The nature of the verbal self-monitor

Ganushchak, A.Y.

Chapter 2

Effects of time pressure on verbal self-monitoring: An ERP study[2]

Abstract

The Error-Related Negativity (ERN) is a component of the event-related brain potential (ERP) that is associated with action monitoring and error detection. The present study addressed the question whether or not an ERN occurs after verbal error detection, e.g., during phoneme monitoring. We obtained an ERN following verbal errors which showed a typical decrease in amplitude under severe time pressure. This result demonstrates that the functioning of the verbal self-monitoring system is comparable to other performance monitoring, such as action monitoring. Furthermore, we found that participants made more errors in phoneme monitoring under time pressure than in a control condition. This may suggest that time pressure decreases the amount of resources available to a capacity-limited self-monitor thereby leading to more errors.

2 This chapter is based on Ganushchak, L. Y., & Schiller, N. O. (2006). Effects of time pressure on verbal self-monitoring: An ERP study. *Brain Research*, 1125, 104-115

Introduction

Error monitoring is an important executive function, which helps to adapt, anticipate, learn, correct, and mend the consequences of actions. The neural basis of error monitoring has become a key issue in cognitive neuroscience due to its importance to the aforementioned cognitive skills. A better understanding of its working may offer new insights into the dysfunctions of self-monitoring seen in a range of clinical conditions such as schizophrenia (Carter, MacDonald III, Ross, & Stenger, 2001), opiate addicts (Forman et al., 2004), and obsessive-compulsive disorder (Gehring, Himle, & Nisenson, 2000).

Progress in identifying the functional characteristics of the error monitoring system has been mainly achieved through the study of an electrophysiological index thought to be associated with error processing, i.e. *Error-Related Negativity* (ERN; Falkenstein, Hohnsbein, Hoorman, & Blanke, 1991; Gehring, Goss, Coles, Meyer, & Donchin, 1993). The ERN is a component of the event-related potential (ERP) that 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). The early onset latency of the ERN with respect to the incorrect response is suggestive of an error monitoring system. The generation of the ERN has been localized in the anterior cingulate cortex (ACC; Dehaene, Posner, & Tucker, 1994; Holroyd & Coles, 2002). Several hypotheses of performance monitoring have been proposed to account for the ERN, for instance, the *mismatch hypothesis* put forward by Falkenstein and colleagues (1991), the *response conflict hypothesis* proposed by Carter and colleagues (1998), and the *reinforcement learning theory* by Holroyd and Coles (2002).

The mismatch hypothesis considers the ERN as the result of a mismatch between the intended and the actual response execution (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).

The conflict hypothesis, in contrast, states that the ERN reflects detection of response conflict and not detection of errors per se (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Carter et al., 1998). A response conflict arises when multiple responses compete with each other 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 is based on the presence of response conflict.

More recently, the reinforcement-learning theory has been developed (Holroyd & Coles, 2002). According to this theory, 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 ACC, where the ERN is elicited. In other words, the ERN is a neurobiological index of comparison processes that are mediated by the dopamine system and responsive to the discrepancy between predicted and actual reward (Holroyd & Coles, 2002).

The majority of studies on the ERN investigated the working of action monitoring. The action monitor is a feed-forward control mechanism that is used to inhibit and correct a faulty response (Desmurget & Grafton, 2000; Rodrígues-Fornells, Kurzbuch, & Münte, 2002). When the wrong selection of the motor command is generated, a copy of an on-line response is produced and compared to the representation of the correct response. If there is a mismatch between the copy of the on-line response and the representation of the correct response, an error signal is generated and a stop command is initiated (Coles, Scheffers, & Holroyd, 2001). The question addressed in the present study is whether or not *verbal* monitoring works in a similar way as action monitoring. It seems plausible that different types of monitoring have the same key mechanisms to monitor different kinds of behavioral output. In such a way, an action monitor may monitor, for example, for motor slips by checking for possible mismatches between representations of actual and desired motor behavior. A verbal monitor, on the other hand, may, for instance, monitor some internal representation as it is produced during speech planning by checking potential mismatches between intended and actual verbal production.

One of the most detailed theories about verbal self-monitoring is the perceptual-loop theory proposed by Levelt (1983, 1989). According to this theory, there is a single, central monitor that is located in the so-called *conceptualizer* (see Figure 1). This monitor receives information from the conceptual loop, the inner loop, and the auditory loop. First, immediately after conceptualization of a verbal message, the conceptual loop checks the message for its appropriateness. Second, the inner loop inspects the speech plan prior to its articulation (Postma & Noordanus, 1996). The inner loop has access to abstract codes, i.e. the phonological planning level (Schiller, 2005, in press; Schiller, Jansma, Peters, & Levelt, 2006; Wheeldon & Levelt, 1995; Wheeldon & Morgan, 2002). For instance, Wheeldon and Levelt (1995) asked participants to silently generate the Dutch translation of an auditorily presented English word and to monitor their covert production for a specific target segment in the Dutch translation. For example, when participants were presented with the word *hitchhiker* and generated the Dutch translation *lifter*, then they were required to press a button if target phoneme was /t/ (since /t/ is a phoneme of *lifter*) but they withheld their response in case the target phoneme was /k/. The findings of Wheeldon and Levelt (1995) demonstrated that participants were faster in detecting onset as opposed to offset phonemes. Based on their findings the authors concluded that participants indeed monitor an abstract internal speech code during a segment/phoneme-monitoring task. The auditory loop, finally, can detect errors via the speech comprehension system after the speech has become overt (Postma, 2000).

Self-monitoring one's own speech is important because 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). Furthermore, verbal-monitoring is often implicated in disorders such as aphasia (for an overview see Oomen, Postma, & Kolk, 2001), stuttering (Lickley, Hartsuiker, Corley, Russell, & Nelson, 2005), and schizophrenia (for overview see Seal, Aleman, & McGuire, 2004).

