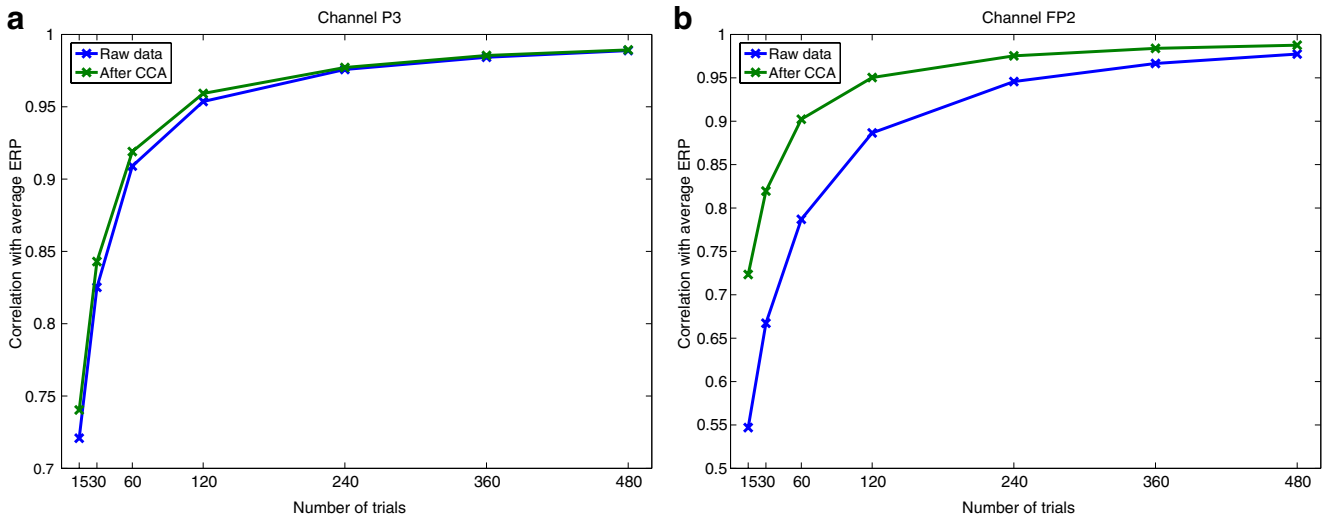


Fig. 5 Correlation between the average of N trials and the grand average of raw and cleaned trials, where N is varied from 15 to 480. (a) on P3 and (b) on Fp2. On Fp2, where there is a lot of

EMG contamination, much less trials are needed to obtain the same ERP quality than without cleaning

then computed Student t -tests to compare the amplitude of the potentials obtained through the various methods.

The results of the comparisons are presented on Fig. 6 and Table 1. Figure 6 indicates a large amplitude reduction of the evoked potentials after low-pass filtering. This is confirmed by statistical analysis: evoked potentials were significantly lower after

heavy low-pass filtering than after BSS-CCA ($t(11) = 5.54, p < 0.001$ on Fp2 and $t(11) = 6.67, p < 0.001$ for P3). By contrast, there was no significant difference between the amplitude measured on the raw data and on BSS-CCA data for Fp2 ($t(11) = 1.51, p = 0.10$), while a significant reduction was observed on P3 ($t(11) = 3.29, p < 0.01$). Note that we did not correct for multiple testing.

Fig. 6 The averaged ERP on two channels. Both the full epoch (from -0.2 ms to 2 s) and a zoomed version (from -0.1 s to 0.4 s) are shown. We compare the unfiltered ERP and the ERP after muscle removal with BSS-CCA and with filtering. BSS-CCA follows much more closely the ERP from the raw data than filtering

