Universiteit Leiden
The Netherlands

# The nature of the verbal self-monitor

Ganushchak, A.Y.

## Citation

Ganushchak, A. Y. (2008, March 12). *The nature of the verbal self-monitor*. Retrieved from https://hdl.handle.net/1887/12635

Version: Not Applicable (or Unknown)

License: Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden

Downloaded from: https://hdl.handle.net/1887/12635

**Note:** To cite this publication please use the final published version (if applicable).

# The nature of the verbal self-monitor

by

Aleksandra Yurievna (Lesya) Ganushchak

Cover design: Yuri and Lesya Ganushchak

“From error to error one discovers the entire truth” – Sigmund Freud (1856 – 1939; ref. Dictionary of Quotations (1998). Ed. C. Robertson, Wordsworth Editions, Ltd, UK, p. 132).

Layout: Lesya Ganushchak
Printed by Ponsen & Looijen B.V.
ISBN 978-90-6464-226-5

# The nature of the verbal monitor

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden
op gezag van Rector Magnificus prof. mr. P. F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 12 maart 2008
klokke 13.45

door

Aleksandra (Lesya) Yurievna Ganushchak
Geboren te Nikolaev (Oekraïne)
in 1978

**Promotor**
Prof. Dr. Niels O. Schiller

**Referent**
Dr. Robert Hartsuiker (Universiteit Gent)

**Overige leden van de promotiecommissie**
Prof. Dr. Richard Ridderinkhof (Universiteit Amsterdam)
Prof. Dr. Bernhard Hommel
Dr. Bernadette Jansma (Universiteit Maastricht)
Dr. Wido La Heij

# Contents

# Chapter 1

# Introduction

> *"The inner workings of a highly complex*
> *system are often revealed by the way*
> *in which the system breaks down…"*
> *Dell, 1986, p, 284*

Language is a crucial part of our everyday lives. For most of us, there is not a day that passes without exposure to some form of language. Speaking is very fast and seemingly effortless process. In speaking aloud, we produce up to 150 words per minute. However, the speech error rate in normal individuals is not more than 1 error in every 1000 words (Levelt, 1989). Such a low error rate indicates that there must be a monitoring system that checks for errors and corrects them if any are found. It is very important to monitor one's own speech, since 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). For example, saying "I want to spank all the thinkers" instead of "I want to thank all the speakers" (Fromkin's Speech Error Database from the Max Planck Institute for Psycholinguistics). 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).

This thesis focuses on investigating various factors that interfere with the working of the verbal monitor in normal individuals. Chapter 2 investigates the effects of time pressure on verbal self-monitoring. In the reported EEG experiment, participants were required to perform a phoneme-monitoring task (see Chapter 2 for detailed description of the task) with and without a time pressure manipulation.

Chapter 3 focuses on how verbal self-monitoring is affected by time pressure when a task is preformed 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.

Chapter 4 investigates how auditory distractors affect workings of the verbal self-monitor. In the described experiment, participants were asked to perform a phoneme-monitoring task. The target stimuli were presented simultaneously with auditory distractors. EEG was recorded through out the whole experiment.

Chapter 5 examines whether motivation and semantic context affects participants' performance using a picture naming task in a semantic blocking paradigm. The semantic context of to-be-named pictures was manipulated; blocks were semantically related (e.g., cat, dog, horse) or semantically unrelated (e.g., cat, table, flute). Motivation was manipulated independently.

Chapter 6 addressed how the Error-Related Negativity (ERN) is affected by conflict in a bilingual context. Dutch-English bilinguals saw Dutch words in white print that needed to be classified according to their grammatical gender and colored words that were to be classified on the basis of their color. Colored words included Dutch common and neuter words, and English translations of those words. EEG was recorded through out the whole experiment.

Finally, in Chapter 7, a short summary is presented, and main findings are discussed.

## 1.1 Verbal self-monitoring in speech production

To this date, the most influential and detailed speech production theory is the theory proposed by Levelt and colleagues (Levelt, 1989, 1999a, 1999b, 2001, Levelt, Roelofs, & Meyer, 1999; see Figure 1). According to this model, speech production consists of a number of steps. First of all, *conceptual preparation* takes place, i.e. the message is generated. A decision is made about what to say and in what order, also word choice and grammatical roles are determined. The output of the conceptual preparation is a preverbal message. The second step of speech production is called *formulation*. Formulation can be divided into *grammatical encoding* and *form encoding*. During grammatical encoding, the syntactic aspects of words are retrieved from the mental lexicon. The mental lexicon stores information about words such as meaning and form. During form encoding, word forms are retrieved based on the output of grammatical encoding. Phonological encoding can start after the word form has been accessed and eventually leads to the phonological representation (i.e. segmental shape and metrical structure) of the morphemes. During *phonetic encoding*, this phonological representation is transformed into a phonetic one, which specifies articulatory commands. Finally, the fully encoded speech plan can be articulated.

One of the crucial processes in speech production is self-monitoring. Speech monitoring or verbal self-monitoring is a process that checks the correctness of the speech flow. Its prime purpose is to detect and correct speech production errors, parts of the speech program or of the actual speech output that do not agree with the speaker's communication purpose or with his/her general linguistic knowledge and standards (Postma, & Kolk, 1993). The monitor, according to Baars (1975) is a mechanism that "listens to" self-produced internal or external feedback, compares this with the intended output, identifies errors, and then computes corrections by using a duplicate copy of the information originally available to the motor system.