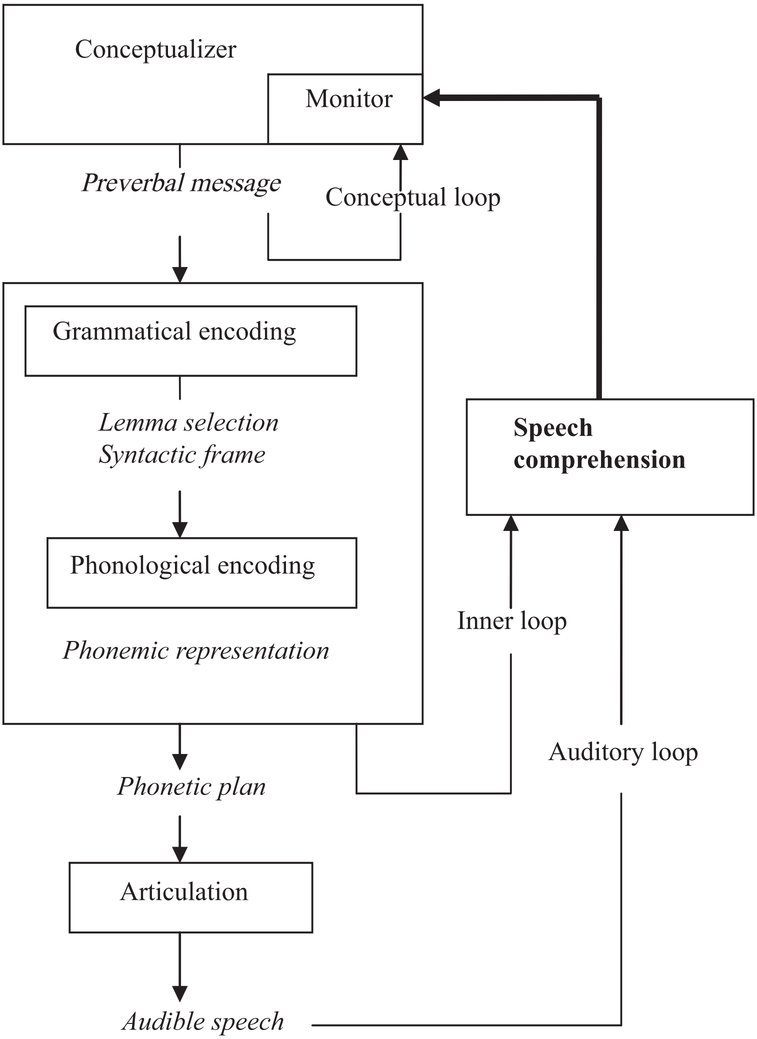

Figure 1. Graphical representation of Levelt's speech production model.

Current study

The objective of the present research is to further our understanding of the verbal self-monitor by examining the relationship between the ERN and errors of the verbal monitor. Considering that the ERN is indifferent to modality of the error information (Holroyd & Coles, 2002), it seems plausible to assume that the ERN will also be generated by verbal errors. One study conducted by Masaki, Tanaka, Takasawa, and Yamazaki (2001) examined whether the ERN occurs in relation to speech errors in the Stroop color-word task. Participants were instructed to overtly name the color of each stimulus as quickly and accurately as possible. Masaki and colleagues found a negative deflection of the ERP signal followed by a positive one shortly after incorrect responses with the same polarity, latency, and scalp distribution as the typical ERN found in motor tasks. Therefore, these authors concluded that ERN-like components can also be found after vocal slips. However, Masaki and colleagues did not apply any manipulations to further investigate whether the ERN after vocal errors shows similar manipulation-dependent alterations in its amplitude and latency as the ERN found after action slips. Furthermore, the Stroop task is a conflict-inducing paradigm and the Stroop effect is not language-specific (for a review see MacLeod, 1991). Therefore, Stroop is a special situation, which may not be representative of general language processing.

In the present study, we will investigate the ERN after errors of the verbal monitor in the presence or absence of a time pressure manipulation. We manipulated time pressure because it has been employed in the ERN as well as in the verbal monitoring literature. Throughout the action monitoring literature, it has consistently been reported that the amplitude of the ERN decreased when time pressure was increased. For example, Gehring and colleagues (1993) used a Flanker task where the speed and accuracy requirements put upon participants were varied. Participants received penalties for errors and rewards for responses faster than a given deadline. Penalties and rewards were varied in such a way that in the speed condition participants responded quickly with little regard for errors, and in the accuracy condition participants responded slowly but more accurately. The results of this study showed that the ERN was largest for the accuracy condition and smallest for the speed condition. Possibly, the representation of the correct response and hence error detection is weaker under high time pressure than in the absence of time pressure (Falkenstein et al., 2000).

Increasing time pressure has also implications for verbal monitoring, more specifically for inner loop monitoring (Oomen & Postma, 2001). According to the perceptual-loop theory, the phonetic plan of the word is temporarily stored in the articulatory buffer. The articulatory buffer serves as the input for the inner loop. The timing relationship between buffer and the articulation stage directly affects the opportunity for the pre-articulatory monitor to timely detect and correct an error (Postma, 2000). In fast speech, buffering is diminished as new output of the formulator is articulated as soon as it becomes available (Oomen & Postma, 2001). Therefore, under time pressure there might be less time to monitor speech and consequently more errors can pass undetected.

Oomen and Postma (2001) investigated how increasing speech rate affects the detection accuracy of the verbal monitor. In their study, participants were presented with visual networks. Networks consisted of colored pictures connected by various lines, with a dot moving along the lines through the network. Participants were required to describe the route of this dot. The rate of describing the movement depended on how fast the dot moved through the network. Oomen and Postma found that speech became more error-prone and less fluent with increased speech rate. However, the percentage of repaired errors was not significantly lower in the fast speech condition than in the normal speech condition. This indicated that the accuracy of error detection, in contrast to production, is not affected by central resource limitations in fast speech (Oomen & Postma, 2001).

In the current study, we investigated not only effects of time pressure on the ERN, but also on workings of the verbal monitor. The task in our study is a phoneme monitoring go/nogo task, previously used in language production and verbal monitoring research (see below). In the phoneme-monitoring task participants are instructed to react to a target phoneme. In the current study participants were required to internally name pictures and press a button when a particular target phoneme occurred in the picture name. For instance, 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 planning.

The phoneme-monitoring task was first used in speech production research by Wheeldon and Levelt (1995). Morgan and Wheeldon (2003) used a similar task to investigate syllable monitoring in internally and externally generated words. Additionally, Schiller (2005) employed the segment-monitoring task to further investigate the phonological encoding processes. Thus, various versions of the phoneme-monitoring task were used to investigate the mechanisms of the verbal self-monitor. We argue that in order to perform this task, participants must monitor their own internal speech by making use of their verbal self-monitoring system. Presumably, however, verbal self-monitoring occurs in a more controlled fashion in the phoneme-monitoring task than in most everyday speech situations.