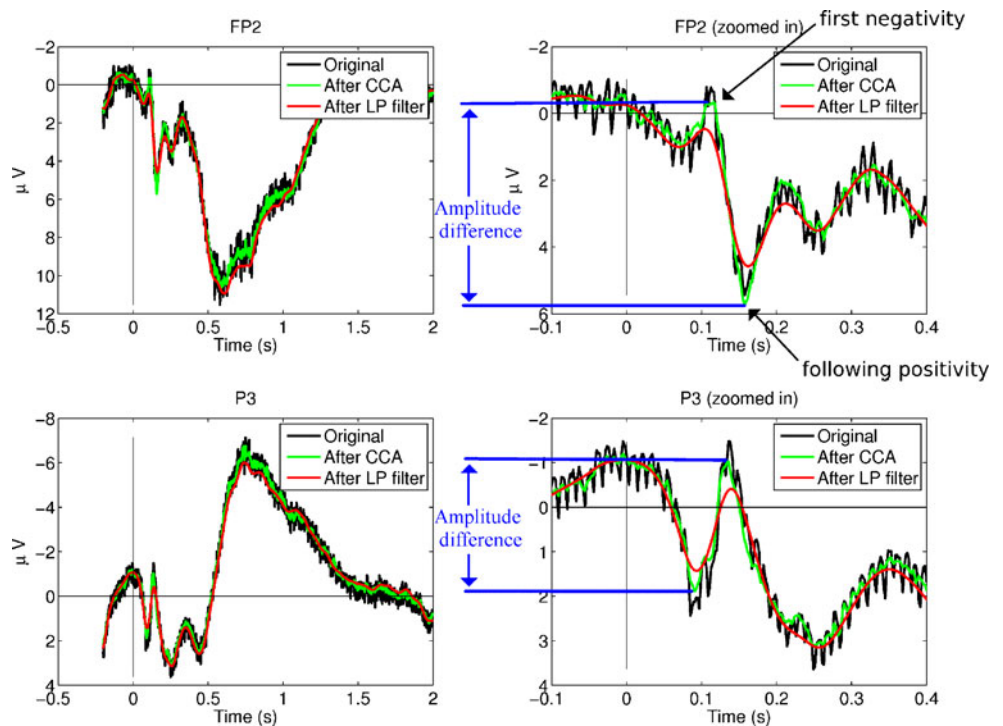


Table 1 The average peak-to-peak value (and its standard deviation) on Fp2 and P3 over the 12 subjects

Channel	Raw data	BSS-CCA	ICA	Filtering
Fp2	8.36 ($\sigma = 3.25$)	7.98 ($\sigma = 3.05$)	7.67 ($\sigma = 2.70$)	6.10 ($\sigma = 2.37$)
P3	5.06 ($\sigma = 2.27$)	4.48 ($\sigma = 2.33$)	4.45 ($\sigma = 2.15$)	2.90 ($\sigma = 1.94$)

The value is defined as the difference between the first negativity and the following positivity for Fp2, and the difference between the first negativity and the preceding positivity for P3. Filtering largely reduced this amplitude difference, while BSS-CCA least reduced the amplitude

We then compared the evoked potentials obtained after cleaning with BSS-CCA and with ICA. Figure 7 indicates that CCA was able to attenuate high frequency noise better. This was confirmed by statistical analyses: there was a significant difference in amplitude between the peak-to-peak values measured on the raw data compared to those measured after ICA processing ($t(11) = 4.44$, $p < 0.05$ and $t(11) = 5.92$, $p < 0.001$, for Fp2 and P3, respectively).

Influence of Selection Criterion on EMG Removal

Contrary to previous reports using BSS-CCA to remove EMG activities (De Clercq et al. 2006; Vergult et al. 2007), we introduced an automatic parameterized criterion to detect muscle components. The criterion is based on a ratio between power in the “EMG” band and power in the “EEG” band. Two parameters are

thus necessary: the lower limit for the “EMG” band and the value for the ratio (see Materials and Methods for more details).