To this date, the most influential and detailed verbal self-monitoring theory is the theory proposed by Levelt and colleagues (Levelt, 1983, 1989, 1999a, 1999b, 2001, Levelt, Roelofs, & Meyer, 1999; see Figure 1). The perceptual loop theory suggests one central monitor localized in the conceptualizer (see Figure 1). According to this theory, only certain-end products in the speech production are monitored. Moreover, these end-products are analyzed in a similar way as a speech of other, in other words, through speech comprehension system (Postma, 2000). A distinction is made between *external* and *internal monitoring*. External monitoring is monitoring of speech after it has been articulated and proceeds through the auditory loop, i.e. the signal enters the auditory system and is then processed by the speech comprehension system where the information is parsed and then sent to the *conceptualizer*. In contrast, internal monitoring is covert monitoring of speech production, i.e. monitoring that occurs prior to articulation and has access to more abstract codes, i.e. the phonological planning level (Schiller, 2005, 2006; 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 function of this inner loop is to inspect the speech plan prior to its articulation (Postma & Noordanus, 1996). The internal monitoring can also proceed through auditory monitoring (e.g., 'listening' to inner speech; Postma, 2000). Finally, there is also a *conceptual loop* (between the preverbal message and the conceptualizer) that checks for semantic appropriateness.

There is empirical evidence supporting the existence of internal monitoring. For instance, it has been found that some repairs in overt speech have a 0 ms cut-off to repair interval (Blackmer & Mitton, 1991). In other words, some speech errors are corrected immediately after having been produced, which indicates that the error must have been detected and repair processes had to be initiated before the word was completely pronounced. Similarly, there are also corrections made in speech that occur after the pronunciation of only the first phoneme of an incorrectly selected word, e.g., *v..horizontal* (Levelt, 1983). In such cases, it is hypothesized that internal monitoring detected the error before overt production, but too late to stop articulation of the initial phoneme.

Furthermore, the perceptual loop monitor is centrally regulated and requires controlled processing. It is generally accepted that controlled processes are capacity-limited and depend upon allocated resources (Shiffrin & Schneider, 1977). Therefore, it is plausible to assume that speech monitoring is a resource-limited process (Postma, 2000). In other words, if resources are allocated to other than monitoring activities, monitoring will not operate on an optimal level. Additionally, the perceptual loop theory maintains that speakers are usually aware of the monitoring process (i.e., monitoring is a conscious process).

In sum, in the perceptual-loop theory, actual error detection takes place at the level of conceptualization, where parsed speech is compared to the intention and linguistic standards and occurs as the speech comprehension system signals some kind of irregularity (Hartsuiker & Kolk, 2001; Oomen & Postma, 2002). This single monitor located within the conceptualizer handles appropriateness monitoring, inner speech monitoring, and external monitoring.

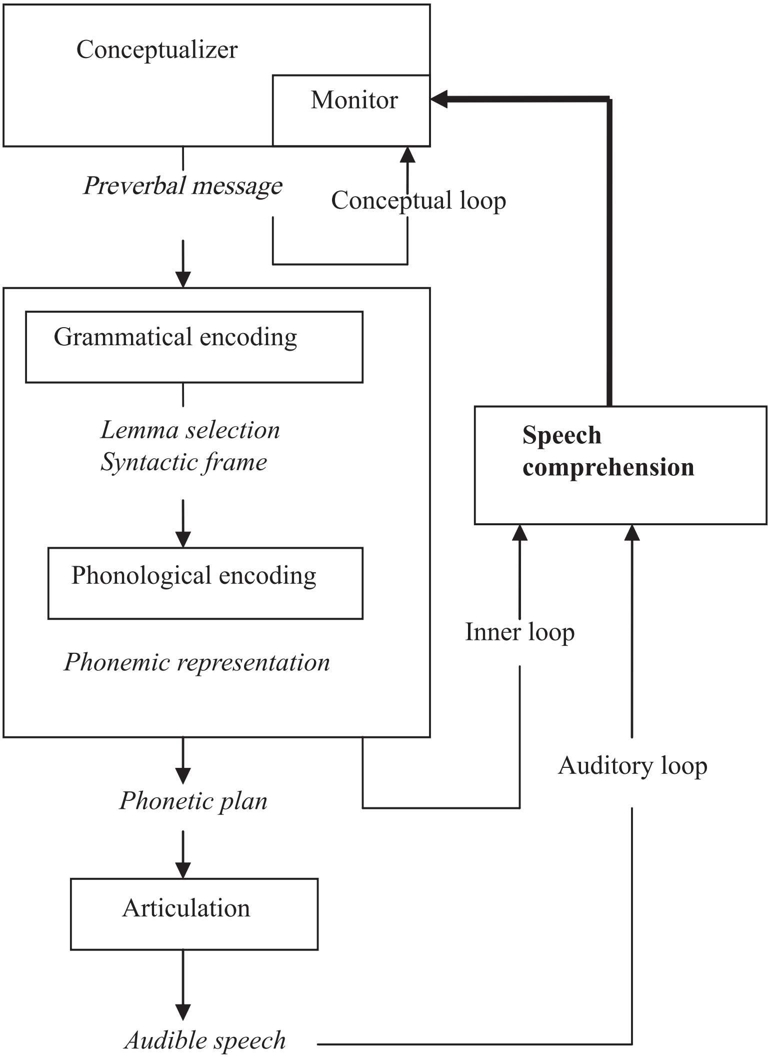

Figure 1. Graphical representation of Levelt's speech production model.

The production-based approach is an alternative account for the self-monitoring process. According to this approach, self-repair processes have access to various stages of speech production (Laver, 1980). In other words, immediate aspects of speech planning, such as components *inside* the formulator, are accessible for monitoring. The production-based view does not limit itself to one monitor, on the contrary, it states that monitors are widely distributed throughout speech production levels. If an error has been detected, further processing is cancelled. This error interruption usually occurs within 150 ms (Oomen & Postma, 2002). The production-based approach, unlike the perceptual loop theory, argues that monitoring is not largely dependent on the resources allocated to it. It is possible that some of the production-based monitors possess their own specialized resources (Postma, 2000). Therefore, one would expect that resource limitations will not lead to a less accurate working of the self-monitor. Another crucial difference between the production-based approach and the perceptual loop theory is that the production-based view states that self-monitoring is an unconscious process.