Our first experiment had three experimental conditions: a control condition (CC), a time pressure 1 (TP1), and a time pressure 2 (TP2) condition. The available response time was manipulated in these conditions; most response time was available in the CC, least in the TP2 condition. Additionally, three lexical retrieval control conditions were added, in which participants were asked to carry out a simple picture naming task with the same time restrictions as in the experimental phoneme-monitoring task. The purpose of these picture naming tasks was to help interpret findings from the experimental conditions (i.e. phoneme monitoring). If more monitoring errors are made during time pressure conditions relative to the control condition, then it is hard to disentangle whether this increase in error rate was due to an incapability of the monitor to detect these errors or due to lexical retrieval failure (i.e. participants not having enough time to retrieve the name of the picture). Therefore, a comparison was made between error rates in the picture naming and the monitoring task.

Additionally, we conducted a second experiment. In this second experiment, we sought to explore whether the effects found in Experiment 1 reflect mechanisms of the monitor or rather result from learning and attention effects. To test this possibility, participants in Experiment 2 performed the same task as in Experiment 1 but without time pressure manipulations. Participants were required to repeat the control condition three times. If participants still become faster and make more errors during the second and third repetitions of the control condition, then the results obtained in Experiment 1 may rather be due to learning and attention effects.

During the entire study, we collected both behavioral and electrophysiological data. As mentioned above, Oomen and Postma (2001) showed that under time pressure more errors were made (though the same percentage of errors was corrected during the time pressure condition as during the control condition). In our study, we also expected to find more errors under time pressure as compared to the control condition. In line with the predictions of the perceptual-loop theory, time pressure might temporarily overload the capacity-limited self-monitoring system and prevent sorting out the competing plans thus leading to more errors (Baars, 1992).

Furthermore, we expected to find slowing of reaction times on correct trials after erroneous responses (i.e. post-error slowing). This would be an important finding because post-error slowing is associated with the initiation of corrective processes (Gehring et al., 1993). Reduction in error slowing (i.e. faster responses after errors) might indicate a dysfunction of the speech monitor. A positive correlation between slowing after errors and performance on post-error trials was also expected. Hajcak, McDonald, and Simons (2003) found, for instance, that participants who showed more slowing after errors also exhibited a better performance on post-error trials.

During the analysis of our EEG data, the Error-Related Negativity (ERN) was of special interest. We expected to obtain an ERN after false alarms (i.e. after participants responded when they should not have responded). During time pressure conditions, we expected to observe a decrease in the amplitude of the ERN, as compared to the amplitude of the ERN during the control condition (see Falkenstein et al., 2001; Gehring et al., 1993). This decrease could potentially mean that the monitor did not have enough time or resources to detect errors.

To summarize, we predicted that participants will make more errors during time pressure conditions than during the control condition. Further, we expected to find post-error slowing and a reduction in this slowing during time pressure conditions. Moreover, we hypothesized to obtain an ERN after erroneous trials across all conditions. However, the amplitude of the ERN should decrease under time pressure. We expected to find none of the above effects in Experiment 2.

Methods

Participants

Twenty-one students of Maastricht University (19 females) participated in Experiment 1 and 20 participants from the same population (18 females) took part in Experiment 2. All participants were right-handed, native Dutch speakers and had normal or corrected-to-normal vision. Participants received course credits or a financial reward for their participation in the experiments. None of them took part in both experiments.

Materials

Eighty-one simple-line drawings were used in this experiment (61 pictures for experimental blocks and 20 pictures for a practice block; see Table 1 for a list of stimuli used in the experimental blocks).

The labels of all the 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), i.e. all picture names had a moderate frequency of occurrence between 10 and 100 per million according to the CELEX database (CEnter for LEXical information, Nijmegen; Baayen, Piepenbrock, & Gulikers, 1995). Furthermore, picture labels all started with consonants. The position of the target phoneme was equated across the stimuli.

Design

Experiment 1 included three experimental conditions: a control condition (CC), a time pressure 1 condition (TP1), and a time pressure 2 condition (TP2). In addition to the experimental conditions, a learning phase and a practice block were administered. Experiment 2 also had three parts, but there was no time pressure manipulation. Instead, response time was identical in each condition.

During the learning phase, participants were familiarized with the pictures and their corresponding names. The names of the pictures were presented auditorily, in order to avoid priming for letters. Then, participants received the practice block, followed by the experimental conditions. In all conditions and after each trial, participants were required to indicate how sure they were about their answer. Participants had to indicate the subjective reliability of their response on a three-point Likert scale that was presented in the middle of the screen after a fixed time interval (1,000 ms) following disappearance of the visual stimulus or after a response to the target picture was made. This scale included the following options: *surely correct*, *do not know*, and *surely incorrect*. However, due to the very low percentage of the incorrect trials during which participants were unaware of their errors (0.17%; on average, there were 0.2 errors during which participants were not aware of their responses), it was impossible to analyze the subjective reliability data statistically.

Table 1. Material employed in the current study.

Target Phoneme							
/t/	/k/	/p/	/n/	/l/	/m/	/s/	/r/
hemd	kom	pan	pan	lamp	kom	mes	muur
pet	broek	plant	nest	film	muur	fles	riem
troon	markt	knop	troon	bloem	riem	slot	dorp
trui	kraan	pet	snor	plant	hemd	nest	trui
baard	kist	kip	knie	naald	bloem	stier	kraan
blad	kip	schaap	pen	plank	mand	schaap	broek
net	wolk	pen	naald	wolk	film	rots	snor
stier	tak	trap	knop	fles	lamp	kist	trap
tak	heks	plank	mand	blad	mes	heks	rots
ster	knie	dorp	net	slot	markt	ster	baard
tram	jurk	schip	band	schaal	maan	fiets	bord
bord	kaars	paard	maan		tram	schaal	rok
fiets	kaart	spoor	kroon			stof	gras
stof	rok	pot	krant			kaas	kaars
kaart	kroon		neus			gras	jurk
trein	krant		schoen			schip	spoor
paard	kruis		hoorn			schoen	hoorn
pot	kraag		ton			neus	kar
band	vork		trein			stok	zwaard
ton	kaas					vuist	vork
kast	kar					kast	kraag
zwaard	stok					kruis	
vuist							

Note. Each stimuli comes twice as a target, but each time with a different target phoneme (e.g., *hemd* ('shirt') has target phoneme /t/ and /m/).

The duration of the stimulus presentation during the control condition was computed separately for each participant, based on their RTs in the practice block. The duration of the stimulus presentation in the control condition was 85% of the RT obtained from the practice block (e.g., if the mean RT during the practice block was 1,000 ms, then the duration of the stimuli in CC was 850 ms). The mean RT of the control condition was used to compute the stimulus duration for TP1 and TP2 conditions in Experiment 1.[3]

3 The reaction times of CC and not of the initial practice block were used for computation of TP1 and TP2 because the average RTs of CC were based on more trials than RTs from the practice block. Participants were also more familiar with the task during CC than during the practice block.

