

General-Purpose Monitoring during Speech Production

Stéphanie Riès¹, Niels Janssen^{1,2}, Stéphane Dufau¹, F.-Xavier Alario¹,
and Borís Burle¹

Abstract

■ The concept of “monitoring” refers to our ability to control our actions on-line. Monitoring involved in speech production is often described in psycholinguistic models as an inherent part of the language system. We probed the specificity of speech monitoring in two psycholinguistic experiments where electroencephalographic activities were recorded. Our focus was on a component previously reported in nonlinguistic manual tasks and interpreted as a marker of monitoring processes. The error negativity (Ne, or error-related negativity), thought to originate in medial frontal areas, peaks shortly after erroneous responses. A component of seemingly comparable properties has been reported, after errors, in tasks requiring access to linguistic knowledge (e.g., speech production), compatible with a generic error-

detection process. However, in contrast to its original name, advanced processing methods later revealed that this component is also present after correct responses in visuomotor tasks. Here, we reported the observation of the same negativity after correct responses across output modalities (manual and vocal responses). This indicates that, in language production too, the Ne reflects on-line response monitoring rather than error detection specifically. Furthermore, the temporal properties of the Ne suggest that this monitoring mechanism is engaged before any auditory feedback. The convergence of our findings with those obtained with nonlinguistic tasks suggests that at least part of the monitoring involved in speech production is subtended by a general-purpose mechanism. ■

INTRODUCTION

The concept of “monitoring” refers to our ability to control our actions on-line. Here we investigated the mechanism by which we monitor what we say. Speech is a seemingly effortless and highly efficient process: Usually speech production elicits around one error every 1000 words (Levelt, 1989). Yet, we continuously control what we are saying and what we are about to say. This speech monitoring ability has been investigated from a psycholinguistic perspective (for detailed reviews, see Postma & Oomen, 2005; Postma, 2000). At stake in this research are issues related to the nature of the linguistic representations involved, the possible automaticity of monitoring processes (central or attention requiring vs. automatic), and the contrast between monitoring occurring before versus after sounds are actually produced. This last point has introduced a distinction between, respectively, the inner loop, which monitors inner linguistic representations, and the outer loop, relying on auditory feedback (e.g., Postma & Oomen, 2005). Beyond these considerations, one implicit (occasionally explicit) assumption that is commonly made is to consider speech monitoring as an inherent part of the language processing system. For example, Postma’s (2000) review suggests that a central perception-based monitor (as defined in Levelt, 1989)

may be common to speech production and other control situations (e.g., manual responding in forced-choice tasks), without discarding the possibility that speech-specific monitoring may also be required.

Consistent with Postma’s suggestion, functional imaging (fMRI or PET) of speech monitoring have reported activities specific to verbal monitoring tasks as well as activities common to speech and action monitoring. Most of these studies have focused on monitoring of verbal auditory feedback (i.e., the outer loop mentioned above) by manipulating the actual feedback that participants heard as they spoke. This research has highlighted the role of the left temporal cortices (for a meta-analysis, see Indefrey & Levelt, 2004). Activation of the superior temporal gyrus is modulated by the nature of distorted auditory feedback (e.g., pitch elevation of the participant’s own voice; Tourville, Reilly, & Guenther, 2008; Fu et al., 2006; Hashimoto & Sakai, 2003; McGuire, Silbersweig, & Frith, 1996). When normal and pink-noise-masked feedback are contrasted, activation in the superior temporal gyrus is attenuated in the normal situation, whereas there is increased activity in other temporal regions and in the medial frontal cortex, including the ACC and the SMA (Christoffels, Formisano, & Schiller, 2007). In another type of experiment, involving a verb generation task where participants produced a verb of their choice in response to a noun, ACC activity was associated with verbal response conflict monitoring (Barch, Braver, Sabb, & Noll, 2000; for a full theoretical development, see Botvinick, Braver, Barch, Carter, & Cohen, 2001). In the tasks

¹Aix-Marseille Université & CNRS, Marseille, France, ²Universidad de La Laguna, Tenerife, Spain

used in the study of Christoffels et al. (2007), investigations of functional connectivity between temporal and medial-frontal regions suggest that their activations are inversely related (Van de Ven, Esposito, & Christoffels, 2009).

In manual nonlinguistic studies of response monitoring, the involvement of medial frontal cortex has been repeatedly observed (for a review, see Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004). ACC activity appears to be modulated by the degree of competition between alternative responses and has been linked to response conflict monitoring (Carter et al., 1998; see, however, Burle, Roger, Allain, Vidal, & Hasbroucq, 2008). To our knowledge, the recruitment of the temporal cortices has solely been reported when speech monitoring was involved. One plausible interpretation of these findings from speech and manual nonlinguistic response monitoring is that medial frontal regions correspond to an aspecific monitoring process (as suggested by Christoffels et al., 2007; Barch et al., 2000), whereas the temporal cortex responds to auditory linguistic processing (Hickok & Poeppel, 2000). However, activities in the temporal cortices and in the ACC seem both modulated by auditory feedback (Christoffels et al., 2007). Indeed, it is not clear from these studies which and whether these brain regions could be involved in inner speech monitoring. Also, a limitation of some of these studies is that acoustic distortions could promote enhanced aspecific cognitive control, which may not be present in normal speech monitoring.

Our main aim in the current study was to test for the recruitment of an aspecific monitoring system in speech production using electroencephalography. Given the temporal precision of the method used, we also consider the type of representations (related to inner vs. outer speech) this type of common system might monitor. Secondarily, by comparing direct overt speech production to a manual linguistic task, we also look for EEG components that may reflect monitoring related to overt speech. We hypothesized that the study of electrophysiological correlates of speech monitoring will help shed light on the specificity of the medial frontal activity and on the nature of the representations it monitors.

The error negativity (Ne) or error-related negativity has been extensively studied in nonverbal tasks (first reports: Ne: Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; error-related negativity: Gehring, Goss, Coles, Meyer, & Donchin, 1993) and more recently in linguistic tasks also (Ganushchak & Schiller, 2006, 2008a, 2008b, 2009; Möller, Jansma, Rodriguez-Fornells, & Münte, 2007; Sebastian Gallés, Rodríguez-Fornells, de Diego-Balaguer, & Díaz, 2006; Masaki, Tanaka, Takasawa, & Yamazaki, 2001). The Ne has been shown to originate in the ACC and/or the SMA (e.g., Debener et al., 2005; Dehaene, Posner, & Tucker, 1994). It consists of an event-related negative potential starting around the onset of the response-related EMG activity and peaking between 100 and 150 msec later (the Ne starts to rise before response onset when time locked to this event, e.g., Falkenstein, Koshlykova, Kiroj, & Hoormann, 1995).

After the Ne, a positive component named the “error positivity” (Pe) was described. Contrarily to the Ne, the Pe is present only when the subject is conscious of his error (Endrass, Reuter, & Kathmann, 2007; O’Connell et al., 2007; Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001). The Ne component was first reported on nonverbal erroneous responses (Gehring et al., 1993; Falkenstein et al., 1991); accordingly, it was first related to an error-detection process. However, an Ne-like wave has been reported on both correct and erroneous trials during nonverbal tasks (Bartholow et al., 2005; Vidal, Burle, Bonnet, Grapperon, & Hasbroucq, 2003; Vidal, Hasbroucq, Grapperon, & Bonnet, 2000). The Ne-like wave on correct trials was smaller, reached its maximum of negativity slightly earlier than on erroneous trials, and was typically followed by a second, generally smaller negativity. The topographical distribution of Ne-like waves is fronto-central, just as it is on erroneous trials. The Ne-like waves observed on correct trials and errors were interpreted to be of comparable origin (Roger, Bénar, Vidal, Hasbroucq, & Burle, 2010). This suggests that the Ne reflects a response monitoring system, secondarily leading to error detection rather than error detection in itself. The early onset of this component, along with data collected on a completely deafferented patient (Allain, Hasbroucq, Burle, Grapperon, & Vidal, 2004), indicates that it is dissociated from sensory feedback.

Under the previously formulated working hypothesis that medial frontal response monitoring is subtended by a generic process involved in different cognitive functions, including language, an Ne-like wave should be observed both on correct and incorrect speech utterances. In contrast with this prediction, the Ne elicited in overt speech tasks has only been reported after errors (Ganushchak & Schiller, 2008a; Möller et al., 2007; Masaki et al., 2001). Although this could be due to intrinsic differences between manual and verbal tasks (direct long-term memory vs. arbitrary stimulus/response relationship, differences in response modalities, etc.), it is also possible that this absence is related to methodological differences: First, most studies reporting an Ne-like wave on correct nonlinguistic trials used current source density analysis (Yordanova, Falkenstein, Hohnsbein, & Kolev, 2004; Vidal et al., 2000, 2003) that enhances the spatial resolution of the EEG signal (Babiloni, Cincotti, Carducci, Rossini, & Babiloni, 2001; Gevins, Leong, Smith, Le, & Du, 1995) and reveals small local activities by removing remote sources contributions (Vidal et al., 2003). Because of volume conduction, overlapping activities hinder the small negativity on correct trials, with conventional monopolar recordings. As a matter of fact, studies that did not report an Ne-like wave on correct trials were not based on the analysis of current source density. Furthermore, because the Ne-like wave on correct trials tends to be smaller in amplitude than on errors, it is more likely to be masked by articulation-related artifacts (Goncharova, McFarland, Vaughan, & Wolpaw, 2003; Brooker & Donald, 1980). Articulation-related artifacts are probably the main reason why speech monitoring has been

studied in paradigms avoiding overt speech and using manual responses (e.g., Ganushchak & Schiller, 2006, 2008b, 2009). Although these tasks involve linguistic representations stored in long-term memory, they differ from overt speech production. Indeed, it is uncertain that the processes they engage (including monitoring) are directly comparable to those naturally occurring in speech production.