In this section, two approaches to self-monitoring during speech production were described: the perceptual loop theory and the production-based theory. Both of these theories have been subject to criticism. For instance, the perceptual loop theory claims that all monitoring proceeds through the comprehension system. However, research with aphasic patients showed that some of the patients have an intact comprehension system but impaired monitoring, and the reversed was also found to be true (Marshall, Rapparport, & Garcia-Bunuel, 1985; Marshall, Robson, Pring, & Chiat, 1998; for overview see Oomen, Postma, & Kolk, 2001). This indicates that the comprehension system is not crucial for monitoring. Furthermore, Oomen and Postma (2001) showed that error detection and repair processes do not suffer from time pressure (i.e. speech has to be produced at higher than normal speed). This is not in accordance with one of the predictions of the perceptual loop theory, which holds that under time pressure less time will be available for inner loop monitoring. The production-based monitors can adjust their functioning based on speech rate. However, one of the main objections to the production-based approach is that it assumes the assistance of separate monitors working on each level of production. Such a system might not be economical, since for optimal functioning it would require several additional monitoring devices (Wheeldon & Levelt, 1995). Despite the differences in described theories, they both agree that if the monitor is not optimally functioning, then more errors will pass undetected and uncorrected. Therefore, under more challenging conditions when monitoring is impaired more errors will be committed.

## 1.2 Speech errors

Speech errors are generally assumed to occur during the construction of an internal representation of the utterance at a semantic, syntactic, phonological, or motor program level. Due to speech planning errors or the incorrect execution of correctly planned programs, for instance, speakers produce errors such as “laboratory in my own computer” instead of “computer in my own laboratory” (Fromkin, 1971; see also Garrett, 1975, 1980).

Errors in which segments, features, or clusters are disordered, deleted or added are related to processes that link phonological units with slots in phonological frames. Examples of such errors are “at the right tace” (intended utterance “at the right time and place”) or “a brun” (instead of “a bread bun”; these examples are taken from the Fromkin’s Speech Error Database from the Max Planck Institute for Psycholinguistics).

According to Baars (1992), speech errors arise from competition between alternative speech plans, time pressure, and momentary overloading of the conscious limited capacity system. Errors might occur when two or more output plans are active at the same time. Solving competition between these plans requires allocation of resources. If a wrong output plan is selected, an error will occur. One of the factors affecting the working of the verbal monitor is time pressure. Levelt (1989) states in his perceptual loop theory that the verbal self-monitor is a controlled process, and hence resource-limited. If there are insufficient resources available for verbal monitoring activities, then functioning of the monitor will not be optimal and lead to more errors. Time pressure adds an extra load on the central-capacity system (Baars, 1992). Hence, time pressure causes overload in the functioning of the verbal self-monitoring and prevents sorting out of the competing plans thus leading to more errors. Previous research showed 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, indicating that most likely speech production mechanisms were affected by time pressure and verbal self-monitoring processes were able to adjust to the demands of the task.

In another study, Oomen and Postma (2002) asked participants to overtly describe visually presented networks and detect errors in the speech of others, in the absence or presence of randomly generated finger taps. The finger typing task continuously demands attention, which decreases the processing resources available for monitoring. Oomen and Postma showed that participants intercepted a small percentage of errors in the dual-task condition compared to the percentage of detected errors in the control condition, indicating that self-monitoring might indeed be a resource-limited process.

## 1.3 Electrophysiological correlates of error processing

The recording of electroencephalography (EEG) is a non-invasive technique that provides an excellent way to investigate online information processing in the brain with milliseconds precision (Fabiani et al., 2000; Luck, 2005). EEG reflects information arising from synchronous activation of neuronal populations in the brain. The neural currents resulting from this activation generate an external electric potential that can be measured with EEG. EEG can be used to examine event-related and ongoing brain activity. Event-related potentials (ERPs) are voltage differences that can be measured by EEG before, during or after a sensory, motor or psychological event. Thus, ERPs reflect electrical activity of the brain time-locked to ongoing information processes of a particular event, e.g., presentation of stimuli or a response to a stimulus (Hillyard & Kutas, 1983; Luck, 2005)

Event-related potentials have been an important tool in investigating neural systems involved in different aspects of behavior, including those related to the processing of an error. Over a decade ago an error-related negativity (ERN) was observed in EEG recordings when participants made an error in a task (e.g. Gehring, Goss, Coles, Meyer, & Donchin, 1993; Falkeinstein, Hohnsbein, Hoorman & Blanke, 1991). Since then it has been found that the ERN is generated about 100 ms after the onset of the electromyographic activity preceding the actual overt response and it peaks approximately 100 ms thereafter (see Figure 2; Scheffers, Coles, Bernstein, Gehring, & Donchin, 1996; Holroyd, & Yeung, 2003). The generation of the ERN is localized in the anterior cingulate cortex (ACC; Holroyd, & Coles, 2002). It has been proposed that the ACC has a monitoring function and initiates other systems to adjust their processing as needed (Holroyd & Coles, 2002).

It has been hypothesized that the ERN is generated as a result of a mismatch between representation of anticipated and actual stimuli/response (Bernstein et al., 1995). This hypothesis was supported by the fact that the ERN occurs only after slips and not after mistakes[1] (Dehaene, Posner, & Tucker, 1994). More recently it was suggested that the ERN is an output of an error-detection system. This error-detection system is hypothesized to be comprised of two main components. The first component is system monitoring, function of which is a detection of errors. The monitoring system includes a so-called *comparator* that compares correct response to the actual response. The second component is *remedial action system* (Coles, Scheffers, & Holroyd, 2001). It receives a signal from the comparator if a mismatch is detected, and starts on corrective actions. The ERN is generated when the error signal arrives at the remedial action system (Coles et al., 2001). This theory is in line with the neurophysiological view, which states that when participants err, the mesencephalic dopamine system (i.e. comparator) sends a negative reinforcement signal to the frontal cortex (anterior cingulate sulcus; i.e. remedial action system), where it generates the ERN by disinhibiting the apical dendrites of motor neurons in the ACC (Holroyd, & Coles, 2002). Support for the importance of the mesencephalic dopamine system (i.e. basal ganglia) comes from studies done with Parkinson's patients, who show a decrease in amplitude of the ERN during error trials (i.e. reduction in error detection; Falkenstein, Hoormann, Christ, & Hohnsbein, 2000).