TP1 was 75% of the RT of CC, and TP2 was 60% of CC reaction time (e.g., if stimulus presentation was 850 ms during CC, then the duration of the stimulus of TP1 and TP2 would be 637.5 ms and 510 ms, respectively).[4] Prior to the experimental blocks, in each condition participants were required to repeat a practice block in order to adapt to the new timing. The time between the onset of the picture presentation and the onset of the confidence question was given as response time. Participants were instructed to press a response button prior to the question about their confidence.

Table 2. Lexico-statistical characteristics of the target words.

Target phoneme	Example (English translation)	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

CC, TP1, and TP2 each consisted of eight experimental blocks and one practice block. 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 the computer screen. Experimental blocks consisted of a total of 300 trials (mean 37.5 trials per block; with the exception of a practice block, which consisted of 20 trials). 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 non-target (nogo trials). Each time, participants were asked to monitor for a different phoneme.

4 The percentages for computing the time pressure deadlines (e.g., 75% and 60% of CC) were derived from the outcome of a pilot study.

For instance, for the word *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 non-target, participants were asked to monitor for /l/ and /n/. Before each block, participants received an auditory sample of the phoneme they were required to monitor (e.g., *Reageer nu op de klank /l/ zo als in tafel, spelen, verhaal* 'React now to the sound /l/ like in table, play, tale'; see Figure 2 for a graphical representation of the task).

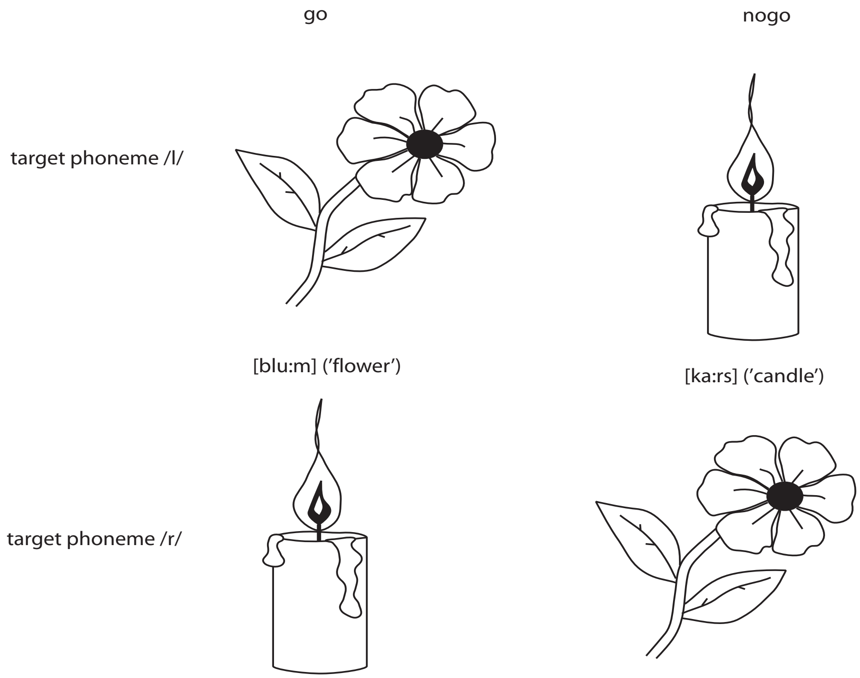

Figure 2. Example of the go and nogo trials for two target phonemes. In the figure, Dutch picture names written in the phonetic code (taken from the CELEX database) and English translations are provided in brackets. Each picture depicted here represents a separate trial. Each picture appeared in the task as a go and as a nogo trial. At the beginning of the block, participants heard for which phoneme they had to monitor.

Procedure

Participants were tested individually while seated in a sound-proof room. They were asked to carry out a learning phase, a practice block, and then the CC, TP1, and TP2 conditions in Experiment 1. In Experiment 2, participants carried out the CC three times. Prior to each condition, participants were required to carry out a picture naming task.

During all blocks, participants were required to press a button if a target phoneme was 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. nogo trials). Button-press latencies were recorded from the onset of the picture.

In the picture naming task, participants saw the same pictures that were used in the phoneme monitoring task and were requested to overtly name them as fast as possible. The picture naming task was also divided into three conditions, i.e. control condition, time pressure 1, and time pressure 2. The set up of this task was identical to the phoneme monitoring task. The purpose of the picture naming task was to assure that participants had enough time to access the name of the picture in the given time window. Participants were instructed to sit as still as possible and to suppress eye blinks while a picture was on the screen and during button presses.

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 band-pass filter from 0.05 to 30 Hz. An electrode at the left mastoid was used for on-line 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 canthus. Eye blinks and vertical eye movements were measured using a bipolar montage of two electrodes placed above and below the left eye. Impedance level for all electrodes was kept below 5kΩ.

Data analysis

Epochs of 1,300 ms (–400 ms to +900 ms) were computed. A 200 ms pre-response baseline was used. The EEG signal was corrected for vertical EOG artifacts, using the ocular reduction method described in Anderer, Satety, Kinsperger, and Semlitsch (1987). The ERN was measured in response-locked ERP averages. For the ERN, averaging was carried out across error trials (i.e. false alarms). For the correct trials, averaging was done for correct go-responses. The amplitude and latency of the ERN was derived from each individual's average waveforms after filtering with a band pass, zero phase shift filter (frequency range: 1 – 12 Hz). The amplitude of the ERN was defined as the difference between the most negative peak in a window from 50 to 150 ms after the response and the most positive peak of the signal from 0 to 50 ms after response onset. The latency of the ERN was defined as a point in time when the negative peak was at its maximum (Falkenstein et al., 2000). The amplitude and latency of the ERN were recorded for each condition (CC, TP1, and TP2) at the following electrode sites: Fz, FCz, Cz, and Pz.

Results

Results – Experiment 1

Behavioral data

Reaction times and error rates

Repeated measures analyses of variance (ANOVAs) were run with Time Pressure as independent variable. Reaction Times (RTs) smaller than 300 ms and larger than 1,500 ms were excluded from the analysis. Mean RTs per are provided in Table 3.

