Speaking one's second language under time pressure: An ERP study on verbal self-monitoring in German-Dutch bilinguals

Schiller, N.O.

Psychophysiology, 46 (2009), 410–419. Wiley Periodicals, Inc. Printed in the USA.

DOI: 10.1111/j.1469-8986.2008.00774.x

Speaking one's second language under time pressure: An ERP study on verbal self-monitoring in German–Dutch bilinguals

LESYA Y. GANUSHCHAK[a] AND NIELS O. SCHILLER[a,b]

[a]Department of Cognitive Psychology, Leiden Institute for Brain and Cognition, Leiden University, Leiden, The Netherlands
[b]Department of Cognitive Neuroscience, Faculty of Psychology, Maastricht University, Maastricht, The Netherlands

Abstract

This study addresses how verbal self-monitoring and the Error-Related Negativity (ERN) are affected by time pressure when a task is performed in a second language as opposed to performance in the native language. German–Dutch bilinguals were required to perform a phoneme-monitoring task in Dutch with and without a time pressure manipulation. We obtained an ERN following verbal errors that showed an atypical increase in amplitude under time pressure. This finding is taken to suggest that under time pressure participants had more interference from their native language, which in turn led to a greater response conflict and thus enhancement of the amplitude of the ERN. This result demonstrates once more that the ERN is sensitive to psycholinguistic manipulations and suggests that the functioning of the verbal self-monitoring system during speaking is comparable to other performance monitoring, such as action monitoring.

Descriptors: Speech production, Verbal self-monitoring, Phoneme monitoring, ERN, Time pressure, Bilingualism

Everyday life cannot be imagined to take place in the absence of errors. Errors are often the basis for new strategies, learning, and adaptation. Therefore, a major part of human performance monitoring research is dedicated to error processing. The neural basis of error monitoring has become a key issue in cognitive neuroscience. An interesting component of the event-related potential (ERP) for exploring the functional characteristics of the error monitoring system is the *Error-Related Negativity* (ERN; Falkenstein, Hohnsbein, Hoorman, & Blanke, 1991; Gehring, Goss, Coles, Meyer, & Donchin, 1993). The ERN has a fronto-central scalp distribution and peaks about 80 ms after an overt incorrect response (Bernstein, Scheffers, & Coles, 1995; Holroyd & Yeung, 2003; Scheffers, Coles, Bernstein, Gehring, & Donchin, 1996).

Originally, the ERN was thought to arise as a result of *error detection* (Bernstein et al., 1995). This hypothesis assumes a comparison between the internal representation of the intended correct response, arising from ongoing stimulus processing, and the internal representation of the actual response, resulting from the efferent copy of the motor activity. If there is a mismatch between these two representations, then an ERN will be generated (Bernstein et al., 1995; Falkenstein, Hoormann, Christ, & Hohnsbein, 2000; Holroyd & Coles, 2002).

This view has been challenged by the *conflict hypothesis*, according to which the ERN reflects detection of response conflict and not detection of errors per se (Botvinick, Braver, Barch, Carter, & Cohen, 2001). Response conflict arises when multiple responses compete for selection. Presence of conflicting responses reflects situations where errors are likely to occur. Thus, according to the conflict hypothesis, error detection is not an independent process but based on the presence of response conflict.

Alternatively, the *reinforcement-learning theory* proposed that the ERN may reflect a negative reward-prediction error signal that is elicited when the monitor detects that the consequences of an action are worse than expected. This reward-prediction error signal is coded by the mesencephalic dopamine system and projected to the anterior cingulated cortex (ACC), where the ERN is elicited (Holroyd & Coles, 2002).

A large set of studies on the ERN investigated the functioning of action monitoring. According to the *action monitoring model*, the action monitor is a feed-forward control mechanism that is used to inhibit and correct a faulty response (Desmurget & Grafton, 2000; Rodríguez-Fornells, Kurzbuch, & Münte, 2002). When the wrong selection of the motor command is generated, a copy of an online response is produced and compared to the representation of the correct response. If there is a mismatch between the copy of the online response and the representation of the correct response, an error signal is generated and a stop command is initiated (Coles, Scheffers, & Holroyd, 2001).

The work presented in this manuscript is supported by NWO grant no. 453-02-006 to Niels O. Schiller.

Address reprint requests to: Lesya Ganushchak, Department of Cognitive Psychology, Leiden University–Institute for Psychological Research (LU-IPR) Wassenaarseweg 52, 2333 AK Leiden, The Netherlands. E-mail: lganushchak@fsw.leidenuniv.nl

If the ERN is associated with error processing in action monitoring, can it also be applied to error processing in *verbal monitoring*? Verbal self-monitoring is a crucial part of speech production, especially when one considers that producing speech errors hampers the fluency of speech and can sometimes lead to embarrassment, for instance when taboo words are uttered unintentionally (Motley, Camden, & Baars, 1982). One prominent theory of verbal self-monitoring is the *perceptual-loop theory* proposed by Levelt (1983, 1989). According to this theory, a speech monitoring system checks the intended message for its appropriateness, inspects the speech plan and detects errors prior to its articulation (Postma & Noordanus, 1996; Schiller, 2005, 2006; Schiller, Jansma, Peters, & Levelt, 2006; Wheeldon & Levelt, 1995; Wheeldon & Morgan, 2002), as well as after the speech has become overt (Postma, 2000). Verbal monitoring is achieved via the speech comprehension system.

Previous studies showed that an ERN can also be elicited by verbal errors (e.g., Ganushchak & Schiller, 2006, 2008; Masaki, Tanaka, Takasawa, & Yamazaki, 2001; Möller, Jansma, Rodríguez-Fornells, & Münte, 2007; Sebastián-Gallés, Rodríguez-Fornells, De Diego-Balaquer, & Díaz, 2006). Importantly, Sebastián-Gallés and colleagues demonstrated the ERN in a bilingual situation. These authors showed that Spanish-dominant bilinguals taking part in a Catalan auditory lexical decision task had great difficulty rejecting nonwords that were phonologically similar to existing Catalan (i.e., their L2) words and did not show an ERN in their erroneous nonword decisions. According to Sebastián-Gallés et al. this suggests that Spanish-dominant bilinguals were unable to distinguish between experimental words and nonwords and therefore exhibited no difference between correct and incorrect responses. In contrast, Catalan-dominant bilinguals showed a clear ERN.

In the present study, we investigated the relationship between the ERN and verbal monitoring in a nonnative language. Nowadays, bilingualism is the rule rather than an exception, certainly in large parts of Europe with its multilingual societies. However, very little is known about monitoring of one's speech in a second language. Increased knowledge about the error monitoring system in monolingual and bilingual speech production may improve our understanding of some disorders where verbal monitoring is implicated, such as aphasia (for an overview, see Oomen, Postma, & Kolk, 2001), stuttering (e.g., Lickley, Hartsuiker, Corley, Russell, & Nelson, 2005), and schizophrenia (for an overview, see Seal, Aleman, & McGuire, 2004).

The present work follows up on earlier work by Ganushchak and Schiller (2006). These authors addressed the questions of whether or not an ERN occurs after verbal error detection and whether a potential ERN is affected by a time pressure manipulation. They employed a phoneme monitoring go/no-go task, previously used in language production and verbal monitoring research (e.g., Schiller, 2005; Wheeldon & Levelt, 1995; Wheeldon & Morgan, 2002). In the particular task employed by Ganushchak and Schiller (2006), participants were required to internally name pictures and press a button if a particular target phoneme was present in the name of the picture. For example, if the target phoneme was /b/ and the target picture was *bear*, then participants were required to press a corresponding button. Thus, participants were asked to monitor their own internal speech production. Ganushchak and Schiller (2006) successfully obtained an ERN following verbal errors and showed a typical decrease in amplitude under conditions of time pressure. The authors suggested that the functioning of the verbal monitor is comparable to other performance monitoring, such as action monitoring.