In any event, an Ne-like wave (or CRN: correct response negativity) has been reported in two manual linguistic studies (Ganushchak & Schiller, 2009; Sebastian Gallés et al., 2006). Both of these experiments tested bilingual subjects in elaborate experimental settings, designed to elicit high error rates. Moreover, neither of these studies tested for the statistical presence of this negativity (see Methods therein). Interestingly, however, both of these studies suggest that the amplitudes of the negativity reported in correct and erroneous trials increased in case of uncertainty of the response. This finding suggests that the Ne and the Ne-like wave could be components reflecting a similar process, although this was not assessed in these studies.

We report below two experiments that test for the presence of postresponse activities with a special focus on the Ne-like wave after correct trials. We used two distinct linguistic tasks: a grammatical gender decision task (manual responses) and an overt picture-naming task (speech production). The fMRI results discussed above suggest that the medial frontal cortex activity, reported in both speech and nonspeech monitoring and supposedly reflected by the Ne, may be shared across different cognitive functions. If the Ne reflects a generic monitoring system also used for speech monitoring, then an Ne-like wave should be observed on correct speech utterances, pending methodological improvements (see below). Moreover, the timing of these waves may be informative about the distinction usually made between inner and outer loops in speech monitoring tasks. For example, if the time course of this negativity is comparable with that of the Ne reported in previous studies (the Ne starts to rise before the response), it will be unlikely that it depends on auditory feedback linking it preferentially to inner speech monitoring (Postma, 2000). This would be consistent with the property of Ne observed in nonlinguistic tasks to be independent from sensory feedback. Finally, given the fMRI results described above, we will also look for EEG activities on electrodes located over the temporal cortices. However, this type of activity has not yet been reported using current source density analysis and will have to be more cautiously interpreted than the better-known medial frontal activity.

METHODS

Experiment 1

In French, nouns are either of masculine or of feminine grammatical gender. To a large extent, this feature is an arbitrary property of the nouns (Corbett, 1991). Accordingly, access to grammatical gender has been repeatedly

used, in French and other languages, as an index of access to lexical information stored in long-term memory (e.g., Navarette, Basagni, Alario, & Costa, 2006; van Turennout, Hagoort, & Brown, 1998). Here participants were asked to press a left or a right button depending of the grammatical gender of the picture names. This task offers a direct comparison with nonlinguistic manual tasks and with previous studies where it was used to investigate speech monitoring. Note that in contrast with the two studies that used manual linguistic tasks and reported an Ne-like wave in correct trials (Ganushchak & Schiller, 2009; Sebastian Gallés et al., 2006), participants were not uncertain of what the correct response for each trial was.

Participants

Fifteen right-handed native French speakers (8 women) with normal or corrected-to-normal vision participated in the experiment (mean age = 21.8 years). The data of five participants were removed from the behavioral and electrophysiological analysis because of problems during the EEG recordings (too many artifacts around the response or too much noise on fronto-central electrodes).

Materials and Design

The stimuli were line drawings of common objects selected from published collections (Bonin, Peereman, Malardier, Mot, & Chalard, 2003; Alario & Ferrand, 1999) or designed by us for the experiment. Their name agreement was above 70%. The stimuli were presented in black on an 11 × 11-cm white square, within a visual angle of 2.22°. The experiment was controlled by a microcomputer running Tscope (Stevens, Lammertyn, Verbruggen, & Vandierendonck, 2006). A total of 116 drawings were used: 108 were the experimental items and 8 were used in warm-up trials.¹ Half of the experimental items had a feminine grammatical gender and half had a masculine grammatical gender (for more information, see Appendix).

Each experimental run comprised 110 trials: Two of the eight warm-up items were followed by the 108 experimental items. Each participant was tested on eight runs.

Within each experimental run, the order in which the groups of items were presented was mixed pseudorandomly using the software MIX (Van Casteren & Davis, 2006). There were never more than four identical expected responses (left thumb or right thumb, for picture names having masculine or feminine grammatical gender, respectively) in a row, and two consecutive items were never phonologically related (i.e., two pictures in a row never had the same initial phoneme).

Procedure

Overall, the experiment lasted for about 3 hours. Participants were tested in a sound-attenuated dimly lit environment. They were seated in a Faraday cage in front of a screen. They

held vertical handgrips on top of which response buttons were fixed.

Participants were asked to perform two tasks during independent/separate blocks: an object orientation task, which will not be detailed here, and a grammatical-gender decision task where participants had to press the right button if the name of the picture was feminine and the left one if the name of the picture was masculine. Participants were instructed to answer as fast and as accurately as possible and to keep their thumbs on the response buttons at all time to minimize RTs. They were also asked to remain as relaxed as possible and to avoid making movements that could generate artifacts on the EEG recordings (e.g., eye blinks, frowning) during the experimental trials.

A trial consisted of the following events: (1) a fixation point (“plus” sign) for 500 msec, (2) the picture appeared and remained on the screen until the participants responded or after a deadline of 1500 msec was reached, and (3) a blank screen for 1000 msec. The next trial started automatically.

There were three parts in the experiment. First, participants were familiarized with the 116 drawings used in the experiment. These were presented one by one in a random order, and the participant was asked to name each picture. The experimenter made verbal corrections when an incorrect or an unexpected response was produced. Second, there was a practice phase in which participants performed one block of the task they had to perform in the experiment proper. Third, the experiment proper started and lasted for about an hour.

Electrophysiological Recordings

The EEG was recorded from 64 Ag/AgCl Active-2 pre-amplified electrodes (BIOSEMI, Amsterdam; 10–20 system positions). The sampling rate was 1024 Hz (filters: DC to 208 Hz, 3 db/octave). The passive reference electrode was placed over the left mastoid. The vertical EOG was recorded by means of two surface electrodes (Ag/AgCl) just above and below the left eye, respectively. The horizontal EOG was recorded with two electrodes positioned over the two outer canthi. For each hand, the EMG activity of the flexor pollicis brevis was recorded by means of two surface electrodes (Ag/AgCl) glued to the skin of the thenar eminence. The distance between the two EMG electrodes was 2 cm.

Data Processing

After acquisition, the EEG and the EMG data were filtered (EEG: high pass = 0.16 Hz; EMG: high pass = 10 Hz). Eye movement artifacts were corrected using the statistical method of Gratton, Coles, and Donchin (1983). All other artifacts were rejected after a trial-by-trial visual inspection of monopolar recordings. The onset of EMG was marked manually after visual inspection, a method that is notoriously more accurate than automated algorithms (Staudé,

Flachenecker, Daumer, & Wolf, 2001; Van Boxtel, Geraats, van den Berg-Lenssen, & Brunia, 1993).

No responses, that is, trials where the participant did not answer within the 1500 msec limit, were removed from further analysis. Trials were classified as correct or erroneous depending on whether the correct or the incorrect button was pressed first. Among correct trials, those containing only one EMG activation on the correct side (referred to as “pure correct” trials) were distinguished from those containing one EMG activation on the incorrect side preceding an EMG activation on the correct side (partial errors, see, e.g., Burle, Possamaï, Vidal, Bonnet, & Hasbroucq, 2002). Clear errors were defined as trials where only one EMG activation was observed on the incorrect side. RTs was taken as the time between stimulus presentation and EMG onset (either correct or incorrect for, respectively, pure correct trials and errors). EEG data corresponding to pure error trials and to “pure correct” trials were averaged separately, first for each subject, and then a grand-average time locked to EMG onset was performed.

Laplacian transformation (which provides an estimate of the current source density), as implemented in BrainAnalyser™ (Brain Products, Munich, Germany), was applied on the individual averages, and a grand average was created from those individual averages. Laplacian transformation has been shown to increase the spatial resolution of the signal providing a good estimation of the corticogram (Nuñez, 1981). The signal was first interpolated with the spherical spline interpolation procedure, then second derivatives in two dimensions of space were computed (Legendre polynomial: 15° maximum). We chose 3 for the degree of spline because this value minimizes errors (Perrin, Bertrand, & Pernier, 1987). We assumed a radius of 10 cm for the sphere representing the head. The resulting unit was $\mu\text{V}/\text{cm}^2$. A 100-Hz low-pass filter was applied off-line on the EEG data.

Analysis

The analysis included the factor “accuracy” (correct or error) and “participants” as random effect. The behavioral data were analyzed using Student’s *t* tests or ANOVAs for comparisons of more than two means. Concerning the accuracy analysis, the pure correct trials were compared with the pure error trials.

For the EEG data, analysis were performed on three electrode sites of interest: FCz, where the Ne and Ne-like waves have been described (Vidal et al., 2000), and T7 and T8, over the left and the right temporal cortices because these regions have also been proposed to be involved in speech monitoring. We performed statistical analysis on three types of measures obtained from the Laplacian transformed data: (1) the latency of the peaks of interest, (2) the peak-to-peak amplitude (i.e., the difference between the amplitude of two consecutive peaks of activity), and (3) the slopes of the activity in the windows of interest. Latencies were measured on smoothed data (the length

of the smoothing window was 40 msec) to minimize the impact of background noise. The peak-to-peak amplitude was preferred over simple amplitude measurements as it is independent from the chosen baseline. Let us consider the Ne: The peak-to-peak amplitude measured here is the difference between the considered negativity and the preceding positive peak. To reduce the impact of background noise on the performed measures, peak-to-peak amplitudes were measured as follows: The latencies of the Ne and the preceding positive peak were measured for each participant for correct and incorrect trials. A 40-msec fixed time window was then determined around the averaged latency for the positive dip (respectively, the negative peak), providing an identical baseline for errors and correct trials. The surfaces between the curves and the baseline were measured for each participant during this time window. Finally, the difference between the surface measured around the preceding positivity and around the Ne was computed, and it is this surface difference that we will refer to as the peak-to-peak amplitude. Slopes were measured, by fitting a linear regression to the data, to attest for the existence of a component by comparing them to zero. This measure was chosen as it is also independent for the baseline and it gives morphological information about the data (see Carbonnell, Hasbroucq, Grapperon, & Vidal, 2004). They were measured on a time window between the average latency value of the first positive dip and the average latency value of the negativity. Such slope measures were not made in the two linguistic manual tasks reporting an Ne-like wave on correct trials (Ganushchak & Schiller, 2009; Sebastian Gallés et al., 2006). What these authors reported are comparisons between surface measures. However, the statistical existence of these surfaces was not tested (i.e., significantly different from zero and therefore from background noise). Here, we wanted to make the statistical presence of activities explicit. For the negativity observed at electrode T7, the same type of measures were performed.