As predicted, RTs were longer during CC, faster during TP1, and fastest during TP2 ($F(2, 38) = 111.24$, $MS_e = 2461.48$, $p < .001$). This decrease in RTs can be interpreted as an increase in participants' efficiency in executing the task. However, if this were true, one would also expect to find fewer errors under time pressure, but the opposite was obtained (see the detailed error analysis below). Hence, it seems reasonable to assume that the experimental task manipulation was successful in inducing time pressure.

Table 3. Overview of behavioral data. Mean (± standard deviation) reaction times (in ms), error rates (%), and post-error slowing (in ms) as a function of time pressure manipulations.

	Control condition	*Time Pressure 1*	*Time Pressure 2*
Reaction times	769 (91)	619 (83)	584 (78)
Error rate	2.6 (11)	4.7 (14)	4.9 (12)
Post-error slowing			
Post-error trials	849 (177)	630 (113)	572 (101)
Post-correct rials	724 (82)	576 (63)	548 (86)

Similar analyses were performed with Time Pressure as independent variable and the number of errors as dependent variable. There was a significant main effect of Time Pressure ($F(2, 38) = 14.44$, $MS_e = 32.25$, $p < .001$; see Table 1 for mean error rates). Overall, participants made more errors during the time pressure conditions than during the control condition. A paired t-test showed that participants made significantly more errors during TP1 as compared to CC (Bonferroni adjusted α-level = .016; $t(19) = 5.50$, $SD = 6.37$, $p < .001$). Participants also made more errors during the TP2 condition than during the TP1 condition, but this difference was not significant ($t(19) < 1$).

To investigate whether or not participants had enough time during TP conditions to retrieve the name of the pictures from the lexicon, a repeated measures ANOVA was run for the picture naming task with Time Pressure as independent variable and number of errors as dependent variable.[5] This analysis showed a significant main effect of Time Pressure ($F(2, 38) = 4.74$, $MS_e = 2.81$, $p < .05$). However, results of the picture naming task are the reverse of the ones obtained in the phoneme monitoring task. Participants made more errors during CC (0.59%) than during TP1 (0.25%) and TP2 conditions (0.21%). This difference could presumably be attributed to participants' inefficiency and unfamiliarity with the stimuli during the CC as compared to the TP conditions. Given that fewer errors were made in naming during the TP conditions than during CC, the effects found in the phoneme monitoring task are likely to be due to the malfunctioning of the verbal monitor under time pressure and not due to lexical retrieval failure.

Post-error slowing

Trials after errors were used for the analysis of the error-related slowing (Gehring et al., 1993; Hajack et al., 2003; Rabbit, 1981). During the task, in every condition each picture was presented twice as a go-trial, which allowed us to select button-press latencies for the same pictures for a post-error trial and a correct trial. For example, if a correct response was given for the picture *heks* ('witch') and this trial appeared after an error, then it was selected for the analysis as a post-error trial. Moreover, if for the same picture a correct response was given which was preceded by another correct response, then the former was selected as a correct trial. Correct trials after errors were compared with correct trials after correct responses. An ANOVA revealed a significant main effect of Post-Error Trial ($F(1, 16) = 9.21$, $MS_e = 12{,}660$, $p < .001$). As expected, participants were slower on post-error trials than post-correct trials. Furthermore, a significant Post-Error Trial by Conditions interaction was found ($F(2, 32) = 4.34$, $MS_e = 14{,}165$, $p < .05$; see Table 1 for an overview of RTs). Further investigation of the interaction revealed a post-error slowing effect for CC ($F(1, 16) = 8.71$, $MS_e = 15{,}121$, $p < .01$). For TP1, the effect of post-error slowing was marginally significant ($F(1, 16) = 4.12$, $MS_e = 5{,}959$, $p < .06$). Finally, in the TP2 condition, there was a trend towards post-error slowing effect, which did not reach significance ($F(1, 16) = 2.71$, $MS_e = 1{,}911$, *n.s.*).

The post-error slowing may possibly be related to corrective processes (Gehring et al., 1993). Therefore, it is plausible to assume that there is a relationship between post-error slowing and the number of errors. To investigate this, Pearson correlations were computed. There was a negative correlation between the number of errors and error-related slowing ($r = -.60$, $p < .001$) indicating that larger post-error slowing was associated with fewer errors. This finding is in accordance with the hypothesis that post-error slowing may reflect corrective processes.

---

5 The purpose of the PN task was to control whether participants could correctly retrieve the name of the picture, despite the presence of time pressure. Therefore, incorrect responses were of most interest for the analysis. Unfortunately, due to technical failure, we could only analyze the naming latencies of 13 participants. These participants named the pictures significantly faster under time pressure than in the control condition (CC: 653 ms; TP1: 575 ms; TP2: 563 ms; $F(2,24) = 5.54$, $MSe = 5{,}614$, $p < .05$).

Electrophysiological data

ERN descriptives

The ERN was revealed in response-locked ERP averages for false alarms. There was no negative deflection observed in the ERP waveforms for correct trials during visual inspection of the EEG waves. A more detailed description of the ERN is given below.

Latency and amplitude analysis

A repeated measures ANOVA was employed with Time Pressure as independent variable and ERN peak latency as dependent variable. This analysis showed no effect of Time Pressure ($F(2, 34) < 1$). The ERN peaked independently of condition at approximately 75 ms after the error was committed. Similar analyses were run to investigate the effect of Time Pressure on the amplitude of the ERN. The analysis revealed a significant effect of Time Pressure ($F(2, 34) = 4.23$, $MS_e = 61.16$, $p < .01$), reflecting the fact that the amplitude of the ERN was smaller during TP2 than during TP1 and CC (see Figures 3 and 4). In addition, there was a significant Electrode Site by Time Pressure interaction ($F(6, 102) = 3.22$, $MS_e = 8.91$, $p < .01$). Follow-up analyses of this interaction revealed that Time Pressure had an effect on the amplitude of the ERN only at electrode sites FCz and Cz ($F(2, 34) = 4.03$, $MS_e = 18.99$, $p < .05$ and $F(2, 34) = 5.19$, $MS_e = 25.48$, $p < .05$, respectively), but not at sites Fz and Pz ($F(2, 34) = 2.78$, $MS_e = 16.55$, *n.s.* and $F(2, 34) = 3.32$, $MS_e = 26.88$, *n.s.*, respectively).

Interestingly, the stimulus-locked ERP averages also showed a negative deflection for incorrect go trials (i.e. misses). This negativity was absent during correct nogo trials (i.e. correct rejections). The negative deflection observed for misses peaked approximately at the time of a potential response (see Figure 5). This may be interpreted as an indication of an ERN-like ERP component in the absence of an overt motor response (i.e. button-press) suggesting that the ERN for verbal errors does not depend on incorrect motor responses.