In the present study, we used the identical setup of the experiment described in Ganushchak and Schiller (2006). However, participants in the current study were German–Dutch bilinguals who were asked to perform a phoneme monitoring task in their second language, that is, Dutch. The main question addressed in the current study was the following: How is the ERN affected by time pressure when a verbal monitoring task is performed in a second language? The reason for using the time pressure manipulation is twofold. First, the present study follows up on an earlier study by Ganushchak and Schiller (2006), and it was important to keep setups as similar as possible between these two studies because the earlier study will serve as a monolingual control group for the current study. More importantly, however, the ERN has been scarcely used before to evaluate cognitive performance in bilinguals. Therefore, it is important to use a manipulation that has been employed in the ERN as well as the verbal monitoring literature, and time pressure is such a manipulation. Throughout the action monitoring literature, it has consistently been reported that the amplitude of the ERN decreased when time pressure was increased (Falkenstein et al., 2000; Gehring et al., 1993). Increasing time pressure has also implications for verbal monitoring; speech became more error prone and less fluent with increased speech rate (Oomen & Postma, 2001; Postma, 2000). There are also indications that the ERN is decreased under time pressure in the verbal monitoring task (Ganushchak & Schiller, 2006).

In the existing literature, there is no evidence suggesting that the ERN would be affected differently in a second language. Moreover, there is increasing evidence that native and foreign languages are based on the same neural substrate (e.g., Klein, Milner, Zatorre, Meyer, & Evans, 1995; Klein, Milner, Zatorre, Zhao, & Nikelski, 1999; Perani et al., 1998; but see Lucas, McKhann, & Ojemann, 2004). Furthermore, in the existing literature there is evidence that functional separation of languages is negatively affected by stressful situations (e.g., Javier & Marcos, 1989). It is possible that under time pressure, languages could not be clearly separated from each other, which would lead to unclear representation of the correct response and thus result in suboptimal comparison between intended and actual response. This, in turn, would lead to lower amplitudes of the ERN. Thus, similar to the Ganushchak and Schiller (2006) study, we expected to find more erroneous responses and a smaller ERN under time pressure than in the absence of time pressure.

Methods

Participants

Twenty-one students of Maastricht University (20 women; mean age: 23.6 years) participated in the experiment. All participants were right-handed, German–Dutch bilinguals, and came from the same population as the bilingual speakers described in Christoffels, Firk, and Schiller (2007). Participants received course credits or a financial reward for their participation in the experiment and gave written informed consent prior to participating in the study. All participants were native German speakers and had completed an intensive Dutch language course prior to starting their undergraduate study in the Netherlands. They had studied in the Netherlands for at least 2 years (mean: 2.8 years) prior to being tested and usually lived in the Netherlands.

Table 1. *Vocabulary Test*

	Mean	*SD*
% correctly recognized words	55.42	15.37
% correctly rejected words	85.89	10.56
Mean of correct words and nonwords	67.56	9.45
ΔM	0.29	0.12

Most classes at the undergraduate level are given in Dutch; teaching materials are in Dutch or English. In their daily lives, the participants typically speak Dutch at the university and German at home.

Their level of proficiency was assessed with a self-rating questionnaire and a vocabulary test based on lexical decision. Both tests were completed after the experiment. Participants rated their language proficiency in two domains (active and passive knowledge) on a 10-point scale (1 = *very low*, 10 = *native level*). The mean score for active and passive knowledge of Dutch was 8.4. The vocabulary test was a Dutch version of an English nonspeeded lexical decision task that was originally developed by Meara (1996). It consisted of 60 items, 40 low-frequency words and 20 nonwords. Participants had to decide whether or not a presented letter string formed a correct Dutch word. Two ways of scoring were employed: the mean percentage of correctly recognized words and correctly rejected nonwords as well as Meara's *M* (ΔM; see Christoffels et al., 2007). ΔM lies between 0 and 1 and represents the proportion of words within the given frequency range that is known by a participant. The results are summarized in Table 1.

Materials

Ninety-two simple line drawings were used as pictures in this experiment (72 for experimental blocks and 20 for a practice block; see the Appendix for the list of stimuli used in the experimental blocks). The labels of all pictures were monosyllabic Dutch words (e.g., *heks* "witch," *brood* "bread," etc.). Per target phoneme, labels were matched on word length and frequency (see Table 2), that is, all picture names had a moderate frequency of occurrence between 10 and 100 per million according to the CEnter for LEXical information database (CELEX, Nijmegen; Baayen, Piepenbrock, & Gulikers, 1995). Picture labels all started with consonants. The position of the target phoneme was equated across the stimuli.

Design

The experiment included two experimental conditions: a control condition (CC) and a time pressure (TP) condition. In addition to the experimental conditions, a learning phase, two practice blocks, and two picture naming tasks were administered. The duration of the stimulus presentation during the control and time pressure conditions was computed separately for each participant. The duration of the stimulus presentation in the control condition was 85% of the reaction time (RT) obtained from the practice block (e.g., if the mean RT during the practice block was 1000 ms, then the duration of the stimuli in the CC was 850 ms). The mean RT of the CC was used to compute the stimulus duration for the TP condition. The RTs of the CC and not of the initial practice block were used for computation of the TP condition because the average RTs of the CC were based on more trials than RTs from the practice block. Participants were also more familiar with the task during the CC than during the practice block. Stimulus presentation in the TP condition was 75% of the RT of the CC (e.g., if stimulus presentation was 850 ms in the CC, then the duration of the stimulus in the TP condition was 638 ms). The percentages for computing the deadlines in this study were identical to the ones used in the previous study by Ganushchak and Schiller (2006). This was done to increase comparability between findings of these two studies. Prior to the experimental blocks, in the CC and TP conditions participants were required to repeat a practice block in order to adapt to the new timing. In practice and experimental blocks, a trial consisted of a fixation point with variable duration (between 500 and 800 ms), a blank screen for 500 ms, and the target stimulus, that is, a picture. Pictures disappeared from the screen as soon as a response was given or after the response deadline expired (depending on the condition; see above). The intertrial interval was variable, depending on the response latency. The time between the onset of the picture presentation and the button press was taken as the response time.

To compare the performance of the bilingual participants to a monolingual control group, we reanalyzed the data reported in Ganushchak and Schiller (2006) in the same way as the data of the current experiment. Note that there are some changes in the way the current data are analyzed with respect to the original data analysis in Ganushchak and Schiller (2006; e.g., mean area analysis instead of peak-to-peak analysis), which made it necessary to reanalyze the monolingual data of Ganushchak and Schiller (2006). Importantly, both the current study and the monolingual control study were carried out in Dutch, and therefore the two studies are well comparable.