We compared three different parameter settings: frequency limit of EMG band set at 15 Hz and ratio set at 1/7, then 16 Hz and 1/10, finally 13 Hz and 1/5. The averaged ERPs for three parameters settings are presented in Fig. 8. Tables 2 and 3 summarize the removed variance and the peak-to-peak amplitudes for the different parameter sets. Clearly, the average ERP is almost unaffected by the parameter choice. Indeed, although changing the selection criterion will to some extent modify the selected sources and hence affect the removed variance, the peak-to-peak values are not significantly different ($p = 0.61$). Note that the fact that the averages are very similar does not imply that single trial denoising was not affected by the different parameters settings. However, such small changes, do

Fig. 7 Averaged ERP on two channels. We compare the unfiltered ERP, the ERP after muscle removal with BSS-CCA and with filtering by ICA. BSS-CCA removes more high-frequency noise and reduces the peak-to-peak values less than ICA

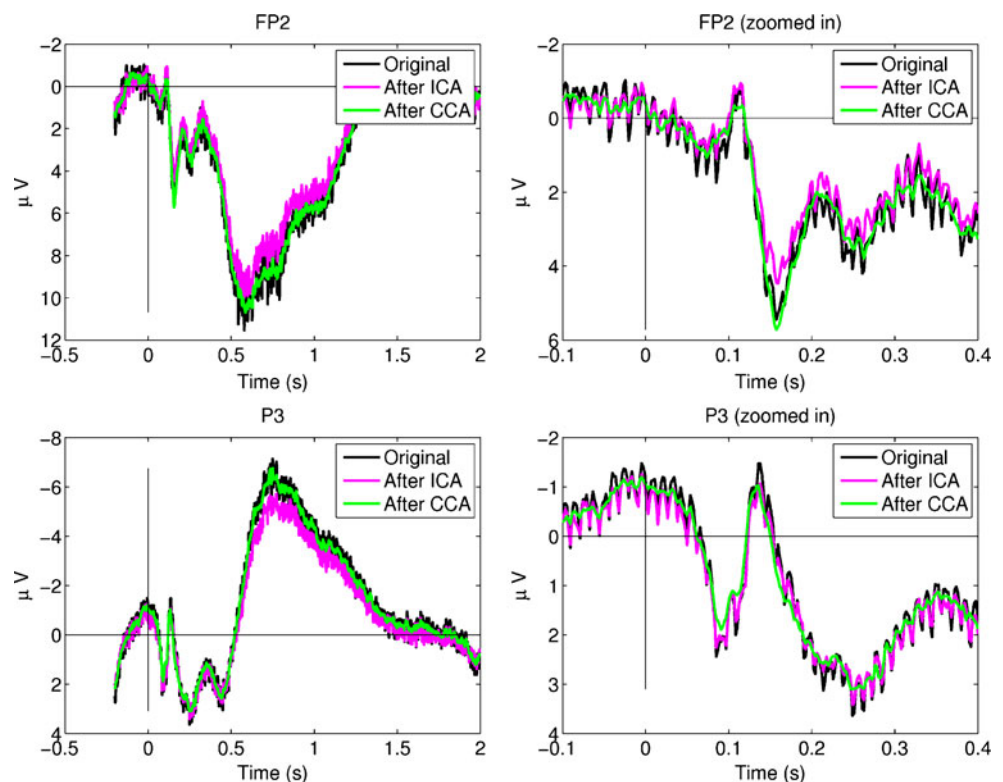
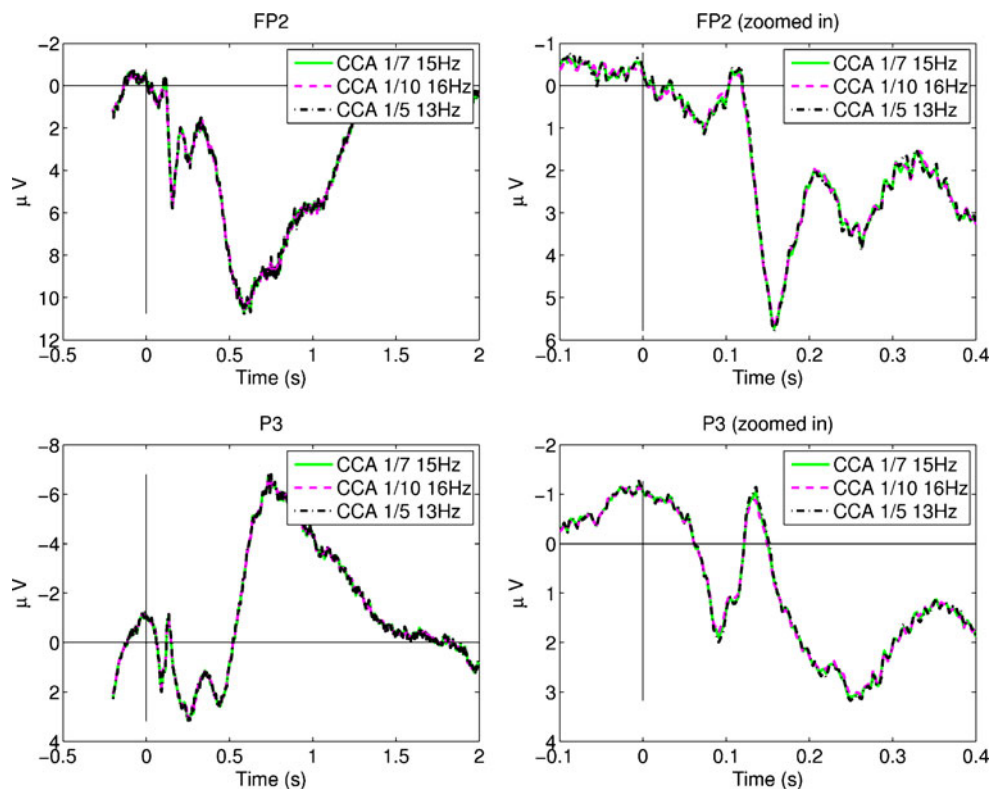