A mean area analysis was used to test whether there was a significant difference between misses and correct rejections. Time windows of interest were derived based on the visual inspection of the grand average waveforms. In such a way, time windows of interest for CC, TP1, and TP2 were 580 – 720 ms, 570 – 670 ms, and 550 – 650 ms, respectively. A 4 (electrodes) by 2 (correct vs. error) ANOVA was run with Correctness of Response as independent variable and the amplitude of the ERN-like response as dependent variable. The analysis showed a significant difference between correct and erroneous responses for CC, TP1, and TP2 ($F(1, 17) = 18.08$, $MS_e = 10.78$, $p < .01$; $F(1, 17) = 30.13$, $MS_e = 5.09$, $p < .01$; and $F(1, 17) = 128.48$, $MS_e = 2.41$, $p < .01$, respectively). In all conditions, erroneous responses had more negative amplitudes compared to correct responses. The amplitude of the ERN after misses showed the same pattern as the ERN after false alarms. Specifically, the amplitude of the negative deflection after misses was more negative during the control condition than during time pressure conditions ($F(2, 34) = 7.93$, $MS_e = 10.54$, $p < .001$; see Figure 4).

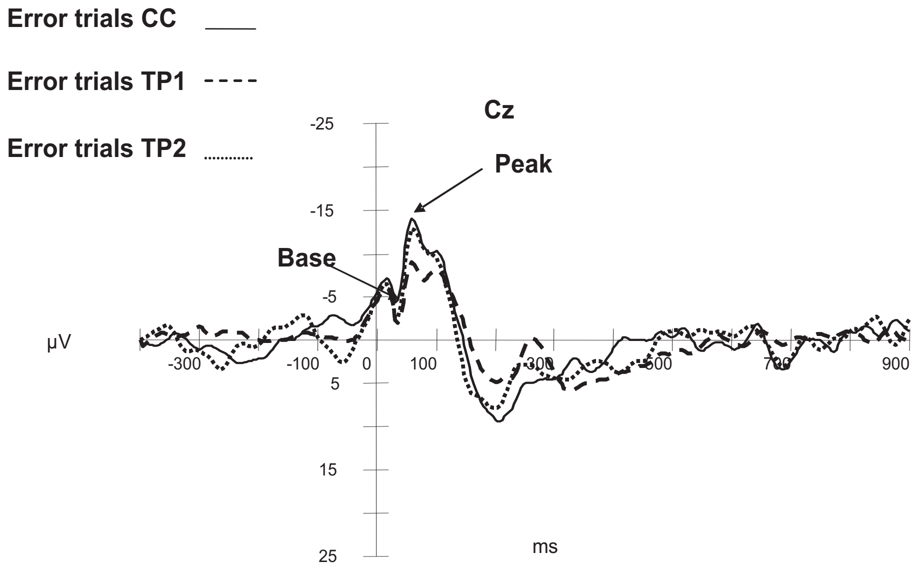

Figure 3. ERP waveforms for a single participant across conditions (CC – control condition, TP1 – time pressure 1 condition, and TP2 – time pressure 2 condition). The amplitude of the ERN is lower at the TP2 condition than during the CC and TP1.

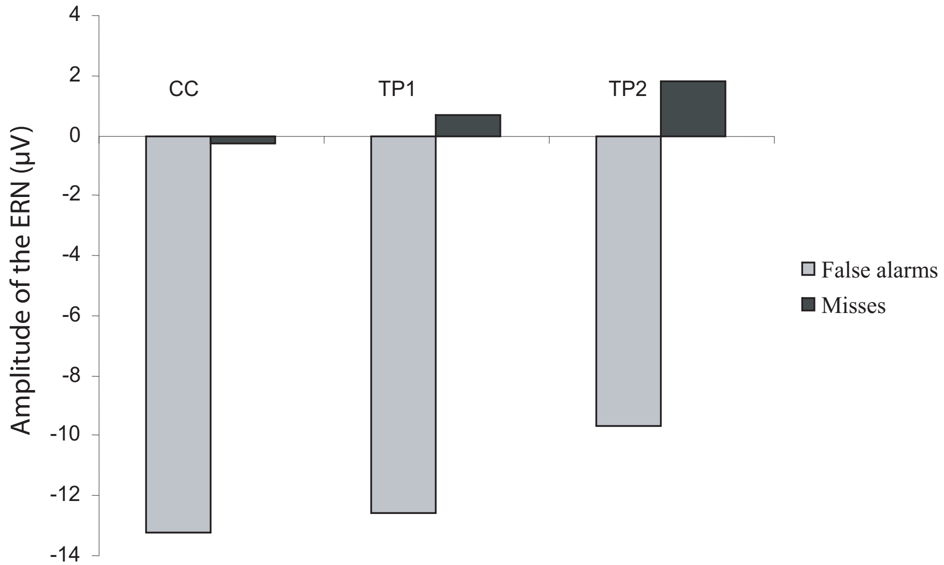

Figure 4. The ERN amplitude across conditions (CC – control condition, TP1 – time pressure 1 condition, and TP2 – time pressure 2 condition) for false alarms and misses.

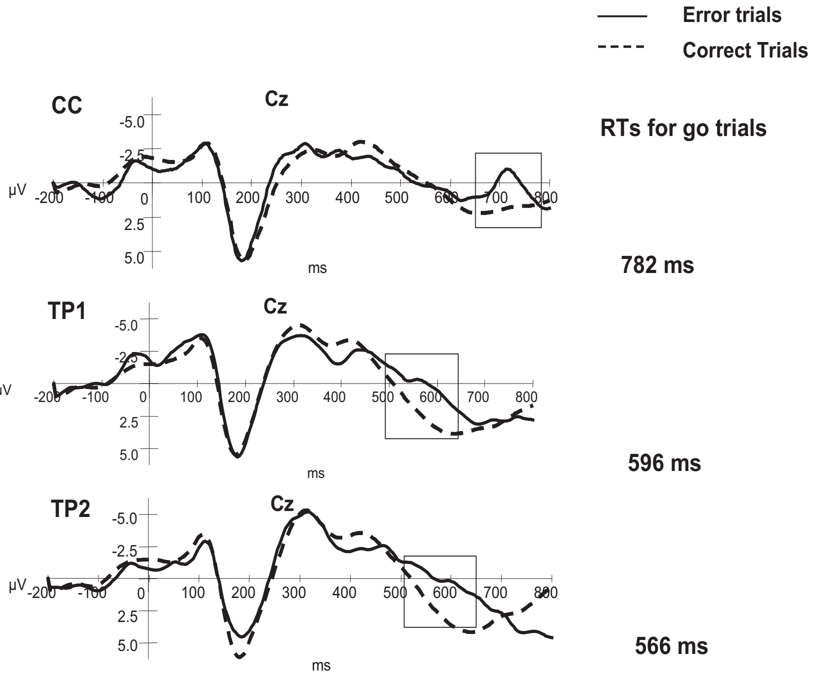

Figure 5. Averaged stimulus-locked ERP waveforms for all misses (solid lines) versus correct rejections (dashed lines) across conditions (CC – control condition, TP1 – time pressure 1 condition, TP2 – time pressure 2 condition). Waveforms depicted for the Cz electrode, where amplitude of the ERN was largest. Correct and incorrect trials were matched on RTs and number of trials. Areas selected by the rectangle depict a time window of average button-press latency for correct responses. Mean RTs for go trials are provided.