CC and TP conditions each consisted of eight experimental blocks and one practice block (see Figure 1). In each block, participants were asked to monitor for a different target phoneme. The target phonemes were /t/, /k/, /p/, /n/, /m/, /l/, /s/, and /r/; the phoneme /b/ was used in the practice trials. In all blocks, pictures were presented one by one on a computer screen. Experimental blocks consisted of a total of 288 trials (mean 36 trials per block, with the exception of the practice block, which consisted of 20 trials). None of the pictures used for the practice block appeared as a target picture in the experimental conditions. Trials (i.e., order of pictures) were randomized across all blocks and for each participant. Each picture was repeated four times: twice as a target (go trials) and twice as a nontarget (no-go trials). Each time, participants were asked to monitor for a different phoneme. For instance, for the picture name *ster* ("star") participants were asked to monitor once for phoneme /t/ and once for the phoneme /s/ when *ster* was a target. When *ster* was a nontarget, participants were asked to monitor for /l/ and /n/.

Table 2. *Lexico-Statistical Characteristics of the Target Words*

Target phoneme	Example (with English translation in parentheses)	Mean CELEX frequency (per one million words)	Mean length in segments
t	troon (throne)	23.2	4.5
k	kraan (faucet)	28.4	4.2
p	paard (horse)	33.1	4.1
n	naald (needle)	30.6	4.2
m	maan (moon)	33.3	4.0
l	lamp (lamp)	33.5	4.6
s	schoen (shoe)	31.9	4.5
r	riem (belt)	29.9	4.3

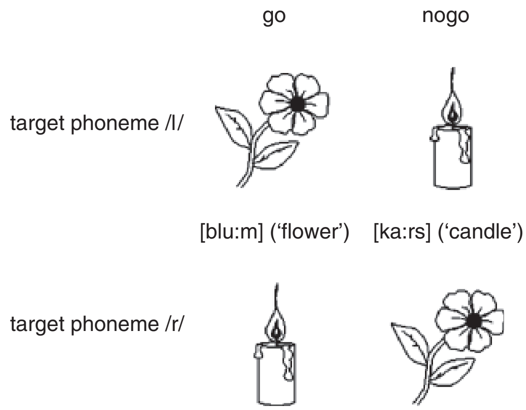

Figure 1. Example of go and no-go trials for two target phonemes. In the figure, Dutch picture names are written in phonetic code (taken from the CELEX database) and English translations are provided in parentheses. Each picture depicted here represents a separate trial. Each picture appeared in the task as a go and a no-go trial. At the beginning of a block, participants were instructed about which phoneme they had to monitor.

During the learning phase, the names of the pictures were presented via headphones. The picture remained in view for 3000 ms or until the response button was pressed. In the picture naming tasks, the pictures were presented without their corresponding names and disappeared from the screen as soon as the voice key was activated or after the response deadline was reached, which was identical to the time set for the control and the time pressure conditions.

Procedure

Participants were tested individually while seated in a soundproof booth. They were asked to carry out a learning phase, a practice block, a picture naming task, and then the CC; this was followed by a second practice block, a second picture naming task, and the TP condition. During the learning phase, participants were familiarized with the pictures and their corresponding names. In the picture naming task, participants were asked to overtly name pictures with the labels they learned during the learning phase. The timing of the second practice block and second picture naming task was identical to the one used in the phoneme-monitoring task in the TP condition. The purpose of the second picture naming task was to assure that participants had enough time to access and retrieve the name of the picture in the given time window.

Prior to practice and experimental blocks, participants received an auditory sample of the phoneme they were required to monitor (e.g., *Reageer nu op de klank /l/ zoals in tafel, spelen, verhaal* "React now to the sound /l/ like in table, play, tale"). Participants were required to press a button if the target phoneme was present in the picture name (i.e., go trials). When there was no target phoneme in the name of the picture, participants were required to withhold a response (i.e., no-go trials). Participants were instructed to give all responses for go trials with their right hand. Button-press latencies were recorded from the onset of the picture. At the end of the experimental session, participants were asked to fill out a questionnaire to assess their proficiency level. Participants were asked to perform the task in Dutch. Dutch was used in the instructions and in the conversations between experimenter and participants.

Apparatus and Recordings

The electroencephalogram (EEG) was recorded from 29 scalp sites (extended version of the 10/20 system) using tin electrodes mounted to an electrode cap. The EEG signal was sampled at 250 Hz with bandpass filter from 0.05 to 30 Hz. An electrode at the left mastoid was used for online referencing of the scalp electrodes. Off-line analysis included re-referencing of the scalp electrodes to the average activity of two electrodes placed on the left and right mastoids. Eye movements were recorded to allow off-line rejection of contaminated trials. Lateral eye movements were measured using a bipolar montage of two electrodes placed on the right and left external canthi. Eyeblinks and vertical eye movements were measured using a bipolar montage of two electrodes placed above and below the left eye. The impedance level for all electrodes was kept below 5 kΩ.

Data Analysis

Epochs of 1300 ms (from 400 ms to+900 ms) were obtained including a 200-ms preresponse baseline. The EEG signal was corrected for vertical electrooculogram (EOG) artifacts, using the ocular reduction method described in Anderer, Safety, Kinsperger, and Semlitsch (1987). To correct for nonocular artifacts, epochs with amplitudes above or below 75 μV were rejected. The amplitude of the ERN was derived from each individual's average waveforms after filtering with a bandpass, zero phase shift filter (frequency range: 1–12 Hz). The ERN was calculated in response-locked ERP averages across false alarms. False alarm trials were compared with correct go trials. The ERN was quantified by peak-to-peak measurements that were calculated to determine baseline-independent amplitudes of negative deflections by subtracting the amplitude of the preceding positive peak from the negative peak of this component (Falkenstein et al., 2000). Thus, the amplitude of the ERN was defined as the difference between the most negative peak in a window from 0 to 150 ms following the response and the most positive peak from50 to 0 ms preceding the ERN (Falkenstein et al., 2000). The amplitude of the ERN was recorded for each condition at Fz, FCz, and Cz electrode sites. For localization of the effects and comparison of correct trials across conditions we used the mean area analysis. The mean amplitude values were calculated per participant and condition in a time interval of 0–150 ms for German participants tested in the present study and for a data set of Dutch participants performing the same task that is described in Ganushchak and Schiller (2006). The time window was determined after careful visual inspection of the grand average ERP waveforms.

All analyses were performed on error and correct trials. Mean reaction times and false alarm rates (i.e., indicating an error on a correct trial) from each participant were submitted to repeated-measures analyses of variance (ANOVAs). The Greenhouse–Geisser correction was used for all repeated-measures ANOVAs. The analysis involved planned comparisons with Time Pressure (time pressure vs. control condition) and Response Type (correct vs. incorrect button press) as independent variables. Group (first vs. second language performance) was defined as a between-subjects factor.

The amplitude of the ERP waveforms was submitted to a repeated-measures ANOVA with Time Pressure (time pressure vs. control condition), Response Type (correct vs. incorrect button press), Location (prefrontal, i.e., Fp1, F3, F4, Fp2, vs. frontal, i.e., F7, FC3, FC4, F8, vs. central, i.e., Fz, Cz, Fcz, Pz, vs. parietal, i.e., TP7, P3, P4, TP8; see also Christoffels et al., 2007; Federmeier & Kutas, 1999), and Group (first vs. second language performance) as independent variables. This analysis was performed for the time windows specified above.