Another theory of the ERN is a conflict-monitoring theory (Botvick, Carter, Braver, Barch, & Cohen, 2001). According to this theory, the ACC monitors for response conflict and then send this information to brain regions that are responsible for cognitive control (e.g. lateral prefrontal cortex).

---

[1] In cognitive psychology there is a difference made between slips and mistakes. Slips are the incorrect executions of correct planned programs. Mistakes are correct execution of the inappropriate programs (Dehaene et al.,1994).

The ERN is generated during continued stimulus processing, after the error causes activation of the correct response, which in turn results in post-error conflict.

The ERN has the following characteristics: the amplitude of the ERN depends on the degree of a mismatch (Bernstein et al., 1995; Holroyd, & Coles, 2002; Falkenstein et al., 2000). The more deviant is the anticipated stimuli/response from the actual stimuli/response the higher the amplitude of the ERN. Furthermore, the amplitude of the ERN largely depends on the motivation of the participants (e.g., the amplitude of the ERN increases when participants concentrate more on the accuracy in the performance than speed; Dehaene et al., 1994). Next, the amplitude of the ERN depends on the confidence that an error had occurred, independently of whether they actually committed an error or not (Scheffers & Coles, 2000). Another characteristic of the ERN is that it can be elicited unconsciously, i.e. when the participants where not aware that they made an error (de Bruijn, Hulstijn, Meulenbroek, & van Galen, 2003; Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001; but see Dehaene et al., 1994, Luu, Flaisch, & Tucker, 2000).

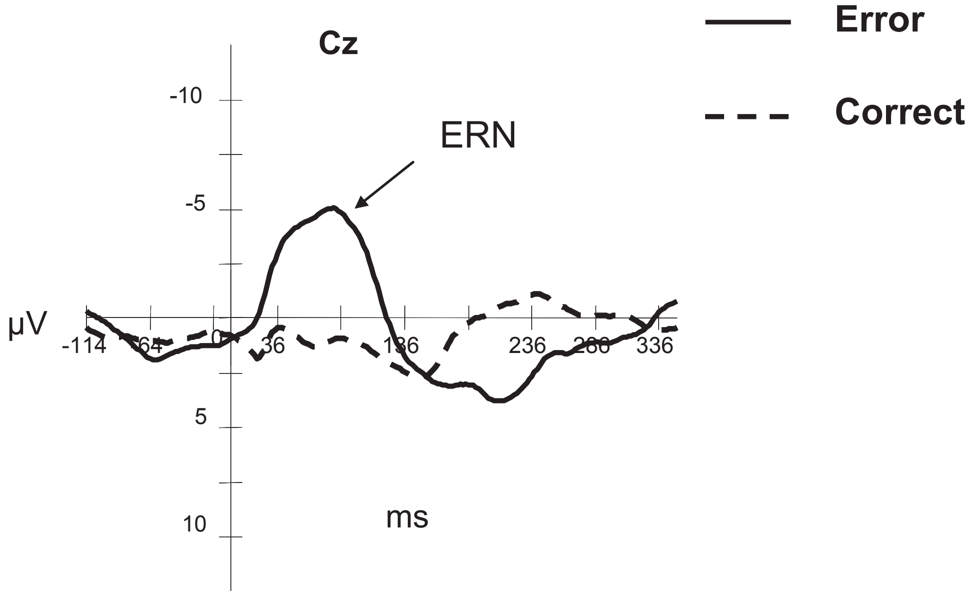

Figure 2. Response-locked ERPs on correct trials (dashed lines) and error trials (solid lines) at Cz. Error trials show a prominent error-related negativity (ERN).

To summarize, the ERN relies on an internal monitoring system (de Bruijn, et al., 2003) and should be indifferent to modality of the error information (Holroyd, & Coles, 2002). Indeed, the ERN was observed in a variety of different areas of research. Note, however, that the majority of the findings reported in this section come from the studies that used action/motor monitoring tasks and not related to the linguistics. There are only few studies that used the ERN after verbal errors (see Masaki et al., 2001; Sebastián-Gallés, Rodríguez-Fornells, Diego-Balaguer, & Díaz, 2006), which we will briefly review below.

Masaki and colleagues (2001) examined whether or not the ERN occurs in relation to speech errors in the Stroop color-word task. Participants in their study were instructed to overtly name the color of each stimulus as quickly and accurately as possible. As a result, Masaki and colleagues found an ERN-like response after speech errors, e.g. when participants named the wrong color.

Sebastián-Gallés and colleagues (2006) assessed Spanish-dominant and Catalan-dominant bilinguals using a lexical decision task in Catalan. In the experimental stimuli, the vowel change involved a Catalan-specific /e – ɛ/ contrast, which is absent in Spanish. The authors showed that Spanish-dominant bilinguals had great difficulty in rejecting experimental non-words, and did not show an ERN in their erroneous non-word decisions either. According to Sebastián-Gallés et al. this suggests that Spanish-dominant bilinguals activated the same lexical entry from experimental words and non-words and therefore showed no differences between correct and erroneous responses. In contrast, Catalan-dominant bilinguals had a clear ERN.

Recently, Möller, Jansma, Rodriguez-Fornells, and Münte (2007) used a laboratory task known to elicit speech errors. In this task, participants are presented with inductor word pairs such as '*ball doze*', '*bash door*', and '*bean deck*', which are followed by a target word pair such as '*darn bore*' (see Motley et al., 1982). The reversal of initial phonemes in the target pair compared to the inductor pairs is supposed to lead to speech errors such as '*barn door*'. Möller and colleagues asked their participants to covertly read the inductor word pairs and vocalize the target word pair immediately preceding a response cue. Möller and colleagues found a negative deflection on error trials, as compared to correct trials, preceding the response cue. They proposed that this activity reflects the simultaneous activation of competing speech plans. However, the authors do not make an explicit link between the negativity they found in their study and the ERN.If the ERN can be observed after speech errors does this mean that the verbal monitor works in a similar way as an action monitor? It seems plausible that different types of monitoring have analogues mechanisms to monitor different kinds of behavioral output. In such a way, an action monitor may monitor for motor slips by checking for possible mismatches between representations of actual and desired motor behavior. A verbal monitor, on the other hand, may monitor some internal representation as it is produced during speech planning by checking potential mismatches between intended and actual verbal production. By investigating the relationship between verbal self-monitoring and the ERN a better understanding of working of verbal monitor may be reached. Additionally, the ERN can provide 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.