Fig. 8 Averaged ERP on two channels for three different source selection parameter settings of BSS-CCA (see text for details). The influence of source selection parameters on the final ERP is minor



not seem to be critical, since the averages were virtually equivalent. This guarantees the robustness of the method.

Discussion

Psycholinguists have long avoided electrophysiological (EEG and MEG) investigations of spoken language because of (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 10s epochs of epilepsy recordings. In the present study, we adapted and automated the method, and applied it to the removal of short bursts of myographic activity related to speech production.

Table 2 The relative removed variance for three different parameter settings in the selection of EMG sources on Fp2 and P3. Relative variance is not much influenced by the parameter choice

Channel	5/13	7/15	10/16
Fp2	0.10 (std 0.08)	0.11 (std 0.09)	0.12 (std 0.09)
P3	0.053 (std 0.037)	0.059 (std 0.041)	0.063 (std 0.044)

We automated BSS-CCA by selecting the muscle components based on their power spectrum. The components that were selected with this criterion were the least autocorrelated. This confirms the appropriateness of the BSS-CCA decomposition.

Validating artifact removal techniques in this context is a challenging task since neither 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 in various ways.

Sensitivity

BSS-CCA provided a considerable reduction of the EMG artifacts in the EEG signal recorded during speech. The benefits of BSS-CCA are visible on a trial by trial basis (Figs. 1 and 2), and also on the power spectra of the grand averages (Fig. 3). Before

Table 3 In order to assess the impact of the different parameters in the selection of EMG sources, we compare peak-to-peak values (average and standard deviation) for the three different source selection parameter settings of BSS-CCA (see text for details)

Channel	5/13	7/15	10/16
Fp2	8.07 (std 3.06)	7.98 (std 3.05)	7.94 (std 3.02)
P3	-4.52 (std 2.29)	-4.47 (std 2.32)	-4.43 (std 2.34)

muscle artifact removal, the power spectra showed high frequency activity. After source separation, the power spectra had a clear $1/f^\alpha$ shape on all electrodes. This is exactly the shape of the EEG power spectrum reported by Goncharova et al. (2003) in the condition where participants were asked not to contract facial muscles (see their Fig. 3). The $1/f^\alpha$ shape is obtained both with filtering and after BSS-CCA, although the actual α value in filtering will be higher, and the spectra after filtering will not follow the original spectra as well as after BSS-CCA.