All the abovementioned measures were compared using exact Wilcoxon signed-rank tests (also known as Wilcoxon T tests) because the measures were based on few error trials errors and the normality of the data could not be assumed. Following Siegel (1956), we reported Wilcoxon T values, corresponding to the sum of the absolute values of the ranks of the least represented sign, and the associated p values.

Experiment 2

Overt picture naming requires lexical access and oral responses. It is one of the most commonly used tasks to study overt speech. Because we wanted to compare our results with those reporting the presence of an Ne-like wave on correct trials, we used a comparable methodology. We calculated the surface Laplacian transformation of the EEG data, enabling to reveal small but consistent activities hindered by remote source in monopolar recordings. Fur-

thermore, we addressed the problem of articulation-related artifacts (most notably, EMG activity) with a blind source separation algorithm. This was expected to discriminate the bursts of EMG activity linked to articulation from the EEG signal. This method has shown successful results for tonic EMG activity induced by epileptic seizures (De Clercq, Vergult, Vanrumste, Van Paesschen, & Van Huffel, 2006) and has been adapted to the situation we are facing here (De Vos et al., 2010; Vanderperren et al., 2008).

Participants

A total of 16 right-handed native French speakers (7 women) with normal or corrected-to-normal vision participated in the experiment (mean age = 23.6 years). None of these participants participated in Experiment 1. The data of four participants were removed from the behavioral and the electrophysiological analysis because of problems during the EEG recordings (similar reasons to those of Experiment 1).

Materials and Design

Forty-five line drawings of common objects were used as stimuli (Alario & Ferrand, 1999). They had a mean name agreement of 96% and were displayed as in the previous experiment (for more information, see Appendix). There were 10 experimental runs. Each of the 45 experimental items appeared in a pseudorandom order twice per block such that two consecutive items were semantically and phonologically unrelated.

Procedure

The experiment was controlled by the software Eprime 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA), which allows on-line recording and voice-key triggering of the participants' verbal responses. The participant's task was to name the pictures presented as fast and as accurately as possible. For the reasons stated in Vidal et al. (2000), participants were informed that no correction was possible in the case of errors.

A trial consisted of the same events as in Experiment 1, except that participants named the pictures aloud instead of providing the grammatical gender. Response latencies were measured from the onset of the stimulus to the beginning of the vocal response by means of a voice key (contained in Eprime 2.0). The accuracy of this measure was checked visually and corrected when necessary, with the software CheckVocal (Protopapas, 2005).

There were three phases in the experiment. First, participants were familiarized with the materials as in Experiment 1. Second, the piezoelectric microphone sensitivity was tested and adjusted to the voice of the participant while he or she was reading words presented on the screen. Third, the experimental instructions were delivered and

the experiment started. The experimental session lasted for about an hour. There was a short break after each run.

Electrophysiological Recordings

They were as in Experiment 1, except for the following points. The sampling rate was 512 Hz (filters: DC to 104 Hz, 3 db/octave). The passive reference was obtained by off-line averaging the signal recorded over the left and right mastoids.

Data Processing

Trials where recording failures occurred, for example, when the voice key triggered for a reason not linked to the participant's voice or when the participant did not answer within the 1500-msec limit, were removed from further analysis. Trials were coded as errors when the participant produced any kind of verbal error: partial or complete production of incorrect words, verbal disfluencies (stuttering, utterance repairs, etc.). These errors were all coded in a single category. Incorrect trials could also be made of two utterances, the first one being a partial or complete error and the second one being correct in most cases. All were included in the general "error" category. Importantly, the marker indicating the onset of the error was always placed at the beginning of the sound waveform of the first utterance.

After acquisition, filtering, and correction of eye movement artifacts, we used a blind source separation algorithm on the basis of canonical correlation analysis (BSS-CCA; De Clercq et al., 2006) to reduce the EMG artifacts induced by articulation. The suitability of BSS-CCA for distinguishing articulatory EMG bursts from EEG signal with considerably improved signal to noise ratio is shown in De Vos et al. (2010). In this application, the BSS-CCA method was applied on nonoverlapping consecutive windows of 1.5 sec corresponding to the duration of a trial (by means of the EEGLAB plug-in Automatic Artifact Removal implemented by Gómez-Herrero available at <http://www.cs.tut.fi/~gomezher/projects/eeg/software.htm#aar>). Components were separated on the basis of their degree of autocorrelation (EMG activity tends to have low autocorrelation), and EMG components were selected according to their power spectral density. Components were considered to be EMG activity if their average power in the EMG band was five times bigger than the average power in the EEG band. This method provided satisfying results on our data (for details, see De Vos et al., 2010; Vanderperren et al., 2008).

After applying the BSS-CCA, all other artifacts were manually rejected (as in Experiment 1). Then, the remaining monopolar recordings were averaged time locked to vocal onset. Finally, Laplacian transformation was applied on these averages using the same parameters as in Experiment 1. A 100-Hz low-pass filter was applied off-line on the EEG data as for Experiment 1.

Analysis

The behavioral and the EEG data were analyzed as in Experiment 1, except that the only factor in the behavioral analysis was "accuracy".

RESULTS

Experiment 1

Behavioral Data

The average RT was 591 msec ($\sigma = 109$ msec) for "pure" correct trials and 595 msec ($\sigma = 126$ msec) for "pure" errors (defined in the Methods section). There was no effect of accuracy on RTs, $t(9) = -0.23$. The average error rate was 9.3% ($\sigma = 5.9\%$). On average, 91% of pure correct trials ($\sigma = 9.3\%$) and 88% of clear errors ($\sigma = 12\%$) were left after artifact rejection.

Electrophysiological Data

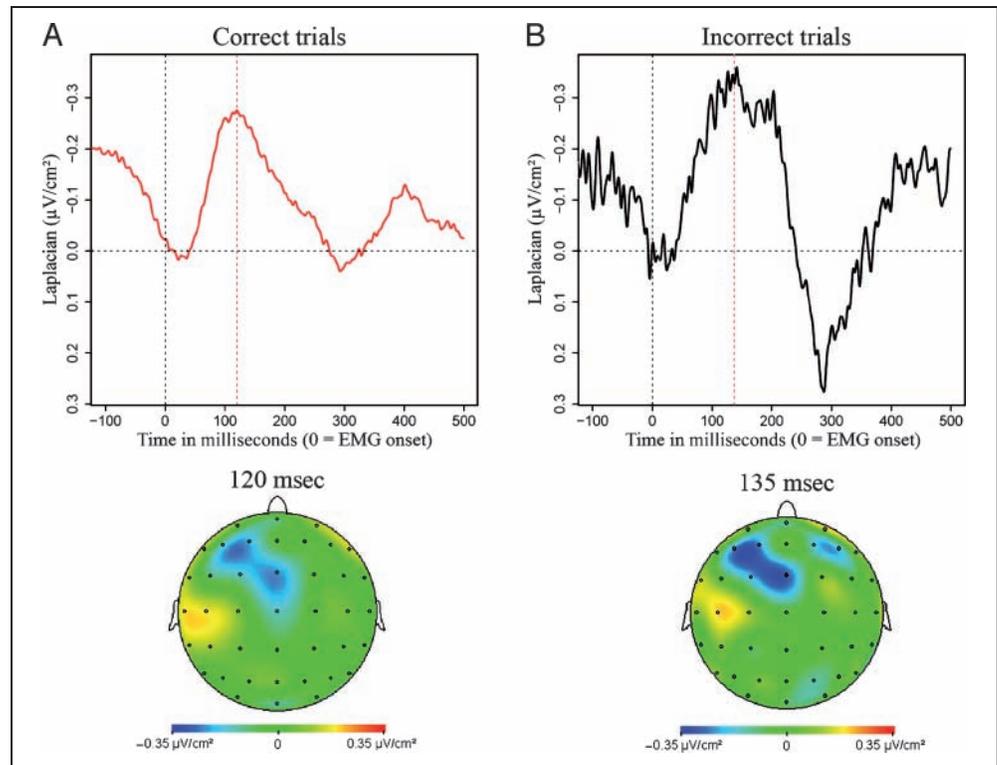
On correct trials, Laplacians showed a clear negativity peaking on average 124 msec after EMG onset ($\sigma = 26$ msec; Figure 1A). The Laplacian cartography indicates that this negativity reaches its maximum fronto-centrally, around electrode FCz. On the time window spanning from 30 to 110 msec, post-EMG onset slopes were significantly different from 0, $T(10) = 0$, $p < .01$. On incorrect trials, we observed a negativity peaking on average 147 msec ($\sigma = 30$ msec) post-EMG onset (Figure 1B). The Laplacian cartography indicates that it reaches its maximum fronto-centrally, especially around electrode FCz. The slopes measured on the same time window as on correct trials were significantly different from 0, $T(10) = 2$, $p < .01$.