Results

Behavioral Data

RTs shorter than 300 ms and longer than 1500 ms were excluded from the analysis, which resulted in a loss of 0.7% of all trials. Table 3 provides an overview of the behavioral results. For button-press latencies, the analyses revealed a significant effect of Time Pressure, $F(1,37) = 750.82$, $MS_e = 764.99$, $p < .001$. Participants were faster during the TP condition than the CC. There was also a significant effect of Group, $F(1,37) = 7.19$, $MS_e = 164.73$, $p < .01$. Dutch participants were faster than German participants. A similar analysis with number of errors as the dependent variable also demonstrated a significant effect of Time Pressure, $F(1,37) = 31.19$, $MS_e = 28.60$, $p < .001$. Participants made more errors in the TP condition than in the CC. There was no significant effect of Group, $F < 1$. Participants made on average 8.75% errors (8.0% false alarms) in the TP condition and 6.9% (5.5% false alarms) errors in the CC (see also Table 3).

The picture naming task was used to assess whether or not participants had enough time to retrieve the name of the picture from their lexicon during the TP condition. To investigate this, a repeated-measures ANOVA was run for the picture naming task with Time Pressure as the independent variable. The number of errors during the picture naming task significantly decreased in the TP condition when compared to the CC, $F(1,38) = 84.42$, $MS_e = 5.09$, $p < .001$. Participants named 91% of the pictures correctly in the CC and 96% of the pictures in the TP. Hence, we argue that in the TP condition there was enough time available for participants to successfully retrieve the name of the pictures from their lexicon.

Electrophysiological Data

The ERN was revealed in response-locked ERP averages for false alarms. Figure 2 provides an overview of the response-locked averaged ERP waveforms for correct and incorrect trials across conditions (CC and TP) and electrodes (Fz, FCz, and Cz), where the ERN was the largest. The ERN obtained in the present study showed a frontal distribution (see Figure 3 for a topographical representation of the ERN across CC and TP conditions).

An ANOVA with Time Pressure as the independent variable and amplitude of the ERN (as determined by the peak-to-peak method) as the dependent variable with Group as a between-subjects factor revealed a significant effect of Time Pressure, $F(1,38) = 4.68$, $MS_e = 46.19$, $p < .05$, and a significant interaction between Time Pressure and Group, $F(1,38) = 4.02$,

Table 3. *Overview of the Behavioral Data*

	Control condition	Time pressure
Dutch		
Reaction times	769 (91)	619 (83)
Error rates	2.6 (11)	4.7 (14)
German		
Reaction times	865 (23)	671 (21)
Error rates	5.5 (6)	8.0 (12)

Note. Mean ($\pm$ *SD*) reaction times (in milliseconds) and percentage of false alarms ($\pm$ *SD*) as a function of Time Pressure and Group. Dutch data were published in Ganushchak and Schiller (2006).

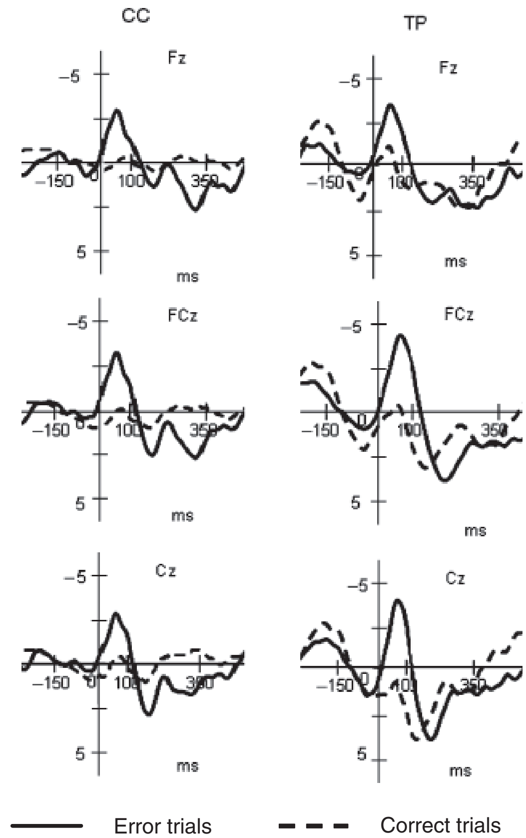

Figure 2. Averaged ERP waveforms for all incorrect versus correct trials across conditions and electrodes (CC: control condition; TP: time pressure condition). Correct and incorrect trials were matched on RTs and number of trials.

$MS_e = 57.01$, $p < .05$. Interestingly, German–Dutch bilinguals showed enhanced amplitudes of the ERN in the TP condition compared to the CC, whereas Dutch participants showed a decrease of the amplitude of the ERN under time pressure.

The mean area analysis showed similar results. There was a significant effect of Group, $F(1,38) = 16.89$, $MS_e = 38.25$, $p < .001$. Overall, German–Dutch participants had more negative amplitudes than native Dutch speakers (see Table 4). Furthermore, the analysis revealed a four-way interaction between Time Pressure, Response Type, Location, and Group, $F(3,114) = 17.06$, $MS_e = 1.63$, $p < .001$. There was no significant interaction between Time Pressure, Response Type, and Group, $F(1,38) = 1.50$, $MS_e = 23.98$, n.s. There were, however, significant interactions between Time Pressure and Group as well as between Response Type and Group, $F(1,38) = 31.19$, $MS_e = 24.98$, $p < .001$, and $F(1,38) = 16.37$, $MS_e = 25.26$, $p < .001$, respectively. To investigate these interactions, we have performed analyses separately for Dutch and German participants. For Dutch participants, effects of Time Pressure and Response Type were quantified in the significant three-way interaction between Time Pressure, Response Type, and Location, $F(3,54) = 4.81$, $MS_e = 1.71$, $p < .01$. To investigate this interaction in more detail, we looked at effect of Time Pressure and Response Type at each location separately. There was no interaction between Time Pressure and Response Type (for prefrontal, frontal, central, and parietal: all $Fs < 1$). Dutch participants showed a significant decrease in the amplitude of the ERN under

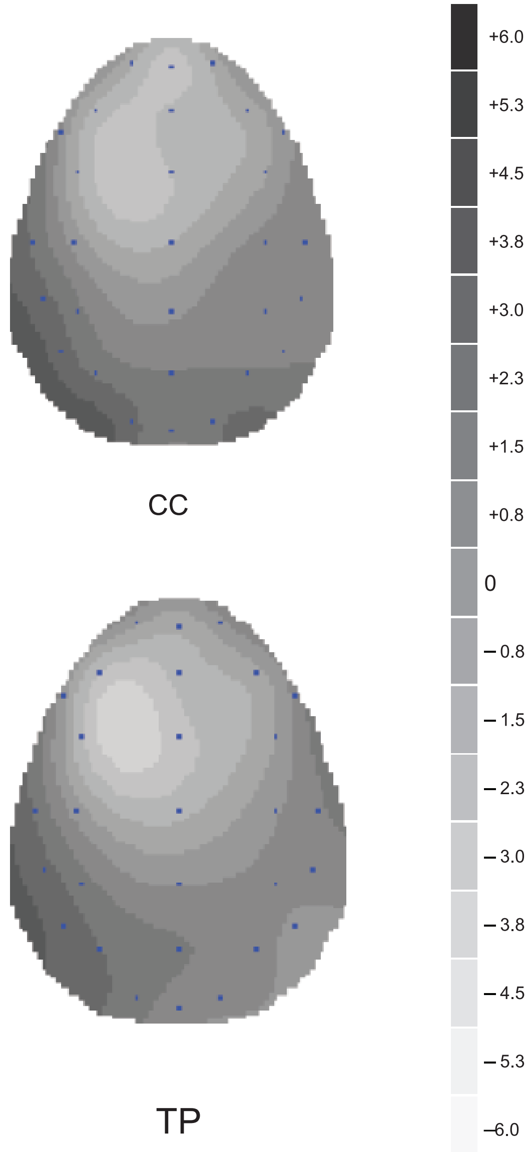

Figure 3. Topographic maps of the ERN amplitude between 0 and 100 ms after response onset. Negative regions depicted in light gray.