In order to further validate the EMG removal by BSS-CCA, we studied the part of the signal that was removed by the algorithm. The average topography (Fig. 4) of the rejected sources was mostly frontal, which fits well with the anatomical positions of the muscles involved in speech production. Interestingly, the locations where most activity was found on these topographies also correspond to where the power spectra of the raw data contained most high frequencies.

Specificity

The results discussed above show the sensitivity of the BSS-CCA method in attenuating EMG activity. A symmetrical question concerns its specificity, i.e. whether the method specifically targets EMG while leaving EEG signal “intact”. To address this issue, we focused on the activities that are most often the subject of attention in psycho-neuro-linguistic EEG studies, namely event-related activities. In this approach, the theoretically relevant differences between experimental conditions are differences in ERP amplitudes (e.g. Osterhout et al. 1997).

On these ERPs, the removal of EMG contamination with BSS-CCA preserved the original amplitudes significantly more than EMG removal through low-pass filtering (Fig. 6) or through ICA (Fig. 7). On Fp2, the statistical test performed on the peak-to-peak amplitudes across participants shows no significant difference between the values on raw data and the values after BSS-CCA. On P3, all the methods induce a significant reduction of the peak-to-peak value in a pair-wise test. However, the reduction in amplitude is the smallest after BSS-CCA.

These observations may be unsurprising, given that ERPs are generally rather phasic and thus contain high frequencies, especially those that lie close to the event to which they are time locked, be it stimulus presentation or overt response triggering. Applying heavy low-pass filtering will remove those high frequency components and hence dramatically reduce the amplitude of the evoked potentials. Differences between conditions

visible on amplitudes may be reduced, or will even disappear as a consequence of heavy low-pass filtering. As BSS-CCA selects a subset of sources based on differential autocorrelations instead of removing all the high frequency components indiscriminately, it does not (or significantly less) flatten these peaks.

Again, although we cannot produce explicit specificity values given the nature of the data at hand, the analysis suggests more specificity with BSS-CCA than with filtering or ICA.

Other Considerations

We also compared BSS-CCA for EMG removal to another BSS algorithm based on higher order statistics: ICA (Zarzoso and Comon 2008). ICA has been proposed and validated as a method to remove artifacts and in particular EMG artifacts (e.g. McMenamin et al. 2010 and references therein). However, there are different reasons why ICA might not be optimal. 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 (as we did), but accurate estimate of high order statistics requires a large number of samples. As every trial is limited in time, there are maybe not enough samples for ICA to reliably decompose the EEG into brain and muscle activity. Although based on our results we can not claim that BSS-CCA is superior to ICA in all applications of EMG removal, our results show that at least for this particular application BSS-CCA better preserves the shape of the ERPs while removing more disturbing signal and is thus more appropriate than ICA.

Practical Application

A direct example of the benefit brought by BSS-CCA to the field of psycholinguistic research comes from a recent study of speech monitoring (Riès et al. 2010). The authors report the observation of an EEG component, known as the error-negativity (Ne or error-related negativity, ERN), which reaches its maximum shortly after response-onset (i.e. precisely during intense articulation movements). Unprecedentedly, Riès et al.

(2010) showed that the error-negativity component is present in correct utterances as well, although with a considerably smaller amplitude. Presumably because of articulation-related artifacts, heavy low-pass filtering procedures were used in previous studies (e.g. 1–12 Hz band-pass filters in Ganushchak and Schiller 2008; 10 Hz low pass in Masaki et al. 2001), and this may be a cause for the non-observation of the Ne in correct utterances. The Ne can still be observed on errors after such filtering, as it generally involves a rather large deflection. By contrast, the Ne observed on correct trials appears to be smaller (in-line with results from studies of non-speech action control) and thus might have been filtered out. In Ries et al.'s study, heavy filters were avoided because articulation artifacts were successfully reduced with the BSS-CCA algorithm.