As stated in the beginning of this general introduction the verbal self-monitoring is a crucial part of speech production processes. However, essential knowledge concerning the precise working of the monitor is still lacking. Investigations of the factors that effect the working of the verbal monitor and relating verbal monitoring to known electrophysiological correlates of error processing will shed a light on processes underlying verbal self-monitoring in speech production. These are the main aims of the present dissertation. The remaining chapters concern experimental studies of verbal-monitoring. As the chapters are based on separate manuscripts they may be read independently of each other. Therefore, there is unavoidably an overlap between chapters.

# References

Baars, B. (1992). A dozen competing-pans 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.

Baars, B. (1975). Output editing for lexical status in artificially elicited sips of the tongue. *Journal of Verbal Learning and Verbal Behavior,* 14, 382-291.

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.

Blackmer, E.R., & Mitton, J.L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition, 39*, 173-194.

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.

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.

de Bruijn, E.R.A., Hulstijn, W., Meulenbroek, R.G.J., & Galen, van G.P. (2003). Action monitoring in motor control: ERPs following selection and execution errors in a force production task. *Psychophysiology, 40*, 786-795.

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

Fabiani, M., Gratton, G., & Coles, M.G.H. (2000). Event-related brain potentials. Methods, theory, and applications. In J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson (Eds.), *Handbook of Psychophysiology*. (pp. 53-77). New York: Cambridge University Press.

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.

Fromkin, V.A. (1971). The non-anomalous nature of anomalous utterances. *Language, 47*, 27-52.

Garrett, M.F. (1975). The analysis of sentence production. In G.H. Bower, (Ed.) *The psychology of learning and motivation* (pp. 133-177).New York, NY: Academic Press.

Garrett, M.F. (1980). Levels of processing in sentence production. In: B.Butterworth (Ed.) *Language production: Speech and talk* (pp. 177-220). New York, NY: Academic Press.

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.

Hartsuiker, R.J., & Kolk, H.H.J. (2001). Error monitoring in speech production: a computational test of the perceptual loop theory. *Cognitive Psychology, 42*, 113-157.

Hillyard, S., & Kutas, M. (1983). Electrophysiology of cognitive processing. *Annual Review of Psychology, 34*, 33-61.

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.

Laver, J. (1980). Monitoring systems in the neurolinguistic control of speech production. In V., Fromkin, (Ed.) *Errors in linguistic performance: Slips of the tongue, ear, pen and hand* (pp. 287-305). New York: Academic 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.

Levelt, W.J.M. (1999a). Producing spoken language: a blueprint of the speaker. In C. Brown, & P. Hagoort (Eds), *The Neurocognition of Language* (pp. 83-122). Oxford, University Press.

Levelt, W.J.M. (1999b). Models of word production. *Trends in Cognitive Sciences, 3*, 223-232.

Levelt, W.J.M. (2001). Spoken words production: a theory of lexical access. *Proceedings of the National Academy of Science, 6*, 13464-13471.

Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and 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.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

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

Marshall, J., Robson, J., Pring, T., & Chiat, S. (1998). Why does monitoring fail in jargon aphasia? Comprehension, judgment, and therapy evidence. *Brain and Language, 63*, 79-107.

Marshall, R.C., Rappaport, B.Z., & Garcia-Bunuel, L. (1985). Self-monitoring behavior in a case of severe auditory agnosia with aphasia. *Brain and Language*, *24*, 297-313.

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., Rodriguez-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.

Nieuwenhuis, S., Ridderinkhof, K.R., Blom, J., Band, G.P.H., Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology, 38*, 752-760.

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. (2002). Limitations in processing resources and speech monitoring. *Language and Cognitive Processes, 17*, 163-184.

Oomen, C.C.E., Postma, A., & Kolk, H. (2001). Prearticulatory and postarticulatory self-monitoring in Broca's aphasia. *Cortex, 37*, 627-641.

Postma, A. (2000). Detection of errors during speech production: a review of speech monitoring models. *Cognition, 77*, 97-131.

Postma, A., & Kolk, H. (1993). The covert repair hypothesis: prearticulary repair processes in normal and stuttered disfluencies. *Journal of Speech, Language, and Hearing Research,* 36, 472-487.

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.

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.

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.

Scheffers, M.K., & Coles, M.G.H. (2000). Performance monitoring in a confusing world: error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology, 26*, 141-151.

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.