time pressure compared to no time pressure at frontal and central sites, $F(1,18) = 19.83$, $MS_e = 2.5$, $p < .001$ and $F(1,18) = 10.39$, $MS_e = 8.40$, $p < .01$, respectively. There was no effect of Time Pressure at prefrontal and parietal sites, $F < 1$ and $F(1,18) = 1.12$, $MS_e = 6.05$, n.s., respectively. Error trials were significantly more negative than correct trials at all locations: prefrontal: $F(1,18) = 16.20$, $MS_e = 9.32$, $p < .01$; frontal: $F(1,18) = 92.13$, $MS_e = 7.88$, $p < .001$; central: $F(1,18) = 55.33$, $MS_e = 15.55$, $p < .001$; parietal: $F(1,18) = 46.07$, $MS_e = 6.17$, $p < .001$ (see Table 4; see also Ganushchak & Schiller, 2006).

Note that statistical methods reported in this article for Dutch participants are slightly different from the ones reported in Ganushchak and Schiller (2006). In Ganushchak and Schiller (2006), we used the so-called peak-to-peak method to analyze the ERN data. In the present study, we reported mean area analyses of the same data set. We chose mean area analyses because these analyses allowed us to better localize the effect on the scalp and to make a better comparison between correct trials for the control

Table 4. *Overview of the Electrophysiological Data*

	Control condition		Time pressure	
	False alarms	Correct go trials	False alarms	Correct go trials
Dutch				
Prefrontal	0.64 (3)	1.40 (1)	0.04 (0.4)	2.09 (2)
Frontal	– 0.63 (2)	1.76 (1)	0.13 (0.4)	2.33 (2)
Central	– 0.86 (3)	2.51 (2)	0.22 (1)	3.77 (3)
Parietal	0.67 (1)	2.38 (2)	0.14 (0.3)	3.03 (2)
German				
Prefrontal	– 0.62 (2)	0.31 (2)	– 1.28 (2)	– 1.03 (2)
Frontal	– 0.01 (2)	1.00 (2)	– 1.31 (2)	– 0.59 (2)
Central	0.11 (2)	1.59 (2)	– 2.06 (2)	– 1.12 (2)
Parietal	0.71 (1)	1.61 (2)	– 0.79 (1)	– 0.31 (2)

Note. Mean (± *SD*) amplitudes (in microvolts) as a function of Time Pressure condition, Type of Response, Location, and Group.

and time pressure conditions. Note that the mean area and peak-to-peak analyses showed the same effects for the ERN, that is, a decrease of the amplitude of the ERN under time pressure.

For German–Dutch bilinguals, there was no significant interaction between Time Pressure, Response Type, and Location, $F < 1$. The interaction between Time Pressure and Response Type was not significant, either, $F < 1$. There was, however, a significant interaction between Response Type and Location as well as between Time Pressure and Location, $F(3,60) = 6.25$, $MS_e = .81$, $p < .001$ and $F(3,60) = 19.54$, $MS_e = 1.57$, $p < .001$, respectively. Further investigation of these interactions showed that the amplitude of the ERN was significantly more negative for error trials compared to correct trials. This difference was significant for all locations: prefrontal: $F(1,20) = 6.02$, $MS_e = 4.85$, $p < .05$; frontal: $F(1,20) = 14.80$, $MS_e = 4.22$, $p < .001$; central: $F(1,20) = 15.78$, $MS_e = 7.83$, $p < .001$; parietal: $F(1,20) = 10.15$, $MS_e = 4.00$, $p < .01$. However, the difference between error and correct trials was largest at central sites.

Further, German–Dutch bilinguals showed enhanced amplitude of the ERN in the TP condition compared to the CC (see Table 4). This difference was significant at all locations: prefrontal: $F(1,20) = 9.54$, $MS_e = 8.85$, $p < .01$; frontal: $F(1,20) = 27.59$, $MS_e = 6.43$, $p < .001$; central: $F(1,20) = 46.93$, $MS_e = 10.69$, $p < .001$; parietal: $F(1,20) = 55.06$, $MS_e = 4.49$, $p < .001$. However, the difference between the amplitude of the ERN in the control condition and time pressure condition was largest at central sites. These results are striking and unexpected. Therefore, we looked at how participants behaved at a single-subject level. We found that 73% of the participants (16 out of 21) showed an enhanced ERN under time pressure compared to the absence of time pressure, whereas 27% of the participants (5 out of 21) showed lower amplitudes of the ERN under time pressure compared to the control condition. See Figure 4 for a comparison between native Dutch speakers and German–Dutch bilinguals.

In the CC, there appeared to be a second negative peak at around 200 ms after the response, which was smaller in the TP condition. To test whether or not there was a significant difference between conditions, we employed a mean area analysis in the time window of 140–270 ms. A 2 (correct vs. error) × 2 (CC vs. TP) ANOVA revealed no significant effects of Time Pressure and Correctness of Response, $F(1,20) = 4.07$, $MS_e = 34.62$, n.s. and $F < 1$, respectively, nor an interaction between these two factors, $F < 1$. It appears from Figure 4 that there seems to be a latency difference between the monolingual and the bilingual

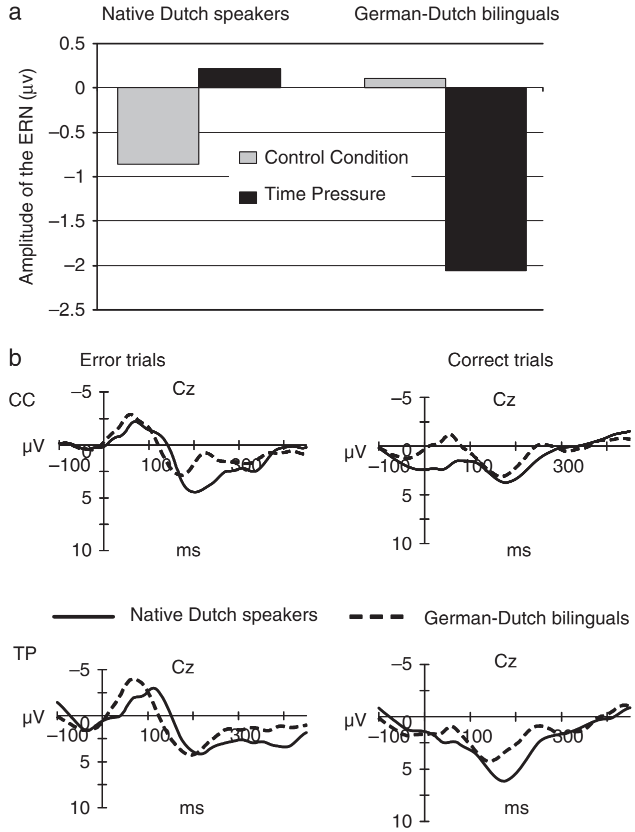

Figure 4. a: The ERN mean amplitude values at electrode site Cz for all incorrect trials across conditions for native Dutch speakers and German–Dutch bilinguals, separately. b: Averaged ERP waveforms at electrode site Cz for incorrect and correct trials compared across conditions and group (CC: control condition; TP: time pressure condition). Note that, for presentation purposes, the ERP averages were filtered with a high-pass filter.

group. However, an ANOVA with latency of the ERN as the dependent variable revealed no significant effects, all $Fs < 1$. Thus, we do not believe that the somewhat earlier latency of the ERN in the bilingual group can account for the enhanced ERN amplitude.