Conclusions

Whether or not part of the brain signal is also removed by BSS-CCA remains an important point of concern and will have to be addressed in future studies. Further validation can be also performed with bio-equivalence tests, described in Shackman et al. (2009) and McMenamin et al. (2009). Future experiments may further clarify possible limitations. Other neuroscientific research groups are also encouraged to also use this algorithm (freely available at www.neurology-kuleuven.be/index.php?id=210) on their event-related potentials and start thorough speech research.

Our study clearly showed that BSS-CCA outperformed heavy filtering for the removal of EMG artifact. This method therefore enables more accurate neuroscientific investigations of spoken language production without avoiding direct overt speech. Besides, the results open perspectives towards new applications, both with continuous and event-related EEG.

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, GOA-MANET, 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), G.0321.06 (Tensors/Spectral Analysis), G.0302.07 (SVM), G.0341.07 (Data fusion), G.0427.10N (Integrated EEG-fMRI), research communities (ICCoS, ANMMM); IWT: TBM070713-Accelerator, 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), FAST

(FP6-MC-RTN-035801), Neuromath (COST-BM0601) ESA: Cardiovascular Control (Prodx-8 C90242), European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013 Grant Agreement no. 241077)

Information Sharing Statement The original BSS-CCA method is available at www.neurology-kuleuven.be/index.php?id=210. The proposed automatization can be obtained after sending an email to the corresponding author.

Appendix: 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 :

$$\rho = \frac{\text{Cov}[\mathbf{a}, \mathbf{b}]}{\sqrt{V[\mathbf{a}]V[\mathbf{b}]}} \quad (6)$$

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} and \mathbf{B} , and two new scalar variables, $\tilde{\mathbf{a}}$ and $\tilde{\mathbf{b}}$, generated as linear combinations of the components in \mathbf{A} and \mathbf{B} :

$$\begin{aligned} \mathbf{A} &= [\mathbf{a}_1(t), \dots, \mathbf{a}_m(t)]^T \\ \mathbf{B} &= [\mathbf{b}_1(t), \dots, \mathbf{b}_n(t)]^T, t = 1, \dots, N \\ \tilde{\mathbf{a}} &= w_{a_1}\mathbf{a}_1 + \dots + w_{a_m}\mathbf{a}_m = \mathbf{w}_a^T \mathbf{A} \\ \tilde{\mathbf{b}} &= w_{b_1}\mathbf{b}_1 + \dots + w_{b_n}\mathbf{b}_n = \mathbf{w}_b^T \mathbf{B} \end{aligned} \quad (7)$$

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 Eq. 7 into the definition of the correlation coefficient (6), and assuming the means of \mathbf{A} and \mathbf{B} zero, we obtain:

$$\rho = \frac{\mathbf{w}_a^T \mathbf{C}_{ab} \mathbf{w}_b}{\sqrt{(\mathbf{w}_a^T \mathbf{C}_{aa} \mathbf{w}_a)(\mathbf{w}_b^T \mathbf{C}_{bb} \mathbf{w}_b)}} \quad (8)$$

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:

$$\begin{aligned} \mathbf{C}_{aa}^{-1} \mathbf{C}_{ab} \mathbf{C}_{bb}^{-1} \mathbf{C}_{ba} \mathbf{w}_a &= \rho^2 \mathbf{w}_a \\ \mathbf{C}_{bb}^{-1} \mathbf{C}_{ba} \mathbf{C}_{aa}^{-1} \mathbf{C}_{ab} \mathbf{w}_b &= \rho^2 \mathbf{w}_b \end{aligned} \quad (9)$$

This system is an eigenvalue decomposition. The matrices $C_{aa}^{-1}C_{ab}C_{bb}^{-1}C_{ba}$ and $C_{bb}^{-1}C_{ba}C_{aa}^{-1}C_{ab}$ have the same eigenvalues. The vectors w_a and 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.