# 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).

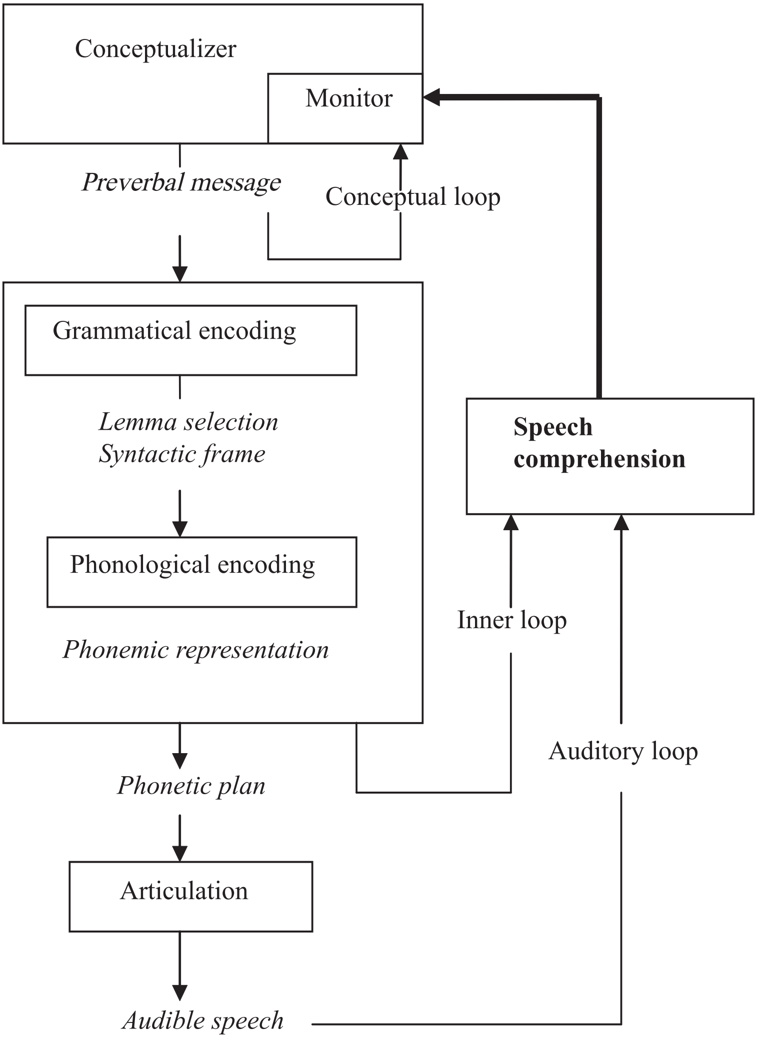

Figure 1. Graphical representation of Levelt's speech production model.

### *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).

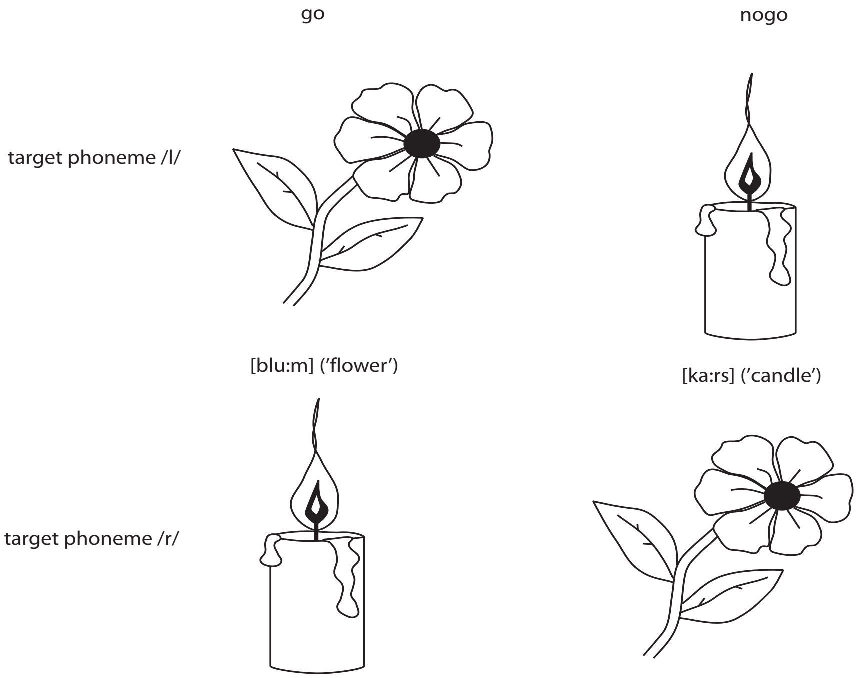

Figure 2. Example of the go and nogo trials for two target phonemes. In the figure, Dutch picture names written in the phonetic code (taken from the CELEX database) and English translations are provided in brackets. Each picture depicted here represents a separate trial. Each picture appeared in the 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).

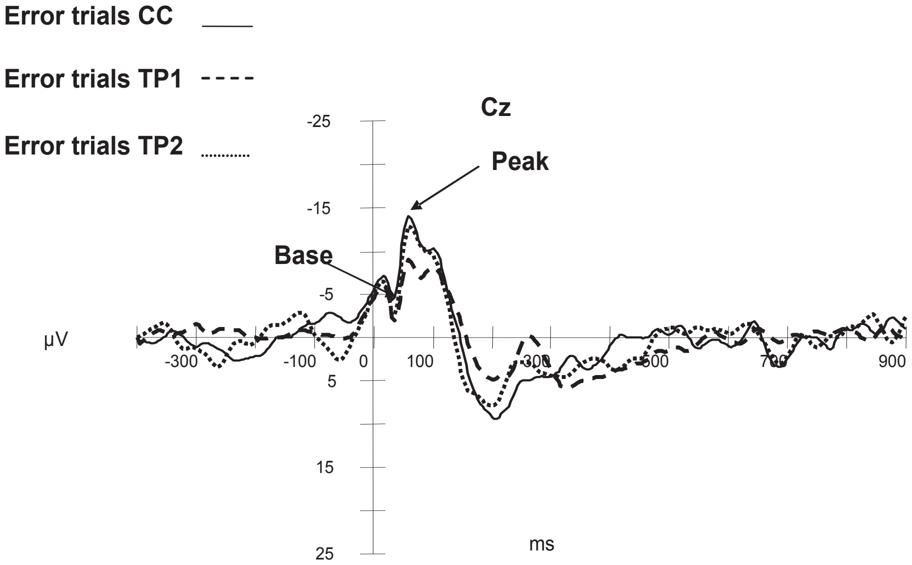

Figure 3. ERP waveforms for a single participant across conditions (CC – control condition, TP1 – time pressure 1 condition, and TP2 – time pressure 2 condition). The amplitude of the ERN is lower at the TP2 condition than during the CC and TP1.