We then compared errors and correct trials. The latencies measured were not significantly different for the considered positivity (positive dip latency on errors: 14 msec post-EMG onset, $\sigma = 44$ msec; on correct trials: 25 msec, $\sigma = 20$ msec; $T(10) = 21$, $p = .56$; average = 20 msec post-EMG onset; $\sigma = 34$ msec). However, the negativity observed on correct trials peaked before the negativity observed on errors, $T(10) = 0$, $p < .01$ (average latency of the negative peak for correct trials = 124 msec, $\sigma = 26$ msec; average latency of the negative peak for errors = 147 msec, $\sigma = 30$ msec), in agreement with previous reports (e.g., Roger et al., 2010).

The amplitude of the negativity on correct trials was smaller than the amplitude of the negativity observed on incorrect trials, $T(10) = 3$, $p < .01$ (average for correct trials = $-0.30 \mu\text{V}/\text{cm}^2$, $\sigma = .16$; average for errors = $-0.44 \mu\text{V}/\text{cm}^2$, $\sigma = .16$; see Figure 2 for individual data).

After the first negativity, we observed a positivity on errors but not on correct trials (greater positive slope on errors: $T(10) = 4$, $p < .05$, time window = 170–270 msec post-EMG onset; see Figure 1).

Figure 1. Ne on correct trials (A) and in errors (B) EEG activity after surface Laplacian transformation, recorded over FCz on (A) correct trials and (B) errors. Zero of time corresponds to EMG onset. Below the temporal data, the cartographies of this negativity time locked to EMG onset. These cartographies show the average of the spatial distribution of this activity on a 40-msec time window centered around the maximum of the peak of the grand average (on correct trials, 125 msec post-EMG onset; on errors, 136 msec post-EMG onset). The baseline was the 100 msec preceding the EMG onset as the activity of interest followed EMG onset. For the topographies, the EEG data were mirrored so that the response given was always right-handed. The lateralized activity therefore appears contralaterally to the response as described in Roger et al. (2010). Electrodes P7 and P8 were removed for all participants because they were too noisy for 1 of the 10 participants included in the analysis, and they could have induced an incorrect spline interpolation.



Experiment 2

Behavioral Data

Overall 0.48% of the trials were removed from further analysis (due to no responses or voice key problems). The average error rate was 1.31% ($\sigma = 0.96\%$). Fifty-five percent of errors were made of two or more responses (e.g., hesitations with filler “hum” sounds). Three participants made less than five errors throughout the experiment. They were therefore excluded from the error and from the comparison between erroneous and correct trials analysis. The average RT for correct trials was 651 msec ($\sigma = 72$ msec). The average RT for incorrect trials was 682 msec ($\sigma = 109$ msec). The trend for errors to yield a longer RT than correct trials was not significant, $t(8) = -0.44, p = .674$.

Electrophysiological Data

On correct trials, we observed a negativity peaking 44 msec after vocal onset on average ($\sigma = 41$ msec; see Figure 3A). The Laplacian cartography at the maximum of the negativity on the grand average (90 msec) indicates that it reached its maximum fronto-centrally, especially around electrode FCz. We then tested the existence of this negativity on the time window spanning from 40 msec prevocal onset to 40 msec postvocal onset (determined as in Experiment 1). The slopes were significantly different from 0, $T(12) = 13, p < .05$.

On incorrect trials, a clear negativity was observed fronto-centrally (on FCz) peaking on average 34 msec after vocal

onset ($\sigma = 24$ msec; Figure 3B). The Laplacian cartography showing the spatial distribution of this negativity at its maximum on the grand average (25 msec) indicates that

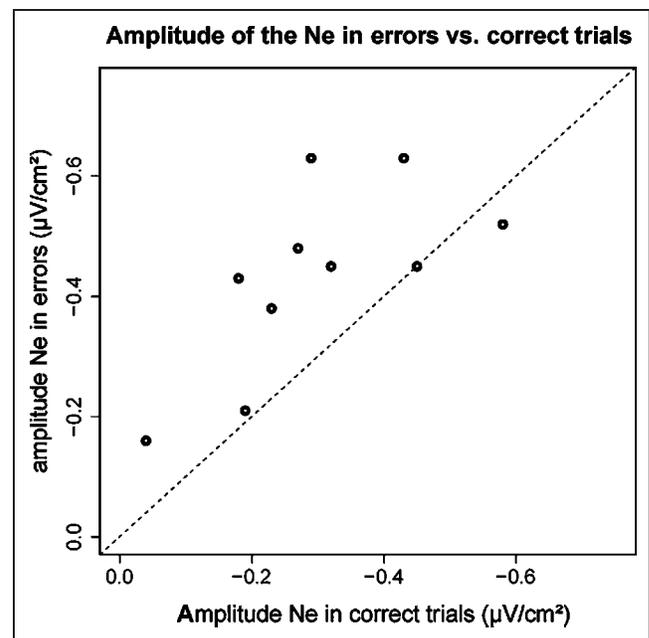
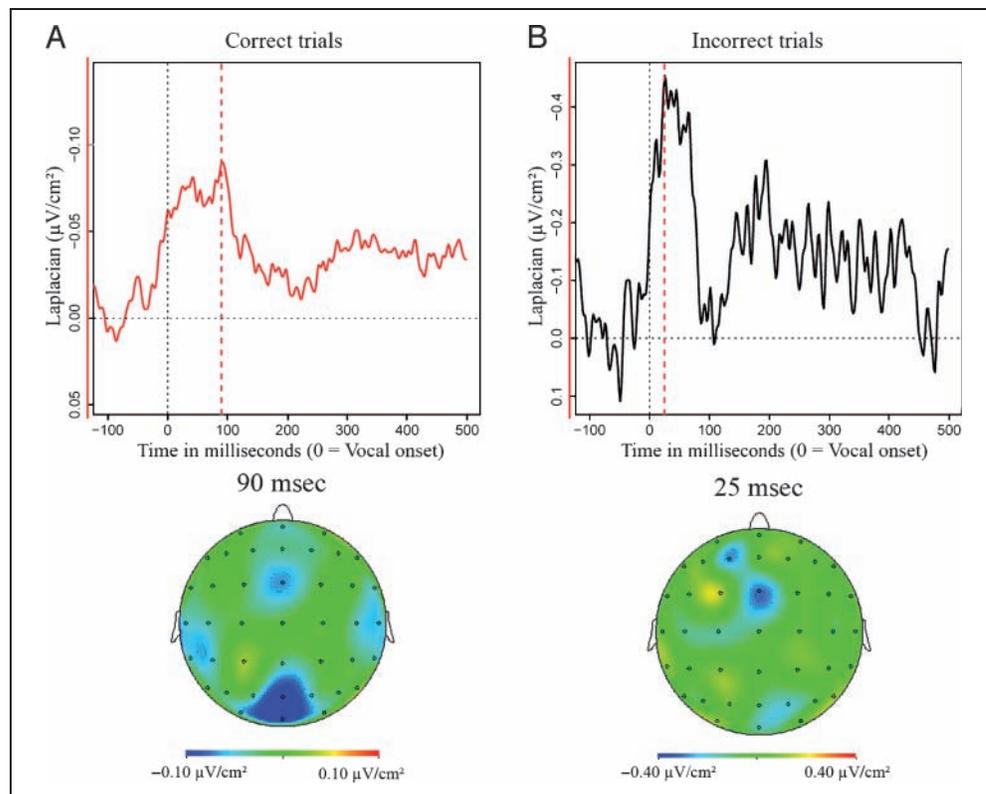


Figure 2. Comparison of the amplitude of the negativity observed on correct trials (x -axis) to that observed on errors (y -axis). Each point represents a participant.

Figure 3. EEG activity after surface Laplacian transformation, recorded over FCz in (A) correct trials and (B) errors. Zero of time corresponds to vocal onset. To see the activities more clearly, a 3-Hz high-pass filter was applied to the EEG data, and the scale of the vertical axis is three times bigger for correct trials. Below the temporal data are the average cartographies corresponding to this activity at the average's maximum of negativity on a 40-msec time window centered around the maximum of the peak of the grand average. The baseline was the 100 msec preceding vocal onset as the activity of interest followed vocal onset. Electrodes P9 and P10 were removed for all participants as they were too noisy for one of the 12 participants.



it is also maximum fronto-centrally, around the electrode FCz. Slopes were different from 0, $T(9) = 6, p = .05$, on the same window as considered for correct trials.

To compare correct trials to incorrect trials, only the nine participants who made more than five errors were considered. On correct trials, the negativity peaked about the same latency, $T(9) = 16, p = .50$, after vocal onset as on errors as well as the preceding positivity (on errors: 46 msec before vocal onset, $\sigma = 52$ msec; on correct trials: 58 msec before vocal onset, $\sigma = 36$ msec; $T(9) = 15, p = .43$; average: 53 msec before vocal onset, $\sigma = 43$ msec). As expected, the amplitude of the negativity on correct trials was weaker than the amplitude of the negativity observed on incorrect trials, $T(9) = 0, p < .01$ (average for correct trials = $-0.12 \mu\text{V}/\text{cm}^2$, $\sigma = .07$; average for errors = $-0.64 \mu\text{V}/\text{cm}^2$, $\sigma = .77$; see Figure 4 for individual data).

In this experiment, on errors, no positivity was visible after the first negativity. Instead, we observed a second negativity peaking at about 200 msec postvocal onset, we will comment this component in the discussion.