References

- 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.
- Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods Instruments & Computers*, 36, 140–155.
- 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.
- Chan, A., Englehart, K., Hudgins, B., & Lovely, D. (2002). Hidden markov model classification of myoelectric signals in speech. *IEEE Engineering in Medicine and Biology Magazine*, 21, 143–146.
- Comon, P. (1994). Independent component analysis, a new concept? *Signal Process*, 36, 287–314.
- Comon, P., & Jutten, C. (2010). *Handbook of blind source separation, independent component analysis and applications*. Academic Press.
- Conway, B. A., Halliday, D. M., Farmer, S. F., Shahani, U., Maas, P., Weir, A. I., et al. (1995). Synchronization between motor cortex and spinal motoneuronal pool during the performance of a maintained motor task in man. *Journal of Physiology (London)*, 489(3), 917–924.
- Crespo-Garcia, M., Atienza, M., & Cantero, J. (2008). Muscle artifact removal from human sleep EEG by using independent component analysis. *Journal of Biomedical Engineering*, 36, 467–475.
- 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.
- Glaser, W. R. (1992). Picture naming. *Cognition*, 42, 61–105.
- Golub, G., Van Loan, C. F. (1996). *Matrix computations* (3rd ed.). Baltimore: John Hopkins University Press.
- 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 offline removal of ocular artifacts. *Electroencephalography and Clinical Neurophysiology*, 55, 468–484.
- Hansen, P. C., & Jensen, S. H. (1998). Fir filter representations of reduced-rank noise reduction. *IEEE Transactions on Signal Processing*, 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.
- 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., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., Sejnowski, T. (2001). Analysis and visualization of single-trial event-related potentials. *Human Brain Mapping*, 14, 166–85.
- Luck, S. L. (2005). *An introduction to the event-related potential technique*. MIT Press.
- 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.
- McMenamin, B., Shackman, A., Maxwell, J., Bachhuber, D., Koppenhaver, A., Greischar, L., et al. (2010). Validation of ica-based myogenic artifact correction for scalp and source-localized EEG. *Neuroimage*, 49, 2416–2432.
- McMenamin, B., Shackman, A., Maxwell, J., Greischar, L., & Davidson, R. (2009). Validation of regression-based myogenic correction techniques for scalp and source-localized EEG. *Psychophysiology*, 46, 578–592.
- Mima, T., & Hallett, M. (1999). Corticomuscular coherence: a review. *Journal of Clinical Neurophysiology*, 16(6), 501–511.
- Morrell, L. K., Huntington, D. A., McAdam, D. W., & Whitaker, H. A. (1971). Electroocortical localization of language production. *Science*, 174(4016), 1359–1360.
- Osterhout, L., McLaughlin, J., & Bersick, M. (1997). Event-related brain potentials and human language. *Trends in Cognitive Sciences*, 1, 203–209.
- Protopapas, A. (2007). Checkvocal: A program to facilitate checking the accuracy and response time of vocal responses from dmdx. *Behavior Research Methods*, 39, 859–862.
- Riès, S., Janssen, N., Dufau, S., Alario, F.-X., & Burle, B. (2010). General purpose monitoring during speech production. *Journal of Cognitive Neuroscience* (in press).
- Shackman, A., McMenamin, B., Slagter, H., Maxwell, J., Greischar, L., & Davidson, R. (2009). Electromyogenic artifacts and electroencephalographic inferences. *Brain Topography*, 22, 7–12.

- 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 Turenout, M., Hagoort, P., & Brown, C. M. (1998). Brain activity during speaking: From syntax to phonology in 40 milliseconds. *Science*, *280*, 572–574.
- Vergult, A., De Clercq, W., Palmini, A., Vanrumste, B., Dupont, P., Van Huffel, S., et al. (2007). Improving the interpretation of ictal scalp eeg: Bss-cca algorithm for muscle artifact removal. *Epilepsia*, *48*, 950–958.
- 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., Pope, K., Fitzgibbon, S. L. T., Clark, C., Loveless, S., Broberg, M. A. W., et al. (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)* (pp. 3344–3347). Vancouver, Canada.