Discussion

The goal of the present study was to investigate how the ERN is affected by time pressure when a verbal self-monitoring task is performed in a second language as opposed to performance in the native language. We demonstrated that bilingual participants made more errors under time pressure. This is in accordance with previous monolingual findings (e.g., Ganushchak & Schiller, 2006; Oomen & Postma, 2001). Contrary to previously reported findings, however, we observed an increase in the amplitude of the ERN under time pressure as compared to a control condition. In the action monitoring as well as verbal monitoring literature, it has been shown that the ERN decreases under time pressure (Falkenstein et al., 1991; Ganushchak & Schiller, 2006; Gehring et al., 1993). Presumably, a monitoring system compares the representation of the correct response with the copy of an online response. If there is a mismatch between actual and intended motor or verbal response, an error signal is generated (e.g., Desmurget & Grafton, 2000; Levelt, 1983). Under time pressure, there might not be enough time available to make an optimal comparison between intended and actual responses. As a result, a weaker signal is sent to the remedial action system thereby decreasing the amplitude of the ERN. In terms of the reinforcement-learning theory, errors induce a phasic decrease in mesencephalic dopaminergic activity when ongoing events are determined to be worse than expected (Holroyd & Coles, 2002). However, under time pressure, due to the lack of time or cognitive resources, the monitoring system might not be able to make an optimal evaluation of current events and events that were predicted. Therefore, a weaker ERN is generated.

Alternatively, according to the new interpretation of the conflict monitoring theory, the ERN reflects conflict that develops after errors as a consequence of continued stimulus processing. This processing results in post-error activation of the correct response and hence conflict with the incorrect response just produced (Yeung, Botvinick, & Cohen, 2004). The difference between this account and the mismatch hypothesis is that the mismatch hypothesis proposes that the ERN reflects the output of a system specifically devoted to error detection. The conflict theory associated the ERN with conflict monitoring that also occurs on correct trials and that may represent the input to, rather than the output from, the error detection system (Yeung et al., 2004). According to Yeung and colleagues, a decrease in the amplitude of the ERN under time pressure is due to a less focused attentional state than under no time pressure, which may result in gaining speed.

Why did we observe an increase in the amplitude of the ERN under time pressure in a bilingual context, but not in a monolingual context? Assuming that verbal self-monitoring works similarly in first and second languages (Kormos, 1999; Poulisse, 2000; Van Hest, 1996), one would predict that a monitoring system can compare the representation of the correct response with the copy of an online response in the second language. If there is a mismatch between actual and intended verbal response, an error signal should be generated, and under time pressure this signal should be weaker, thereby decreasing the amplitude of the ERN in bilinguals as well as monolinguals.

However, we obtained an enhanced ERN under time pressure compared to the absence of time pressure. How can we explain this reversed effect of time pressure on the ERN in bilinguals? We would like to propose the following possibility: Participants, in the present study, were bilingual German–Dutch students, who were requested to perform a phoneme-monitoring task in their second language, that is, Dutch. To perform this task, participants presumably had to suppress their more dominant mother tongue to generate a Dutch name of the picture and determine whether or not the target phoneme was present in the name of the picture. It has long been known that switches between languages can occur unintentionally, for instance, in aphasic bilingual speakers (e.g., Fabbro, Skrap, & Aglioti, 2000), when bilinguals undergo brain stimulation (e.g., Holtzheimer, Fawaz, Wilson, & Avery, 2005), or under psychological stress (e.g., Dornic, 1979, 1980; Grosjean, 1982). According to Levelt (1989), monitoring involves controlled processing that requires attentional control. In a second language, a considerably lower number of cognitive processes are automatic and thus need more attention than in the first language (Kormos, 1999). It is possible that, under time pressure, participants had more difficulty inhibiting their dominant native language and experienced more intrusions from it.

Rodríguez-Fornells et al. (2005) demonstrated that bilinguals cope with second language interference during language production by recruiting ''executive function'' brain areas, that is, the left prefrontal cortex, the supplementary motor area, and the left middle prefrontal cortex. These areas might be crucial in inhibiting the production of irrelevant, nontarget language words (Rodríguez-Fornells et al., 2005). It is possible that under time pressure, inhibition of the nontarget words was less successful than in the absence of time pressure. There is evidence from bilingual word recognition that even in a monolingual task alternative lexical candidates in the other language are accessed (for a review, see Kroll & Dijkstra, 2002) and phonologically activated (Costa, Caramazza, & Sebastián-Gallés, 2000; Colomé, 2001; Rodríguez-Fornells et al., 2005; but see also Hermans, Bongaerts, De Bot, & Schreuder, 1998). Hence, it is possible that at the time of the response, there was not only the Dutch name of the picture active but also the German name. During execution of the monitoring task in a native language (Ganushchak & Schiller, 2006), it is unlikely that there were intrusions from a less dominant second language, which means that the monitor did not need to deal with resolving a competition between multiple responses. In contrast, performing the task in a second language could have required a resolution of response competition between an inappropriate response (e.g., a phoneme from a German word) and a correct response (e.g., a phoneme from a Dutch word). This is in accordance with Yeung and colleagues' (2004) interpretation of the conflict theory: Continued stimulus processing after the response could have resulted in the activation of multiple candidates for the correct response, for example, Dutch and German words, which would have led to higher conflict and higher amplitudes of the ERN. This may also explain why we did not find an interaction between Time Pressure and Response Type, because the conflict between multiple correct responses could have been present on both correct and error trials. Activation of both German and Dutch names could have resulted in more response conflict and thus higher amplitudes of the ERN (e.g., Botvinick et al., 2001; Yeung et al., 2004). Interestingly, in their study, Möller and colleagues (2007) showed a negative deflection prior to vocalization of errors, which was absent prior to vocalization of a correct response. Möller and colleagues argued that this negativity was a result of a conflict that arose at a processing level related to the phonetic encoding or articulatory planning of speech output.

Our results are also in agreement with the Sebastián-Gallés and colleagues (2006). They showed a larger ERN and an increased negativity at correct trials for less dominant bilinguals compared to more proficient bilinguals. In our study, we showed that the German–Dutch participants, when performing a task in their less dominant second language (i.e., Dutch), showed enhanced ERN and correct-related negativity (CRN) on correct trials compared to the native Dutch speakers, who performed the task in their dominant language (i.e., Dutch).

However, our findings are in disagreement with the error detection theory (Bernstein et al., 1995), according to which the ERN under time pressure should be of lower amplitude compared to the absence of time pressure due to the lack of time to make an optimal comparison between intended and actual responses. Similarly, the reinforcement-learning theory (Holroyd & Coles, 2002) cannot fully account for our findings. As stated above, the reinforcement-learning theory assumes that errors induce a phasic decrease in mesencephalic dopaminergic activity when ongoing events are determined to be worse than expected (Holroyd & Coles, 2002). Time pressure may result in the inability of the monitoring system to make an optimal evaluation of current events and events that were predicted, therefore predicting smaller ERN amplitudes under time pressure compared to the absence of time pressure.