Because temporal regions have also been proposed to be involved in speech monitoring (e.g., Indefrey & Levelt, 2004), we investigated possible activities on electrodes above those regions as they could correspond to speech monitoring. A left-lateralized temporal activity at electrode T7 was observed on correct trials (starting on average 48 msec before vocal onset, $\sigma = 37$ msec, and reaching its maximum 216 msec after vocal onset, $\sigma = 85$ msec). On the contra-lateral site, T8, there was a very slow negative slope

but no clear maximum was reached within 500 msec after vocal onset (see Figure 5). Unfortunately, on errors, the EEG data at these electrode sites were too noisy to enable the observation of this type of activity.

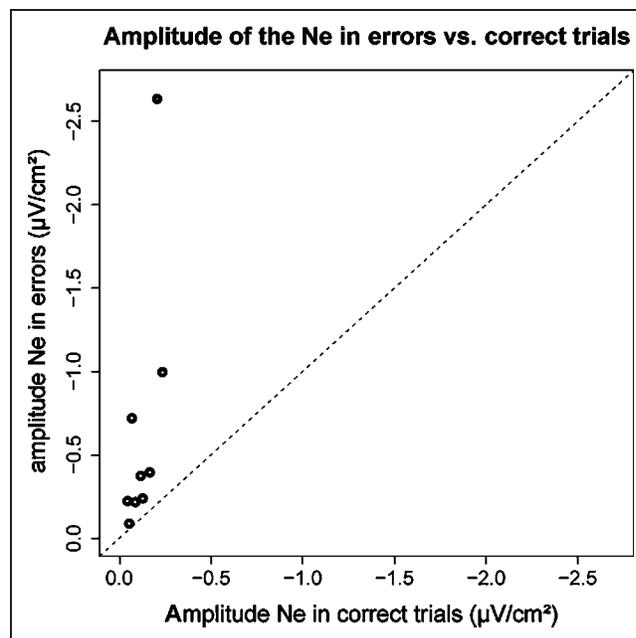
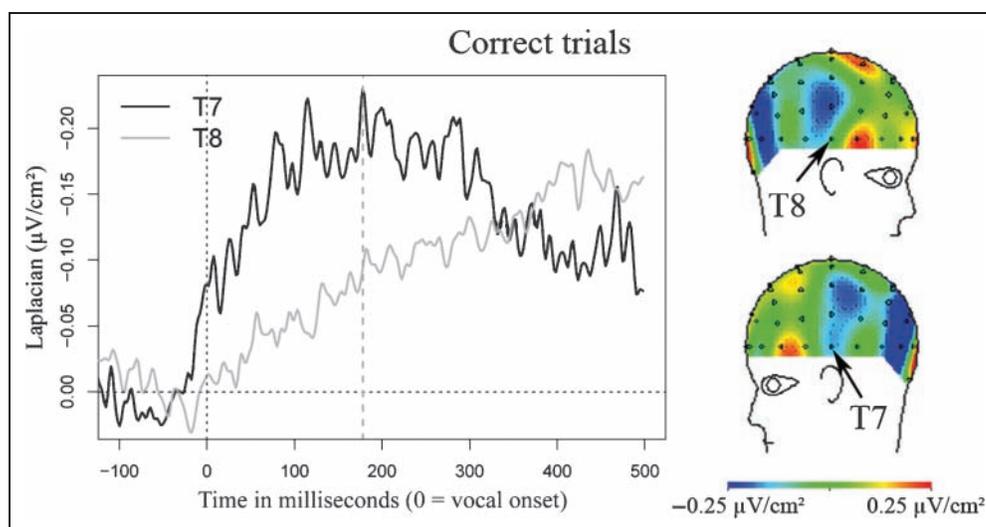


Figure 4. Comparison of the amplitude of the negativity observed on correct trials (x-axis) to that observed on errors (y-axis). Each point represents a participant.

Figure 5. EEG activity after surface Laplacian transformation, recorded over T7 (black line) and T8 (gray line) in correct trials (EEG data were too noisy in errors). Zero of time corresponds to vocal onset. Below the temporal data, the cartographies showing average of the spatial distribution of this activity on a 40-msec time window centered around the maximum of the peak of the grand average (i.e., 270 msec post-vocal onset on errors, 290 msec post-EMG onset, electrodes P7 and P8 removed).



At electrode T7, the slope of the negativity was significantly different from zero on a time window going from 20 msec before vocal onset until 80 msec after response onset (time window corresponding to the part of the curve that had the strongest slope on the grand average), $T(12) = 13, p < .05$. At electrode T8, the slope of the EEG data on the same time window was not significantly different from zero, $T(12) = 24, p = .26$.

DISCUSSION

In line with our main prediction, electrophysiological data revealed the presence of a post-response fronto-central negative wave on both errors and correct trials. The shape, latency, and local topography of this negativity were quite similar for errors and correct trials. Moreover, the local topographies of both negativities were fronto-central (Figures 1 and 3). Surface Laplacian transformation indicated a delimitation of this topography around FCz, which can correspond to the SMA/ACC origin as suggested by Möller et al. (2007) and by Dehaene et al. (1994). Finally, as had previously been observed in nonlinguistic tasks, the negativity on correct trials was smaller than on errors (Bartholow et al., 2005; Vidal et al., 2000, 2003). These observations were made in both experiments, with minor variations across experiments detailed below.

In Experiment 1, the negative wave observed after correct trials peaked earlier than after errors, consistently with nonlinguistic studies (see Roger et al., 2010; Vidal et al., 2000). The difference in amplitude between correct trials and errors was smaller in our study than in previous studies. However, as discussed later on, a rather large error rate could account for this observation. The negative component observed on errors is undoubtedly the well-known Ne seen in numerous nonlinguistic studies (e.g., Vidal et al., 2000; Falkenstein et al., 1991, 1995; Gehring et al., 1993) and in linguistic tasks with manual response modality (e.g.,

Ganushchak & Schiller, 2006, 2008b). Our observations further suggest that the negative component after correct trials is the Ne-like wave reported mainly in nonlinguistic manual tasks (Roger et al., 2010; Bartholow et al., 2005; Vidal et al., 2000, 2003) and incidentally in linguistic tasks (Ganushchak & Schiller, 2009; Sebastian Gallés et al., 2006).

In Experiment 2, the negativity observed on correct trials was clearly smaller than on erroneous trials, matching previous results (see earlier references). The negative wave after incorrect utterances peaked in the same latency range as in Ganushchak and Schiller (2008a) and Masaki et al. (2001), that is, between 50 and 100 msec postvocal onset. The peak of the negativity was reached earlier in Experiment 2 than in Experiment 1; this difference likely stems from the response marker on which the EEG data were averaged (vocal onset in Experiment 2, EMG onset in Experiment 1). In the correct trials of this experiment, we also observed a left temporal activity starting to rise slightly before vocal onset and reaching its maximum about 200 msec after vocal onset (data were too noisy on errors). Here again, the characteristics of the fronto-central negative wave observed on errors correspond to the Ne reported previously in overt speech errors (Ganushchak & Schiller, 2008a; Möller et al., 2007; Masaki et al., 2001). We will therefore refer to this component as the Ne wave. As above, our observations on correct trials strongly suggest that the negative component is the Ne-like wave reported in nonlinguistic manual tasks.

To our knowledge, this is the first observation of an Ne-like wave on correct trials in an overt speech production experiment. We suggest two methodological explanations. First, in covert and overt language studies where the Ne was reported, heavy low-pass filters have been applied on the EEG data (e.g., 1–12 Hz band-pass filters in Ganushchak & Schiller, 2006, 2008a, 2008b; 10 Hz low pass in Masaki et al., 2001). Such severe low-pass filtering is known to distort the EEG signal (Luck, 2005) and especially reduces the amplitude of phasic evoked potentials. It is not

surprising that the Ne observed on errors could still be observed after such filtering because it generally involves a rather large deflection. By contrast, the Ne observed on correct trials, which appears to be smaller, might have been filtered out. The use of such heavy filters was motivated in overt language experiments where contamination of the EEG signal by articulation-induced artifacts was feared. In our study, we did not use such heavy filters because we were able to successfully reduce EMG bursts with the BSS-CCA algorithm. Second, the current source density analysis (Laplacian transformation), acting as a spatial high-pass filter, has been shown to remove the low spatial frequency contribution of remote sources, hence revealing local, higher spatial frequency activities (Vidal et al., 2003).

Different Amplitude Ratios between Ne on Errors and Correct Trials

The Ne amplitude ratio between correct and incorrect trials was relatively small in Experiment 1 and relatively large in Experiment 2. Previous observations have reported large amplitude ratios (Vidal et al., 2000, 2003; Falkenstein, Hoormann, Christ, & Hohnsbein, 2000; Luu, Flaisch, & Tucker, 2000). In nonlinguistic tasks, error rate is one of the parameters that affects the amplitude of the Ne: lower error rates induce larger Ne after incorrect responses (Falkenstein et al., 2000; Gehring et al., 1993). In manual linguistic studies of speech monitoring, the amplitude of the Ne is similarly affected by error rate (Ganushchak & Schiller, 2006, 2009) and by time pressure (lower under time pressure). In the two manual linguistic tasks where an Ne-like wave on correct trials was reported, the amplitude was affected by the error rate and by the certainty of the response (Ganushchak & Schiller, 2009; Sebastian Gallés et al., 2006). In a phoneme-monitoring task, Ganushchak and Schiller (2006) observed a decrease in the amplitude of the Ne (on errors) under severe time pressure, where error rate was higher. Modulations of the Ne-like wave amplitude on correct trials have been less studied. We can speculate that increasing error rates will lead to an amplitude decrease in the Ne on errors and an increase in the Ne-like wave on correct trials (indeed, with 50% of errors in a two alternative choice, errors and correct trials would become indistinguishable). Our two studies are compatible with this tentative proposal because the higher error rate in Experiment 1 (compared with Experiment 2) coincides with an important modulation of the observed amplitude ratio. Such a link, however, deserves a specific investigation.