Results – Experiment 2

The analysis showed that participants do in fact become faster during the second and third block of the task compared to the first block ($F(2, 36) = 76.49$, $MS_e = 551.93$, $p < .001$). During the first time (CC1), participants did the task slower (746 ms, *SD* = 86) than during the second (CC2; 675 ms, *SD* = 80) and third time (CC3; 656 ms, *SD* = 76). However, contrary to what we observed in Experiment 1, participants do not make significantly more errors in repetitions of the control condition ($F(2, 36) = 2.61$, $MS_e = 17.58$, *n.s.*). During CC1, participants made 5.2% (*SD* = 9) errors and during CC2 and CC3 they made 6.2% (*SD* = 11) and 6.4% (*SD* = 7), respectively. Furthermore, there were no significant differences in the amplitude of the ERN between the first, second, and third time participants performed the task ($F(2, 36) = 1.81$, MS_e

= 78.91, *n.s.*). The amplitudes of the ERN for CC1, CC2, and CC3 were –7.68 μV, –10.02 μV, –10.09 μV, respectively. In contrast, in Experiment 1 we showed that the amplitude of the ERN was significantly lower in the time pressure conditions than in the control condition. Thus, even though participants became faster in Experiment 2 during the repetitions of the control condition, they did not make more errors and the amplitude of the ERN remained unaffected by repetitions of the task. Therefore, we conclude that results of Experiment 1 cannot be fully attributed to attention and learning effects, but are more likely to be due to the malfunction of the verbal self-monitor.

Discussion

The present study aimed at investigating the electrophysiological correlates of verbal monitoring in the presence and absence of time pressure. Previously, it has been shown that verbal monitoring might be affected by the presence of time pressure (Oomen & Postma, 2001). The ERN is also known to be sensitive to time pressure manipulations (e.g., Gehring et al., 1993). In line with our predictions, we found that participants made more errors and showed a decrease in amplitude of the ERN under severe time pressure.

The main manipulation used 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 is more important. Such conditions are rather common and people can often control their level of SAT. Recently, it has been proposed that SAT is controlled by changing the duration of a stage that verifies the already selected and prepared response (Osman, Lou, Muller-Gethmann, Rinkenauer, Mattes, & Ulrich, 2000). Specifically, Osman and colleagues suggested that people may select one response alternative after a tentative decision and then re-check the selected response. Slow, but accurate performance would result when the final execution of the response was withheld until re-checking was completed. Speed stress would shorten the RT interval and decrease accuracy by inducing participants to skip or reduce re-checking. Interestingly, in the computational implementation of Levelt's model of speech production (WEAVER++; see Levelt, Roelofs, & Meyer, 1999; Roelofs, 1992) it is argued that speech errors may occur when WEAVER++ skips verification to gain speed in order to obtain a higher speech rate.[6] Thus, more errors are to be expected at higher speech rates (Levelt et al., 1999). It seems plausible to assume that the shifting along the speed-accuracy continuum might to some extent be controlled by a monitor.

It is likely that errors observed in the current study resulted from the substitution, addition, or deletion of phonemes. For example, in the word *kaart* 'card' the phoneme /t/ could have

6 Verification is a binding-by-checking process. Each node in the speech production network has a procedure attached to it that checks whether the node, when active, links up to the appropriate active node one level up (Levelt et al., 1999).

been substituted by phoneme /s/ which would have resulted in *kaars* 'candle'. Similarly, if the phoneme /r/ was deleted from *kaars* 'candle', it would have resulted in *kaas* 'cheese'. All these examples would lead to an inaccurate decision about the presence or absence of the target phoneme in the name of the picture. As mentioned above, the time pressure manipulation resulted in a higher error rate than the no-time pressure manipulation. It is possible that time pressure resulted in the reduced monitoring.

Why is verbal monitoring affected by time pressure? According to Levelt (1989), verbal self-monitoring is a controlled process, and therefore resource-limited. Controlled processes require resource allocation (Shiffrin & Schneider, 1977). If there are not enough resources available, controlled processes do not function at an optimal level. Hence, if there are insufficient resources available for verbal monitoring activities, then functioning of the monitor will not be optimal and this may potentially lead to more errors. In terms of Levelt's (1989) model, it is possible that under time pressure the inner loop has less time to monitor the phonetic plan. Under such conditions, more errors pass undetected or corrective processes are not activated fast enough. It seems reasonable to hypothesize that in the current study under time pressure verbal self-monitoring was reduced. This, in turn, may support the idea that verbal self-monitoring is a resource-limited process, given that time pressure decreases the amount of resources and time available for the functioning of the monitor.

Previous research revealed that under time pressure conditions more errors were made (Oomen & Postma, 2001). In the same study, however, Oomen and Postma also showed that the same percentage of errors was corrected during a time pressure condition as during a control condition. In the present study, the experimental task did not give participants the opportunity to correct their errors. However, during time pressure conditions an error-related slowing was found, which might be interpreted as a form of corrective action (see Gehring et al., 1993). Furthermore, participants who showed larger error-related slowing made overall fewer errors. This is in line with previous findings (e.g., Hajack et al., 2003).

As mentioned above, our electrophysiological data are in line with our predictions and previous findings. Specifically, the amplitude of the ERN decreased under time pressure. The button presses in our study were dependent on a decision about the presence or absence of the target phoneme in the name of the picture. It is likely that the verbal monitor monitors an abstract phonological representation by checking for mismatches between intended and actual verbal responses. Thus, the verbal outcome is compared with the original intention, and if there is a mismatch, then an error is detected. This is in accordance with the *perceptual-loop theory* by Levelt (1983). Similarly, the action monitor compares the representation of the correct response with the copy of an on-line response. If there is a mismatch between actual and intended response, an error signal is generated (e.g., Desmurget & Grafton, 2000). Under time pressure, there might not be enough time available to make an optimal comparison between correct 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.