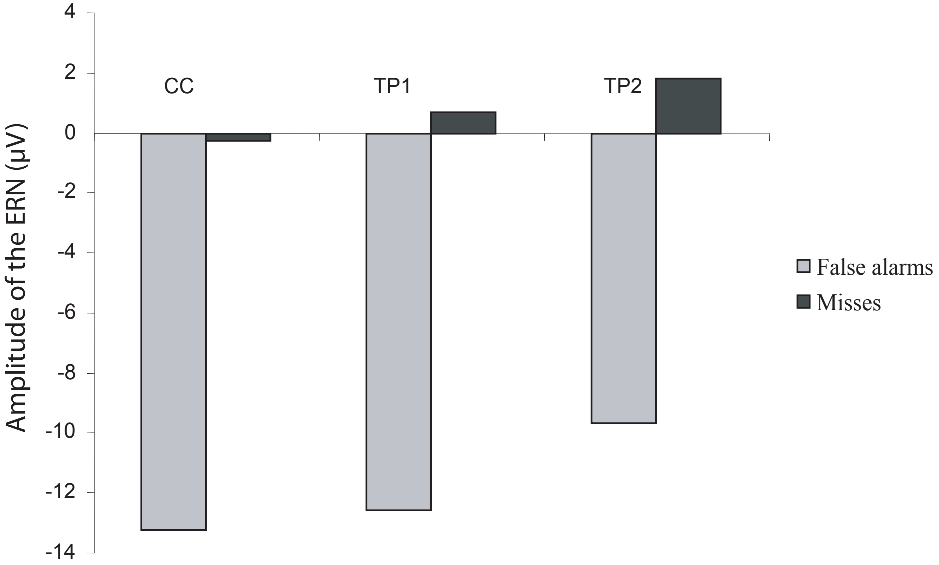

Figure 4. The ERN amplitude across conditions (CC – control condition, TP1 – time pressure 1 condition, and TP2 – time pressure 2 condition) for false alarms and misses.

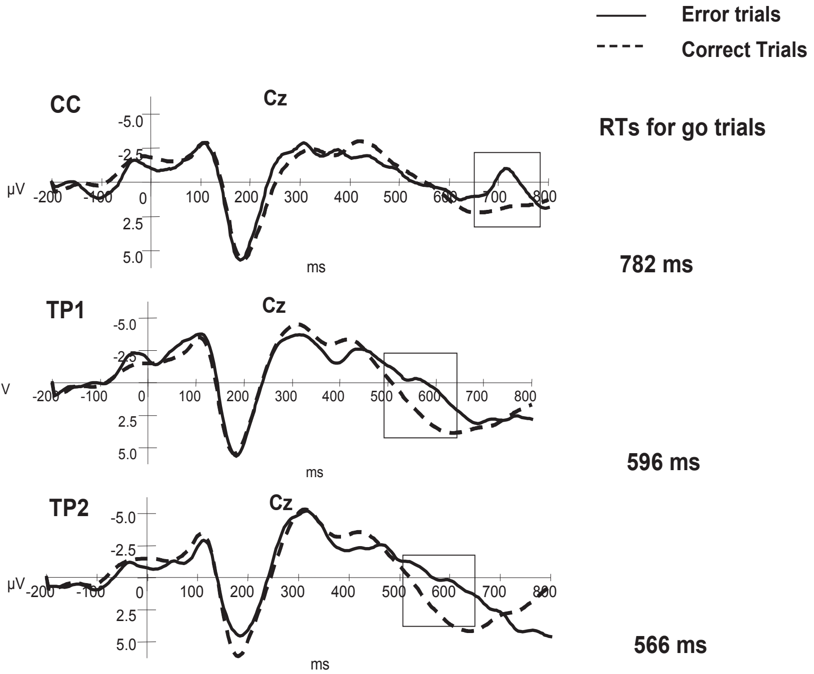

Figure 5. Averaged stimulus-locked ERP waveforms for all misses (solid lines) versus correct rejections (dashed lines) across conditions (CC – control condition, TP1 – time pressure 1 condition, TP2 – time pressure 2 condition). Waveforms depicted for the Cz electrode, where amplitude of the ERN was 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.

# Chapter 3

# Effects of time pressure on verbal self-monitoring in German-Dutch bilinguals[7]

## Abstract

This study addressed how verbal self-monitoring and the Error-Related Negativity (ERN) are affected by time pressure when a task is preformed 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 which showed an atypical increase in amplitude under time pressure. This may 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.

---

[7] This chapter is based on Ganushchak, L. Y. & Schiller, N. O. (submitted). Effects of time pressure on verbal self-monitoring in German-Dutch bilinguals.

# Introduction

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 conscious *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*, which states that the ERN reflects detection of response conflict and not detection of errors *per se* (Botvinick et al., 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 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).

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, in press; Masaki, Tanaka, Takasawa, & Yamazaki, 2001; Möller, Jansma, Rodríguez-Fornells, & Münte, 2007; Sebastián-Gallés, Rodríguez-Fornells, Diego-Balaguer, & Díaz, 2006). In the present study, we investigated the relationship between the ERN and verbal-monitoring in a non-native language. Nowadays, bilingualism is the rule rather than an exception (Costa & Santesteban, 2006), 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 overview see Seal, Aleman, & McGuire, 2004).

The present work is a follow-up of a study by Ganushchak and Schiller (2006). These authors addressed the questions 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/nogo 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, 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 set up of the experiment described in Ganushchak and Schiller (2006). However, participants in the current study were German-Dutch bilinguals and were asked to perform a phoneme monitoring task in their second language, i.e. Dutch. The main question addressed in the current study was: How is the ERN affected by time pressure when a verbal monitoring task is performed in a second language? The monitoring task might be more difficult when performed in a second language, as compared to performing the same task in a native language. This difficulty might arise from co-activation of both native and second languages thereby leading to a higher conflict between potential responses. However, there is increasing evidence that native and foreign languages are based on the same neural substrate (e.g., Klein et al., 1995, 1999; Perani et al., 1998). Moreover, in the existing literature, there is no evidence suggesting that the ERN would be affected differently in a second language. Therefore, it is plausible to assume that verbal monitoring would work in a similar way in bilingual context as compared to monolingual context. 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 females; 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 completed an intensive Dutch language course prior to starting their undergraduate study in the Netherlands. They studied in the Netherlands for at least 2 years (mean: 2.8) and usually lived in the Netherlands. 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 university.