Post-Ne Activity

On overt errors, the Ne is usually followed by a positivity peaking about 300 msec post-EMG onset. This positivity, known as the Pe, is specific to overt and detected errors

(e.g., Falkenstein et al., 1991; for a review, see Overbeek, Nieuwenhuis, & Ridderinkhof, 2005). The Pe is clearly present in Experiment 1 (see Figure 1), but it is absent in Experiment 2. Instead, we observed a second negativity peaking about 200 msec postvocal onset. Remarkably, a similar second negativity is also visible in the other studies reporting an Ne after overt speech errors (Masaki et al., 2001, Figure 2B; Ganushchak & Schiller, 2008a, Figures 2 and 3), although it is not discussed. This suggests that the second negativity might be a reliable phenomenon.

One tentative interpretation could be that this second negative component is linked to auditory feedback as the well-documented feedback-related negativity (i.e., FRN, e.g., Holroyd & Coles, 2002). Indeed, this component is not observed on incorrect responses of Experiment 1, where no auditory feedback is available. However, this second negativity seems present even when the voice of the subject is masked by pink noise (Masaki et al., 2001), making this hypothesis less likely. A more plausible explanation might be linked to error correction phenomena. In three studies where the Ne is reported in overt speech (Ganushchak & Schiller, 2008a; Masaki et al., 2001; and our study), incorrect trials included both completely incorrect utterances (without any overt self-correction) and partially incorrect utterances followed by an overt self-correction. Masaki et al. (2001) mentioned this explicitly. We can suppose it is also the case in Ganushchak and Schiller's (2008a) study, as it is difficult to imagine that their participants made no such errors (no mention of the removal of trials of this kind, e.g., including two utterances, is made). Notice also that this second negativity is not visible in the Möller et al. (2007) data, where self-corrected trials were not included in the analysis. In a nonlinguistic manual task (i.e., Flanker task), Fielher, Ullsperger, and von Cramon (2005) reported that the Ne is followed by a negativity peaking between 200 and 240 msec in the case of corrected errors. They interpreted this negativity as an error correction process. It is thus possible that the second negativity reported on speech production errors reflect such an error correction process that could mask the Pe. Unfortunately, this account is not fully straightforward either. For example, the grand averages in Figure 3 seem to indicate that the second negativity observed on errors is larger than the one observed on correct trials. It would remain to be seen whether the accuracy of self-corrections modulates this second component. Our data do not allow this detailed analysis.

Determining whether a Pe follows overt speech errors is an important but difficult question. Indeed, as reviewed above, in all three studies that reported the second negativity, a Pe might have been masked by this second negativity. It might also be, however, that the Pe is smaller than that in manual linguistic tasks or even absent in speech production. As a matter of fact, there is a debate regarding the awareness of the detection and self-correction of speech errors (for a review, see Postma, 2000). Because in manual tasks the Pe occurs only on consciously detected errors

(Endrass et al., 2007; O'Connell et al., 2007; Nieuwenhuis et al., 2001), it might not always be present in speech if some errors are not detected consciously. This would lead to its amplitude reduction on the grand averages.

In Experiment 2, we also observed a slowly rising negativity over the left temporal cortex. On the contra-lateral site, the activity was virtually absent. Similarly to the Ne and Ne-like wave, this activity started to rise slightly before vocal onset. This activity could only be studied on correct trials because, on errors, the data were too noisy to assess its existence. We cannot therefore decide whether this component is specific to correct trials. However, the fact that this negativity was clearly observed in the overt picture-naming task but not in the gender decision task is of interest. Indeed, the presence of this component could be linked to the production of overt speech. A vocalization-related cortical potential has been described (e.g., Wohlert, 1993; Schafer, 1967). However, this component starts to rise long before vocal onset and its generators (SMA and bilateral laryngeal motor areas; Galgano & Froud, 2006) have been identified in regions somewhat incompatible with the topography we reported. An alternative interpretation of this left-lateralized component is to relate it to overt speech monitoring, as postulated in the outer-loop component of speech monitoring models (Postma, 2000). The topography of this component is in agreement with the brain regions identified in fMRI studies of speech monitoring manipulating auditory feedback (see Introduction). The fact that this negativity starts to rise before the actual vocal onset (20 msec before), however, moderates this interpretation. It is possible that this component reflects monitoring of both inner representations (e.g., phonological) and outer speech. However, we are not aware of previous EEG study having reported this type of component in overt speech. Possible reasons being again of a methodological order (use of BSS-CCA and Laplacian transformation in our study but not in other language production EEG studies). Because we did not assess what parameters could affect this component, we cannot be certain of the interpretation proposed, and further research will be needed to clarify this finding.

Implications of Our Observations

Experiment 1 extended to a linguistic task the observation of an Ne-like on correct trials (for previous incidental findings, see Ganushchak & Schiller, 2009; Sebastian Gallés et al., 2006). More importantly, Experiment 2 extended this observation to overt speech production. In the two experiments, a fronto-medial negative component after both correct and incorrect trials was observed, at similar latencies, with similar local topography and with variable amplitudes. As suggested for nonlinguistic tasks (Roger et al., 2010; Vidal et al., 2000), this can be taken as an indication that the Ne observed in linguistic tasks reflects a monitoring mechanism rather than an error-detection phenomenon.

The Ne observed in speech production therefore appears to be a valuable tool to study speech monitoring.

The similarity between our findings and previous observations suggests that the mechanism reflected by the Ne is the same in linguistic tasks, which require access to long-term memory, and in nonlinguistic tasks, which often involve more arbitrary stimulus-response associations. This means that the monitoring process reflected by the Ne can operate over representations involving access to different types of memory. This interpretation is in agreement with the prediction stated in the Introduction that at least part of the monitoring involved in speech may not be specific to speech but may instead be shared across different cognitive functions.

Our data do not rule out that other brain regions may be implied in speech monitoring. As a matter of fact, we also reported activity in the left-temporal area in the overt speech task that seems to be specific to this task. The medial frontal region as well as the temporal cortices has been associated with verbal self-monitoring in several fMRI studies (see Introduction). The Ne has been shown to stem from ACC/SMA region (Roger et al., 2010; Debener et al., 2005; Van de Veen & Carter, 2002; Dehaene et al., 1994) but certainly not from the temporal cortices. Moreover, activity in the temporal cortices has not been associated to monitoring outside linguistic studies. Our results would thus support the hypothesis that speech monitoring comprises two components: aspecific processes, stemming from the medial-frontal cortex (reflected by the Ne and the Ne-like), which are shared with general action monitoring, and language-specific processes originating from the left temporal cortex.

A critical issue in the study of speech monitoring is to determine when monitoring takes place. As stated in the Introduction, most theories of speech monitoring assume the involvement of both an inner and an outer loop. The inner loop can detect errors before the utterance is actually articulated, as it relies on inner representations. The outer (or auditory) loop goes through the auditory system as it depends on auditory feedback to detect errors (e.g., Postma & Oomen, 2005). In their meta-analysis of brain imaging studies, Indefrey and Levelt (2004) proposed the left temporal cortex to be the cerebral basis of external speech monitoring. They argue that the most economical assumption would be that the same brain region would subserve inner speech monitoring as well. As described in Introduction, most functional imaging studies (fMRI and PET) have focused on the monitoring of verbal auditory feedback and do not provide strong arguments for distinguishing inner and outer speech monitoring.

Electrophysiological evidence may be more appropriate to settle this issue. Masaki et al. (2001) masked the voice of their participants by a loud pink noise to avoid the contamination of their data by voice-related cortical potentials. They observed incorrect partial utterances followed by corrections, despite the fact that participants could not hear themselves, suggesting that inner loop monitoring

was operational. Notably, they observed an Ne (on errors), suggesting that this component is relatively independent from auditory feedback (i.e., outer loop).

The results of our study, along with data available in the literature, support the proposal that the Ne might reflect the inner loop. First, in the grammatical decision task, “speech production” monitoring occurs without verbal auditory feedback, and an Ne was observed in both errors and correct trials. Second, in the experiment involving speech production proper, the Ne starts before participants can actually hear themselves. Indeed, the Ne starts before vocal onset (53 msec before vocal onset, $\sigma = 43$ msec) and reaches its maximum about 40 msec postvocal onset, whereas the first cortical auditory response is known to arise around 50 msec after auditory presentation. This is in agreement with findings showing that, in manual response tasks, the Ne (on both error and correct trials) is generated in the absence of any sensory reafferences (Allain et al., 2004). Although the left temporal cortex activity also starts before vocal onset, it peaks much later than the Ne (around 200 msec). It is therefore possible that its later part reflects feedback processing. In this case, this component might be a good candidate for the outer-loop monitoring system. This point however needs further investigation. In terms of previously proposed theories of speech monitoring, our

findings suggest that the Ne reflects a mechanism tied to internal rather than external monitoring. More generally, this interpretation suggests that the Ne reflects monitoring of internal, preexecution representations of a linguistic or nonlinguistic nature, whereas postresponse left temporal cortex activity may reflect, at least in part, the outer-loop monitoring process.

Conclusions

On the basis of the available evidence concerning the involvement of medial frontal areas in monitoring processes as well as the reflection of these processes in the Ne component, we predicted that comparable negativities should be observed on correct and incorrect trials in linguistic tasks. We report two experiments in which this hypothesis was confirmed. We suggest that the monitoring processes reflected by the Ne are shared across different cognitive functions. This implies that a substantial part of the monitoring system involved in speech is not inherent to the language system. We further argued that this multimodal mechanism must be engaged before overt responses are made (i.e., before auditory feed-back is available). This is in line with the prearticulatory inner loop postulated in previously proposed theories of speech monitoring.