It seems that there is conceptual overlap between verbal monitoring and general performance monitoring theories. As stated above, both monitoring theories independently state that in order to detect an error a monitor compares the representation of a correct response with the copy of an on-line response. In the present study, we showed a typical ERN in a task where performance is dependent on a verbal judgment. Additionally, there is further recent evidence that errors during verbal tasks activate the anterior cingulated cortex (ACC) and medial frontal cortex (SMA; Möller, Jansma, Rodriguez-Fornells, & Münte, 2007). This latter result is in accordance with the claim that the ERN is generated within the ACC/SMA region (Dehaene et al., 1994). Based on this evidence, we suggest that the verbal monitoring is not a process separate from but rather a special case of general performance monitoring.

Interestingly, in the present study we also demonstrated an ERN-like response on incorrect go-trials (i.e. misses). This negativity was present only after misses and not during the correct nogo trials. It had the same characteristics as a typical ERN, i.e. it peaked at fronto-central sites and it initiated at the time of average response latency. Additionally, the negative deflection after misses was affected by time pressure in a similar way as the ERN after false alarms. In other words, the amplitude of the negativity after misses decreased under time pressure, as compared to its amplitude during the control condition. For these reasons, we think that the negative deflection after misses can be interpreted as an ERN-like response. This is particularly interesting since the literature on the ERN available so far mainly reports a negative deflection after *overt* motor errors. Critically, however, misses are errors where such an overt motor response is absent.

This finding is not necessarily in disagreement with existing theories about the ERN. For instance, during misses participants failed to detect the target phoneme in the name of the picture. It is possible that after participants made that decision, further processing of the stimuli revealed that there actually was a target phoneme in the name of the picture. This, in turn, resulted in the mismatch between actual and desired response, which led to a higher conflict during the miss-trial as compared to the correct nogo response. Hence, the ERN is generated.

Interestingly, Luu, Flaisch, and Tucker (2000) showed that the ERN can also be elicited in late responses. Luu and colleagues used a deadline reaction task in which participants were told to respond within a given time interval or the response will be considered late and scored as an error. Luu and colleagues found that as the responses become increasingly late, the self-monitoring of these responses became increasingly strong and the amplitude of the ERN increased linearly. Similar, in our study it is possible that at least during some of the misses participants were uncertain about their response (i.e. had difficulty distinguishing error and correct responses), which in turn led to missing the response deadline. Thus, participants became increasingly aware of making an error as the response became increasingly late and eventually missed the deadline. However, there is a crucial difference between the late responses of Luu and colleagues and the misses of our study. The late responses of the Luu et al. study varied from those barely missing the deadline to ones that were made much later, but nevertheless the overt motor response was made. In contrast, we found similar ERN-like responses in the absence of any overt button presses.

Overall, our findings provide evidence that verbal self-monitoring is a resource-limited process supporting the *perceptual-loop theory* (Levelt, 1983, 1989). More importantly, we demonstrated a link between the verbal self-monitoring system and the ERN. The ERN after verbal errors reflected the same changes afflicted by enhanced time pressure as the ERN found in typical action monitoring studies. Therefore, we conclude that the processes of verbal monitoring might be analogous to the processes of action monitoring.

This provides researchers with a new useful tool for psycholinguistic research. For example, one of the major problems with studying covert errors in language production is that one can never be sure about compliance of the participants. By using the ERN as an electrophysiological marker, this problem can be eliminated. However, more research is needed to make a clear cut separation between errors of verbal monitoring and errors of other action monitoring. We believe that most of the errors found in the current study are the result of incorrect decisions about the presence or absence of the target phoneme in the picture name. However, it cannot be completely excluded that some of the errors arose from action slips and are not slips of verbal monitoring *per se*.

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.

Baars, B. (1992). A dozen competing-plans techniques for inducing predictable slips in speech and action. In B. Baars (Ed), *Experimental Slips and Human Error*. (pp. 129-150). New York: Plenum Press.

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.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D. C., & Cohen, J. D. (1998). Anterior cingulated cortex, error detection, and the online monitoring of performance. *Science*, 280, 747-749.

Carter, C.S., MacDonald III, A.W., Ross, L.L., & Stenger, V.A. (2001). Anterior cingulated cortex and impaired self-monitoring of performance in patients with schizophrenia: an event-related fMRI study. *American Journal of Psychiatry*, 158, 1423-1428.

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.

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science*, 5, 3-23.

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

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.

Forman, S. D., Dougherty, G., Casey, B. J., Siegle, G., Braver, T., Barch, D., Stenger, V. A., Wick-Hull, C., Pisarov, L. A., & Lorensen, E. (2004). Opiate addicts lack error-dependent activation of rostral anterior cingulated. *Biological Psychiatry*, 55, 531-537.

Gehring, W. J., Himle, J., & Nisenson, L. G. (2000). Action-monitoring dysfunction in obsessive-compulsive disorder. *Psychological Science*, 11, 1-6.

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.

Hajcak, G., McDonald, N., & Simons, R. F. (2003). To err is autonomic: error-related brain potentials, ANS activity, and post-error compensatory behavior. *Psychophysiology*, 40, 895-903.

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.

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.

Levelt, W. J. M., Roelofs, A., & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral Brain Science*, 22, 1-75.

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.

Luu, P., Flaisch, T., & Tucker, D. (2000). Medial frontal cortex in action monitoring. *Journal of Neuroscience*, 20, 464-469.

MacLeod, C. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109, 163-203.

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

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.

Morgan, L. J., Wheeldon, L. R. (2003). Syllable monitoring in internally and externally generated English words. *Journal of Psycholinguistic Research*, 32, 269-296.

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.

Osman, A., Lou, L., Muller-Gethmann, H., Rinkenauer, G., Mattes, S., & Ulrich, R. (2000). Mechanisms of speed-accuracy tradeoff: evidence from covert motor processes. *Biological Psychology*, 51, 173-199.

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.

Rabbit, P. M. A. (1981). Sequential reactions. In D. Holding (Ed), *Human Skills*, Wiley, New York, pp. 153-175.

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.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107-142.

Scheffers, M. K., Coles, M. G. H., Bernstein, P. S., Gehring, W. J., & Donchin, E. (1996). Event-related brain potential and error-related processing: and 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.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127-190.

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.