Suggestively, the increased amplitude of the ERN under time pressure in bilingual situations might be dependent on the proficiency of second-language speakers. Proficiency is a determining factor in the ease with which bilinguals control and regulate their two (or more) languages (Meuter, 2005). Participants in the present study completed a course of Dutch language and studied at a Dutch university. However, they were not balanced bilinguals. It is possible that highly proficient, balanced bilinguals will be more successful in suppressing a language not required for the task and thus have less or no interference of the native language in the second-language context. Therefore, it is plausible that the amplitude of the ERN will show a typical decrease under time pressure when highly proficient second-language speakers perform the task.

One potential problem of the current study is the order of experimental conditions; that is, the time pressure condition was always preceded by the control condition. It is possible that in the TP condition, participants were more experienced in the task than in the CC, and therefore the findings of the experiment could be attributed to a practice effect. However, if practice played a significant role here, then one would expect that participants would perform the task more accurately and make fewer errors in the TP condition than in the CC. The findings of the current study demonstrate the opposite; that is, participants made more errors under time pressure than in the absence of time pressure. Moreover, Ganushchak and Schiller (2006) showed that simple repetition of the control condition, without time pressure manipulation, does not influence the amplitude of the ERN. Thus, we believe that in the present study practice did not have a large influence on performance and amplitudes of the ERN. However, it cannot be completely excluded that order of conditions had some effect on the performance.

Note that, to compute the ERN in the present study, we had on average 13 error trials in the control condition and 20 error trials in the time pressure condition per participant, which might be considered a relatively low number of trials and thus a potential limitation of the current study. However, the ERN is a robust component and can easily be seen even on an individual trial-by-trial basis. Even though our error rate is relatively low, we do find reliable effects. Some of the previous research in this area also showed reliable effects of the ERN with similar error rates. For instance, Vidal, Hasbroucq, Grappenron, and Bonnet (2000) had error rates of 2.4% and 3.2% (about 15 and 20 trials on average, respectively) and stated that, to permit error analysis, at least five trials were enough. Besides this support from the literature, we would like to emphasize that in our study, the ERN component was clearly visible on erroneous trials and the signal-to-noise ratio was good enough to compute statistical comparisons between conditions. However, we cannot completely exclude the possibility that with more error trials, the ERN could have a slightly different morphology than shown in the present study.

The main manipulation employed in the present study was time pressure. In speeded tasks, there is obviously the possibility of a speed–accuracy trade-off (SAT). One way in which people control their actions occurs when speed or accuracy are more important. As stated above, previous studies that investigated the ERN under time pressure demonstrated that the amplitude of the ERN decreases when participants select speed over accuracy (Falkenstein et al., 1991; Gehring et al., 1993). However, in the present study, we obtained the opposite pattern. The amplitude of the ERN was enhanced under time pressure compared to the absence of time pressure. Therefore, our results cannot be fully accounted by SAT effects.

In summary, we showed that the ERN can successfully be elicited by errors of verbal monitoring and is sensitive to the linguistic context. Performing the task in a second language led to an enhancement of the ERN under time pressure as compared to when time pressure was absent. This effect is reversed when the task is performed in a native language; that is, the amplitude of the ERN is lower under time pressure than in the absence of time pressure. This provides further evidence that the ERN is sensitive to verbal manipulations and could be used as an electrophysiological marker of error processing in language research.

As a note of caution, we would like to mention that in the present study the required responses were button presses. We believe that the majority of errors observed in the current study are errors of the verbal monitoring system and are based on the incorrect decision about the target phoneme. We cannot completely rule out the possibility, however, that some of the errors could have been due to action slips (i.e., slips of the hand) and not slips of verbal monitoring per se. However, this seems unlikely because action slips did not lead to an enhancement of the ERN under time pressure in previous research (Falkenstein et al., 1991; Gehring et al., 1993). The reversal effect of time pressure on the ERN in multilingual situations merits further investigation, for example, by manipulating the proficiency of participants in their second language.

REFERENCES

Anderer, P., Safety, B., Kinsperger, K., & Semlitsch, H. (1987). Topographic brain mapping of EEG in neuropsychopharmacology—Part 1. Methodological aspects. *Methods and Findings in Experimental Clinical Pharmacology*, *9*, 371–384.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database* (CD-ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Bernstein, P. S., Scheffers, M. K., & Coles, M. G. H. (1995). 'Where did I go wrong?' A psychophysiological analysis of error detection. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1312–1322.

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.

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research*, *1147*, 192–208.

Coles, M. G. H., Scheffers, M. K., & Holroyd, C. B. (2001). Why is there an ERN/Ne on correct trials? Response representations, stimulus-related components, and the theory of error-processing. *Biological Psychology*, *56*, 173–189.

Colomé, Á. (2001). 'Lexical activation in bilinguals' speech production: Language-specific or language independent? *Journal of Memory and Language*, *45*, 721–736.

Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect: Implications for the models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1283–1296.

Desmurget, M., & Grafton, S. T. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in Cognitive Science*, *4*, 423–431.

Dornic, S. (1979). Information processing in bilinguals: Some selected issues. *Psychological Research*, *40*, 329–348.

Dornic, S. (1980). Information processing and language dominance. *International Review of Applied Psychology*, *29*, 119–140.

Fabbro, F., Skrap, M., & Aglioti, S. (2000). Pathological switching between languages following frontal lesions in a bilingual patient. *Journal of Neurology, Neurosurgery and Psychiatry*, *68*, 650–652.

Falkenstein, M., Hohnsbein, J., Hoorman, 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.

Federmeier, K., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, *41*, 469–495.

Ganushchak, L. Y., & Schiller, N. O. (2006). Effects of time pressure on verbal self-monitoring. *Brain Research*, *1125*, 104–115.

Ganushchak, L. Y., & Schiller, N. O. (2008). Brain error-monitoring activity is affected by semantic relatedness: An event-related brain potentials study. *Journal of Cognitive Neuroscience*, *20*, 927–940.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*, 385–390.

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Cambridge, MA: Harvard University Press.

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, *1*, 213–230.

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.

Holroyd, C. B., & Yeung, N. (2003). Alcohol and error processing. *Trends in Neurosciences*, *26*, 402–404.

Holtzheimer, P., Fawaz, W., Wilson, C., & Avery, D. (2005). Repetitive transcranial magnetic stimulation may induce language switching in bilingual patients. *Brain and Language*, *94*, 274–277.

Javier, R. A., & Marcos, L. R. (1989). The role of stress on the language-independent and code-switching phenomenon. *Journal of Psycholinguistic Research*, *18*, 449–472.

Klein, D., Milner, B., Zatorre, R. J., Meyer, E., & Evans, A. C. (1995). The neural substrates underlying word generation: A bilingual functional-imaging study. *Proceedings of the National Academy of Sciences, USA*, *92*, 2899–2903.

Klein, D., Milner, B., Zatorre, R. J., Zhao, V., & Nikelski, J. (1999). Cerebral organization in bilinguals: A PET study of Chinese-English verb generation. *NeuroReport*, *10*, 2841–2846.

Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning*, *49*, 303–342.

Kroll, J. F., & Dijkstra, T. (2002). The bilingual lexicon. In R. B. Kaplan (Ed.), *Handbook of applied linguistics* (pp. 301–324). Oxford, UK: Oxford University Press.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, *14*, 41–104.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Lickley, R. J., Hartsuiker, R. J., Corley, M., Russell, M., & Nelson, R. (2005). Judgment of disfluency in people who stutter and people who do not stutter: Results from magnitude estimation. *Language and Speech*, *48*, 299–312.