APPENDIX

Materials and Design

<i>Picture Name (French)</i>	<i>Picture Name (English)</i>	<i>Frequency</i>	<i>Number of Syllables</i>	<i>Gender (F = Feminine; M = Masculine)</i>
<i>Experiment 1</i>				
botte	boot	6.38	1	F
brosse	brush	7.29	1	F
chaise	chair	32.7	1	F
douche	shower	32.56	1	F
flèche	arrow	8.21	1	F
fraise	strawberry	5.28	1	F
lampe	lamp	22.22	1	F
lettre	letter	108.79	1	F
malle	chest	6.53	1	F
moufle	mitten	0.28	1	F
niche	doghouse	2.17	1	F
palme	flipper	1.25	1	F
pelle	shovel	8.75	1	F
poire	pear	5.67	1	F
pomme	apple	19.77	1	F
porte	door	288.39	1	F
robe	dress	72.72	1	F

APPENDIX *(continued)*

<i>Picture Name (French)</i>	<i>Picture Name (English)</i>	<i>Frequency</i>	<i>Number of Syllables</i>	<i>Gender (F = Feminine; M = Masculine)</i>
tasse	cup	18.52	1	F
tente	tent	14.4	1	F
trousse	pencil-case	3.77	1	F
baignoire	bathtub	11.9	2	F
batterie	drum-set	10.61	2	F
bouilloire	water-boiler	1.52	2	F
brouette	wheelbarrow	1.1	2	F
casquette	cap	8.64	2	F
casserole	pot	2.91	2	F
chaussette	sock	3.29	2	F
chaussure	shoe	12.49	2	F
chemise	shirt	36.48	2	F
commode	dresser	3.21	2	F
cravate	tie	15.99	2	F
cuillère	spoon	5.18	2	F
fenêtre	window	70.2	2	F
fourchette	fork	4.98	2	F
guitare	guitar	12.78	2	F
moto	motorcycle	22.61	2	F
passoire	colander	1.27	2	F
perceuse	drill	0.97	2	F
sandale	sandal	0.22	2	F
serrure	door-lock	7.4	2	F
serviette	towel	17.16	2	F
souris	mouse (computer)	21.94	2	F
théière	teapot	0.8	2	F
tomate	tomato	7.88	2	F
tondeuse	lawnmower	1.72	2	F
trompette	trumpet	5.71	2	F
valise	suitcase	33.21	2	F
voiture	car	388.87	2	F
cafetière	coffeepot	1.92	3	F
caméra	video-recorder	41.64	3	F
gazinière	stove	0.01	3	F
mitrailleuse	machine-gun	2.62	3	F
télévision	television	25.45	4	F
banc	bench	8.96	1	M
bus	bus	50.63	1	M
cor	french horn	2.57	1	M

APPENDIX *(continued)*

<i>Picture Name (French)</i>	<i>Picture Name (English)</i>	<i>Frequency</i>	<i>Number of Syllables</i>	<i>Gender (F = Feminine; M = Masculine)</i>
four	oven	13.95	1	M
lit	bed	176.1	1	M
livre	book	112.43	1	M
masque	mask	23.16	1	M
phare	lighthouse	7.17	1	M
plat	dish	21.87	1	M
seau	bucket	7.02	1	M
tank	tank	3.8	1	M
toit	roof	42.63	1	M
train	train	244.4	1	M
vase	vase	9.83	1	M
balcon	balcony	9.9	2	M
bateau	boat	106.55	2	M
biberon	nursing-bottle	5.25	2	M
bonnet	beanie	6.62	2	M
bureau	desk	156.68	2	M
camion	truck	50.06	2	M
canon	canon	14.89	2	M
carton	box	10.92	2	M
chapeau	hat	48.61	2	M
château	castle	40.51	2	M
citron	lemon	8.1	2	M
clavier	keyboard	1.62	2	M
colis	package	7.5	2	M
couteau	knife	51.08	2	M
kiwi	kiwi	0	2	M
maison	house	570.3	2	M
manteau	coat	36.16	2	M
micro	microphone	11.25	2	M
miroir	mirror	24.89	2	M
moulin	windmill	6.8	2	M
panier	basket	13.82	2	M
piano	piano	21.5	2	M
rasoir	razor	8.18	2	M
râteau	rake	0.77	2	M
roller	rollerblade	0.88	2	M
sabot	wooden shoe	1.79	2	M
tampon	stamp	2.96	2	M
tiroir	drawer	12.18	2	M

APPENDIX (continued)

<i>Picture Name (French)</i>	<i>Picture Name (English)</i>	<i>Frequency</i>	<i>Number of Syllables</i>	<i>Gender (F = Feminine; M = Masculine)</i>
T-shirt	T-shirt	7.59	2	M
violon	violin	11.56	2	M
bouclier	shield	10.47	3	M
canapé	sofa	17.66	3	M
pantalon	pants	31.49	3	M
pistolet	revolver	31.63	3	M
robinet	faucet	5.12	3	M
saladier	salad-bowl	0.68	3	M
sécateur	clippers	0.15	3	M
téléphone	telephone	155.68	3	M
calendrier	calender	5.37	4	M
ventilateur	ventilator	2.43	4	M
réfrigérateur	refrigerator	2.94	5	M
<i>Experiment 2</i>				
bague	ring	26.14	1	F
fraise	strawberry	5.28	1	F
hache	ax	8.73	1	F
lampe	lamp	22.22	1	F
luge	sled	1.19	1	F
moufle	mitten	0.28	1	F
niche	doghouse	2.17	1	F
robe	dress	72.72	1	F
scie	saw	5.21	1	F
vis	screw	6.89	1	F
baleine	whale	11.52	2	F
chenille	caterpillar	1.38	2	F
église	church	60.2	2	F
fenêtre	window	70.2	2	F
girouette	windvane	0.47	2	F
hélice	propeller	1.9	2	F
lunettes	glasses	31.61	2	F
maison	house	570.3	2	F
montagne	mountain	36.76	2	F
voiture	car	388.87	2	F
balançoire	swing	1.93	3	F
arbre	tree	49.29	1	M
bras	arm	149.26	1	M
clown	clown	0.06	1	M

APPENDIX (continued)

Picture Name (French)	Picture Name (English)	Frequency	Number of Syllables	Gender (F = Feminine; M = Masculine)
coeur	heart	224.98	1	M
livre	book	112.43	1	M
nuage	cloud	11.81	1	M
piéd	foot	105.51	1	M
train	train	244.4	1	M
verre	glass	154.13	1	M
bureau	desk	156.68	2	M
chapeau	hat	48.61	2	M
moulin	windmill	6.8	2	M
poivron	pepper	0.51	2	M
râteau	rake	0.77	2	M
soleil	sun	120.72	2	M
tracteur	tractor	2.87	2	M
violon	violin	11.56	2	M
ananas	pineapple	2.02	3	M
artichaut	artichoke	1.45	3	M
champignon	mushroom	3.34	3	M
entonnoir	funnel	0.81	3	M
microscope	microscope	3.71	3	M
papillon	butterfly	8.12	3	M
robinet	faucet	5.12	3	M

Acknowledgments

This research was supported by grant ANR-07-JCJC-0074 from the Agence Nationale de la Recherche (France), by the European Research Council under the European Community's Seventh Framework Program (FP7/2007-2013, grant 241077) and by an MNRT grant from the Ministère de l'Enseignement et de la Recherche (France) awarded to Stéphanie Riès. The authors thank Frank Vidal for helpful discussions and comments, Bruno Nazarian for his help with the piezoelectric microphone, and an anonymous reviewer for helpful comments.

Reprint requests should be sent to Stéphanie Riès, Laboratoire de Psychologie Cognitive (LPC), UMR 6146, Pôle 3C-Case D, Centre St Charles, 3, Place Victor Hugo, 13331 Marseille Cedex 03, or via e-mail: stephanie.ries@univ-provence.fr.

Note

1. For a different purpose related to an investigation of pre-response processing, we manipulated the semantic context in which the stimuli were presented: The experimental items belonged to 18 semantic categories (e.g., pieces of furniture, fruits, clothing, etc.). Pictures were presented in semantically related or unrelated groups. Grammatical gender (i.e., masculine or feminine) was equally distributed across these groups. Importantly, our semantic context manipulation was different than the one performed by

Ganushchak and Schiller (2006, 2008a, 2008b). In our study, the amplitude of the Ne was not different between conditions.

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*, 531–552.
- Allain, S., Hasbroucq, T., Burle, B., Grapperon, J., & Vidal, F. (2004). Response monitoring without sensory feedback. *Clinical Neurophysiology*, *115*, 2014–2020.
- Babiloni, F., Cincotti, F., Carducci, F., Rossini, P. M., & Babiloni, C. (2001). Spatial enhancement of EEG data by surface Laplacian estimation: The use of magnetic resonance imaging-based head models. *Clinical Neurophysiology*, *112*, 724–727.
- Barch, D. M., Braver, T. S., Sabb, F. W., & Noll, D. C. (2000). Anterior cingulate and the monitoring of response conflict: Evidence from an fMRI study of overt verb generation. *Journal of Cognitive Neuroscience*, *12*, 309.
- Bartholow, B. D., Pearson, M. A., Dickter, C. L., Sher, K. J., Fabiani, M., & Gratton, G. (2005). Strategic control and

- medial frontal negativity: Beyond errors and response conflict. *Psychophysiology*, *42*, 33–42.
- Bonin, P., Peereman, R., Malardier, N., Mot, A., & Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments, and Computers*, *35*, 167.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652.
- Burle, B., Possamaï, C. A., Vidal, F., Bonnet, M., & Hasbroucq, T. (2002). Executive control in the Simon effect: An electromyographic and distributional analysis. *Psychological Research*, *66*, 324–336.
- Burle, B., Roger, C., Allain, S., Vidal, F., & Hasbroucq, T. (2008). Error negativity does not reflect conflict: A reappraisal of conflict monitoring and anterior cingulate cortex activity. *Journal of Cognitive Neuroscience*, *20*, 1655.
- Carbonnell, L., Hasbroucq, T., Grapperon, J., & Vidal, F. (2004). Response selection and motor areas: A behavioral and electrophysiological study. *Clinical Neurophysiology*, *115*, 2164–2174.
- Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, *280*, 747–749.
- Christoffels, I. K., Formisano, E., & Schiller, N. O. (2007). Neural correlates of verbal feedback processing: An fMRI study employing overt speech. *Human Brain Mapping*, *28*, 868–879.
- Corbett, G. G. (1991). *Gender*. Cambridge, MA: Cambridge University Press.
- 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., Ries, S., Vanderperren, K., Vanrumste, B., Alario, F.-X., Van Huffel, S., et al. (2010). Removal of muscle artifacts from EEG recordings of spoken language production. *Neuroinformatics*, *8*, 135–150.
- Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, Y., & Engel, A. K. (2005). Trial-by-trial coupling of concurrent EEG and fMRI identifies the dynamics of performance monitoring. *Journal of Neuroscience*, *25*, 11730–11737.
- Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science*, *5*, 303–305.
- Endrass, T., Reuter, B., & Kathmann, N. (2007). ERP correlates of conscious error recognition: Aware and unaware errors in an antisaccade task. *European Journal of Neuroscience*, *26*, 1714–1720.
- Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components: II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, *78*, 447–455.
- Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: A tutorial. *Biological Psychology*, *51*, 87–107.
- Falkenstein, M., Koshlykova, N. A., Kiroj, V. N., & Hoormann, J. (1995). Late ERP components in visual and auditory go/nogo tasks. *Electroencephalography and Clinical Neurophysiology: Evoked Potentials*, *96*, 36–43.
- Fielher, K., Ullsperger, M., & von Cramon, D. Y. (2005). Electrophysiological correlates of error correction. *Psychophysiology*, *42*, 72–82.
- Fu, C. H., Vythelingum, G. N., Brammer, M. J., Williams, S. C., Amaro, E., Andrew, C. M., et al. (2006). An fMRI study of verbal self-monitoring: Neural correlates of auditory verbal feedback. *Cerebral Cortex*, *16*, 969–977.
- Galgano, J., & Froud, K. (2006). Evidence of the voice-related cortical potential: An electroencephalographic study. *Neuroimage*, *41*, 1313–1323.
- Ganushchak, L. Y., & Schiller, N. O. (2006). Effects of time pressure on verbal self-monitoring: An ERP study. *Brain Research*, *1125*, 104–115.
- Ganushchak, L. Y., & Schiller, N. O. (2008a). Motivation and semantic context affect brain error-monitoring activity: An event-related brain potentials study. *Neuroimage*, *39*, 395–405.
- Ganushchak, L. Y., & Schiller, N. O. (2008b). Brain error-monitoring activity is affected by semantic relatedness: An event-related brain potentials study. *Journal of Cognitive Neuroscience*, *20*, 927–940.
- Ganushchak, L. Y., & Schiller, N. O. (2009). Speaking one's second language under time pressure: An ERP study on verbal self-monitoring in German-Dutch bilinguals. *Psychophysiology*, *46*, 410–419.
- Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*, 385–390.
- Gevens, A., Leong, H., Smith, M. E., Le, J., & Du, R. (1995). Mapping cognitive brain function with modern high-resolution electroencephalography. *Trends in Neurosciences*, *18*, 429–436.
- Goncharova, I. I., McFarland, D. J., Vaughan, T. M., & Wolpaw, J. R. (2003). EMG contamination of EEG: Spectral and topographical characteristics. *Clinical Neurophysiology*, *114*, 1580–1593.
- Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifacts. *Electroencephalography and Clinical Neurophysiology*, *55*, 468–484.
- Hashimoto, Y., & Sakai, K. L. (2003). Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: An fMRI study. *Human Brain Mapping*, *20*, 22–28.
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, *4*, 131–138.
- Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*, 679–709.
- Indefrey, P., & Levelt, W. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*, 101–144.
- Levelt, W. J. M. (1989). *Speaking, from intention to articulation*. Cambridge, MA: MIT Press.
- Luck, S. L. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.
- Luu, P., Flaisch, T., & Tucker, D. M. (2000). Medial frontal cortex in action monitoring. *Journal of Neuroscience*, *20*, 464–469.
- Masaki, H., Tanaka, H., Takasawa, N., & Yamazaki, K. (2001). Error-related brain potentials elicited by vocal errors. *Neuroreport*, *12*, 1851–1855.
- McGuire, P. K., Silbersweig, D. A., & Frith, C. D. (1996). Functional neuroanatomy of verbal self-monitoring. *Brain*, *119*, 907–917.

- Möller, J., Jansma, B. M., Rodriguez-Fornells, A., & Münte, T. F. (2007). What the brain does before the tongue slips. *Cerebral Cortex*, *17*, 1173–1178.
- Navarette, E., Basagni, B., Alario, F.-X., & Costa, A. (2006). Does word frequency affect lexical selection in speech production? *Quarterly Journal of Experimental Psychology*, *59*, 1681–1690.
- Nieuwenhuis, S., Ridderinkhof, R. K., Blom, J., Band, G. P. H., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology*, *38*, 752–760.
- Núñez, P. (1981). *Electric fields of the brain*. Oxford, UK: Oxford University Press.
- O’Connell, R. G., Dockree, P. M., Bellgrove, M. A., Kelly, S. P., Hester, R., Garavan, H., et al. (2007). The role of cingulate cortex in the detection of errors with and without awareness: A high-density electrical mapping study. *European Journal of Neuroscience*, *25*, 2571–2579.
- Overbeek, T. J. M., Nieuwenhuis, S., & Ridderinkhof, K. R. (2005). Dissociable components of error processing: On the functional significance of the Pe vis-à-vis the ERN/Ne. *Journal of Psychophysiology*, *19*, 319–329.
- Perrin, F., Bertrand, O., & Pernier, J. (1987). Scalp current density mapping: Value and estimation from potential data. *IEEE Transactions on Biomedical Engineering*, *34*, 283–288.
- Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition*, *77*, 97–131.
- Postma, A., & Oomen, C. E. (2005). Critical issues in speech monitoring. In R. J. Hartsuiker, R. Bastiaanse, A. Postma, & F. Wijnen (Eds.), *Phonological encoding and monitoring in normal and pathological speech* (pp. 157–186). New York: Psychology Press.
- Protopapas, A. (2005). CheckVocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*, 859–862.
- Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*, 443–447.
- Roger, C., Bénar, C. G., Vidal, F., Hasbroucq, T., & Burle, B. (2010). Rostral Cingulate Zone and correct response monitoring: ICA and source localization evidences for the unicity of correct- and error-negativities. *Neuroimage*, *51*, 391–403.
- Schafer, E. W. P. (1967). Cortical activity preceding speech: Semantic specificity. *Nature*, *216*, 1338–1339.
- Sebastian Gallés, N., Rodríguez-Fornells, A., de Diego-Balaguer, R., & Díaz, B. (2006). First- and second-language phonological representations in the mental lexicon. *Journal of Cognitive Neuroscience*, *18*, 1277–1291.
- Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.
- Staude, G., Flachenecker, C., Daumer, M., & Wolf, W. (2001). Onset detection in surface electromyographic signals: A systematic comparison of methods. *EURASIP Journal on Applied Signal Processing*, *2*, 67–81.
- Stevens, M., Lammertyn, J., Verbruggen, F., & Vandierendonck, A. (2006). Tscope: A C library for programming cognitive experiments on the MS Windows platform. *Behavior Research Methods*, *38*, 280–286.
- Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage*, *39*, 1429–1443.
- Van Boxtel, G. J. M., Geraats, L. H. D., van den Berg-Lenssen, M. M. C., & Brunia, C. H. M. (1993). Detection of EMG onset in ERP research. *Psychophysiology*, *30*, 405–412.
- Van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods*, *38*, 584–589.
- Van de Veen, V., & Carter, C. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience*, *14*, 593–602.
- Van de Ven, V., Esposito, F., & Christoffels, I. K. (2009). Neural network of speech monitoring overlaps with overt speech production and comprehension networks: A sequential spatial and temporal ICA study. *Neuroimage*, *47*, 1982–1991.
- van Turennout, M., Hagoort, P., & Brown, C. M. (1998). Brain activity during speaking: From syntax to phonology in 40 milliseconds. *Science*, *280*, 572–574.
- Vanderperren, K., Ries, S., De Vos, M., Vanrumste, B., Alario, F.-X., Burle, B., et al. (2008, April). Removal of muscle artifacts from eeg recordings of spoken language production. In *Proceedings of the Neuromath Workshop, Advanced Methods for the Estimation of Human Brain Activity and Connectivity* (pp. 1–2), Jena, Germany.
- Vidal, F., Burle, B., Bonnet, M., Grapperon, J., & Hasbroucq, T. (2003). Error negativity on correct trials: A reexamination of available data. *Biological Psychology*, *64*, 265–282.
- Vidal, F., Hasbroucq, T., Grapperon, J., & Bonnet, M. (2000). Is the “error negativity” specific to errors? *Biological Psychology*, *51*, 109–128.
- Wohlert, A. B. (1993). Event-related brain potentials preceding speech and non-speech oral movements of varying complexity. *Journal of Speech and Hearing Research*, *36*, 897–905.
- Yordanova, J., Falkenstein, M., Hohnsbein, J., & Kolev, V. (2004). Parallel systems of error processing in the brain. *Neuroimage*, *22*, 590–602.