Lucas, T. H. II, McKhann, G. M., & Ojemann, G. A. (2004). Functional separation of languages in the bilingual brain: A comparison of electrical stimulation language mapping in 25 bilingual patients and 117 monolingual control patients. *Journal of Neurosurgery*, *101*, 449–457.

Meara, P. (1996). *English Vocabulary Test 10K*. Swansea, UK: Centre for Applied Language Studies, University of Wales.

Masaki, H., Tanaka, H., Takasawa, N., & Yamazaki, K. (2001). Error-related brain potentials elicited by vocal errors. *NeuroReport*, *12*, 1851–1855.

Meuter, R. F. I. (2005). Language selection in bilinguals: Mechanisms and processes. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 349–370). Oxford, UK: Oxford University Press.

Möller, J., Jansma, B. M., Rodríguez-Fornells, A., & Münte, T. F. (2007). What the brain does before the tongue slips. *Cerebral Cortex*, *17*, 1173–1178.
Motley, M. T., Camden, C. T., & Baars, B. J. (1982). Covert formulation and editing of anomalies in speech production: Evidence from experimentally elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, *21*, 578–594.
Oomen, C. C. E., & Postma, A. (2001). Effects of time pressure on mechanisms of speech production and self-monitoring. *Journal of Psycholinguistic Research*, *30*, 163–184.
Oomen, C. C. E., Postma, A., & Kolk, H. (2001). Prearticulatory and postarticulatory self-monitoring in Broca's aphasia. *Cortex*, *37*, 627–641.
Perani, D., Paulesu, E., Sebastián-Gallés, N., Dupoux, E., Dehaene, S., Bettinardi, V., et al. (1998). The bilingual brain: Proficiency and age of acquisition of the second language. *Brain*, *121*, 1841–1852.
Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition*, *77*, 97–131.
Postma, A., & Noordanus, C. (1996). Production and detection of speech errors in silent, mouthed, noise-masked, and normal auditory feedback speech. *Language and Speech*, *39*, 375–392.
Poulisse, N. (2000). Slips of the tongue in first and second language production. *Studia Linguistica*, *54*, 136–149.
Rodríguez-Fornells, A., Kurzbuch, A. R., & Münte, T. F. (2002). Time course of error detection and correction in humans: Neurophysiological evidence. *Journal of Neuroscience*, *22*, 9990–9996.
Rodríguez-Fornells, A., Van der Lugt, A., Rotte, M., Britti, B., Heinze, H. J., & Münte, T. F. (2005). Second language interferes with word production in fluent bilinguals: Brain potential and functional imaging evidence. *Journal of Cognitive Neuroscience*, *17*, 422–433.
Scheffers, M. K., Coles, M. G. H., Bernstein, P. S., Gehring, W. J., & Donchin, E. (1996). Event-related brain potential and error-related processing: An analysis of incorrect responses to go and no-go stimuli. *Psychophysiology*, *33*, 42–53.
Schiller, N. O. (2005). Verbal self-monitoring. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 245–261). Mahwah, NJ: Lawrence Erlbaum Associates.
Schiller, N. O. (2006). Lexical stress encoding in single word production estimated by event-related brain potentials. *Brain Research*, *1112*, 201–212.
Schiller, N. O., Jansma, B. M., Peters, J., & Levelt, W. J. M. (2006). Monitoring metrical stress in polysyllabic words. *Language and Cognitive Processes*, *21*, 112–140.
Seal, M. L., Aleman, A., & McGuire, P. K. (2004). Compelling imagery, unanticipated speech and deceptive memory: Neurocognitive models of auditory verbal hallucinations in schizophrenia. *Cognitive Neuropsychiatry*, *9*, 43–72.
Sebastián-Gallés, N., Rodríguez-Fornells, A., De Diego-Balaquer, R., & Díaz, B. (2006). First- and second-language phonological representation in the mental lexicon. *Journal of Cognitive Neuroscience*, *18*, 1277–1291.
Van Hest, E. (1996). *Self-repair in L1 and L2 production*. Tilburg, the Netherlands: Tilburg University Press.
Vidal, F., Hasbroucq, T., Grappenron, J., & Bonnet, M. (2000). Is the 'error negativity' specific to errors. *Biological Psychology*, *51*, 109–128.
Wheeldon, L. R., & Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language*, *34*, 311–334.
Wheeldon, L. R., & Morgan, J. L. (2002). Phoneme monitoring in internal and external speech. *Language and Cognitive Processes*, *17*, 503–535.
Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the Error-Related Negativity. *Psychological Review*, *111*, 931–959.

(Received December 21, 2007; Accepted June 23, 2008)

APPENDIX

The following is a list of stimuli used in the experimental blocks. The approximate English translation is given in parentheses. Each stimulus appears twice as a target, but each time with a different target phoneme (e.g., *hemd* ["shirt"] has the target phonemes /t/ and /m/; due to final devoicing, the <d> in *hemd* is pronounced as /t/).

TARGET PHONEME /t/: hemd (shirt), pet (cap), troon (throne), trui (sweater), baard (beard), blad (leaf), net (net), stier (bull), tak (branch), ster (star), tram (tram), bord (plate), fiets (bike), stof (material), kaart (card), trein (train), paard (horse), pot (pot), band (tire), ton (barrel), kast (closet), zwaard (sword), vuist (fist)

TARGET PHONEME /k/: kom (bowl), broek (trousers), markt (market), kraan (tap), kist (chest), kip (chicken), wolk (cloud), tak (branch), heks (witch), knie (knee), jurk (dress), kaars (candle), kaart (card), rok (skirt), kroon (crown), krant (newspaper), kruis (cross), kraag (collar), vork (fork), kaas (cheese), kar (wagon), stok (stick)

TARGET PHONEME /p/: pan (pan), plant (plant), knop (button), pet (cap), kip (chicken), schaap (sheep), pen (pen), trap (stairs), plank (shelf), dorp (village), schip (ship), paard (horse), spoor (rail), pot (pot)

TARGET PHONEME /n/: pan (pan), nest (nest), troon (throne), snor (moustache), knie (knee), pen (pen), naald (needle), knop (button), mand (basket), net (net), band (tire), maan (moon), kroon (crown), krant (newspaper), neus (nose), schoen (shoe), hoorn (horn), ton (barrel), trein (train)

TARGET PHONEME /l/: lamp (lamp), film (film), bloem (flower), plant (plant), naald (needle), plank (shelf), wolk (cloud), fles (bottle), blad (leaf), slot (lock), schaal (dish)

TARGET PHONEME /m/: kom (bowl), muur (wall), riem (belt), hemd (shirt), bloem (flower), mand (basket), film (film), lamp (lamp), mes (knife), markt (market), maan (moon), tram (tram)

TARGET PHONEME /s/: mes (knife), fles (bottle), slot (lock), nest (nest), stier (bull), schaap (sheep), rots (rock), kist (chest), heks (witch), ster (star), fiets (bike), schaal (dish), stof (material), kaas (cheese), gras (grass), schip (ship), schoen (shoe), neus (nose), stok (stick), vuist (fist), kast (closet), kruis (cross)

TARGET PHONEME /r/: muur (wall), riem (belt), dorp (village), trui (sweater), kraan (tap), broek (trousers), snor (moustache), trap (stars), rots (rock), baard (beard), bord (plate), rok (skirt), gras (grass), kaars (candle), jurk (dress), spoor (rail), hoorn (horn), kar (wagon), zwaard (sword), vork (fork), kraag (collar)