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 non-speeded lexical decision task that was originally developed by Meara (1996). It consisted of 60 items, i.e. 40 low-frequency words and 20 non-words. 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 non-words and Meara's M ($\Delta$ M; see Christoffels et al., 2007). $\Delta$ M 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*

Eighty-one simple line drawings were used as pictures in this experiment (61 for experimental blocks and 20 for a practice block; see 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), 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). Picture labels all started with consonants. The position of the target phoneme was equated across the stimuli.

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 non-words | 67.56 | 9.45 |
| Δ M | 0.29 | 0.12 |

### *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 1,000 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 in order to increase comparability between findings of these two studies. Prior to the experimental blocks, in the CC and TP 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 next trial was taken as response time.

CC and TP condition 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 a computer screen. Experimental blocks consisted of a total of 300 trials (mean 37.5 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 non-target (nogo 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 non-target, participants were asked to monitor for /l/ and /n/.

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 |

During the learning phase, the names of the pictures were presented via head phones. The picture remained in view for 3,000 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 sound-proof 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 had to press a button if a 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. nogo 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 asses 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 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 5 kΩ.

### *Data analysis*

Epochs of 1,300 ms (from –400 ms to +900 ms) were obtained including a 200 ms pre-response baseline. The EEG signal was corrected for vertical electrooculogram (EOG) artifacts, using the ocular reduction method described in Anderer, Satety, Kinsperger, and Semlitsch (1987). For the ERN, averaging was done across false alarms. False alarm trials were compared with correct go trials. The amplitude of the ERN was derived from each individual's response-locked average waveforms after filtering with a band pass, zero phase shift filter (frequency range: 1-12 Hz). 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 from –50 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.

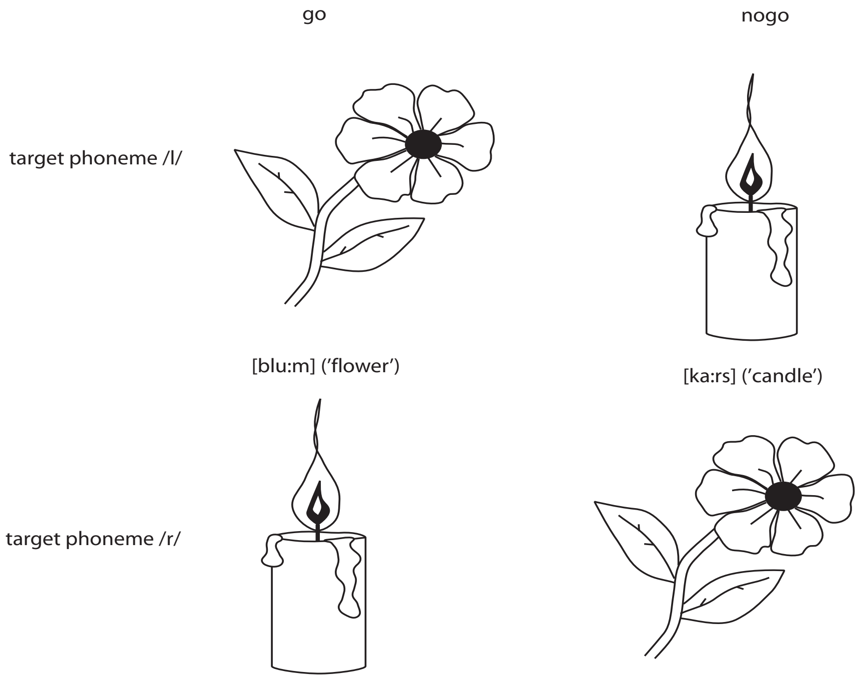

*Figure 1*. Example of go and nogo trials for two target phonemes. In the figure, Dutch picture names are written in phonetic code (taken from the CELEX database) and English translations are provided in brackets. Each picture depicted here represents a separate trial. Each picture appeared in the task as a go and a nogo trial. At the beginning of a block, participants were instructed for which phoneme they had to monitor.

# Results

## *Behavioral data*

RTs shorter than 300 ms and longer than 1,500 ms were excluded from the analysis, which resulted in a loss of 0.7% of all trials. An ANOVA with Time Pressure as independent variable and button-press latencies as dependent variable revealed a significant effect of Time Pressure ($F(1, 38) = 348.63$, $MS_e = 1008.12$, $p < .001$). Participants were significantly faster during the TP condition (671 ms, SD = 21) than the CC (865 ms, SD = 23). A similar analysis with number of errors as dependent variable also demonstrated a significant effect of Time Pressure ($F(1, 38) = 22.80$, $MS_e = 58.46$, $p < .001$). 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.

The picture naming task was used to asses whether 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 independent variable. Number of errors during the picture naming task significantly decreased in TP condition when compared to the CC ($F(1, 38) = 84.42$, $MS_e = 5.09$, $p < .001$). 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. There was no negative deflection observed in the ERP waveforms for correct trials during visual inspection of the EEG waves. Figure 2A 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). The ERN obtained in the present study showed a frontal distribution (see Figure 2B for a topographical representation of the ERN across CC and TP conditions).

An ANOVA with Time Pressure as independent variable and amplitude of the ERN as dependent variable revealed a significant effect of Time Pressure ($F(1, 38) = 4.68$, $MS_e = 46.19$, $p < .05$). Interestingly, German-Dutch bilinguals showed enhanced amplitude of the ERN in the TP condition compared to the CC. To investigate whether or not Time Pressure elicited higher ERP amplitudes in general rather than specifically on the ERN, we ran a 2 (correct vs. error) by 2 (CC vs. TP) ANOVA. More specifically, we used a mean area analysis for investigating effects of Time Pressure on correct and erroneous trials, as it was impossible to identify peaks for correct trials on a trial-by-trial basis. The time window of interest was 0 – 100 ms after the overt response. The analysis showed a significant interaction between Correctness of Response and Time Pressure ($F(1, 20) = 5.58$, $MS_e = 9.20$, $p < .05$). Time Pressure had an effect only for erroneous trials and not for correct trials ($F(1, 20) = 6.96$, $MS_e = 7.87$, $p < .05$ and $F < 1$, respectively). 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, while 27% of the participants (5 out of 21) showed lower amplitudes of the ERN under time pressure compared to the control condition.

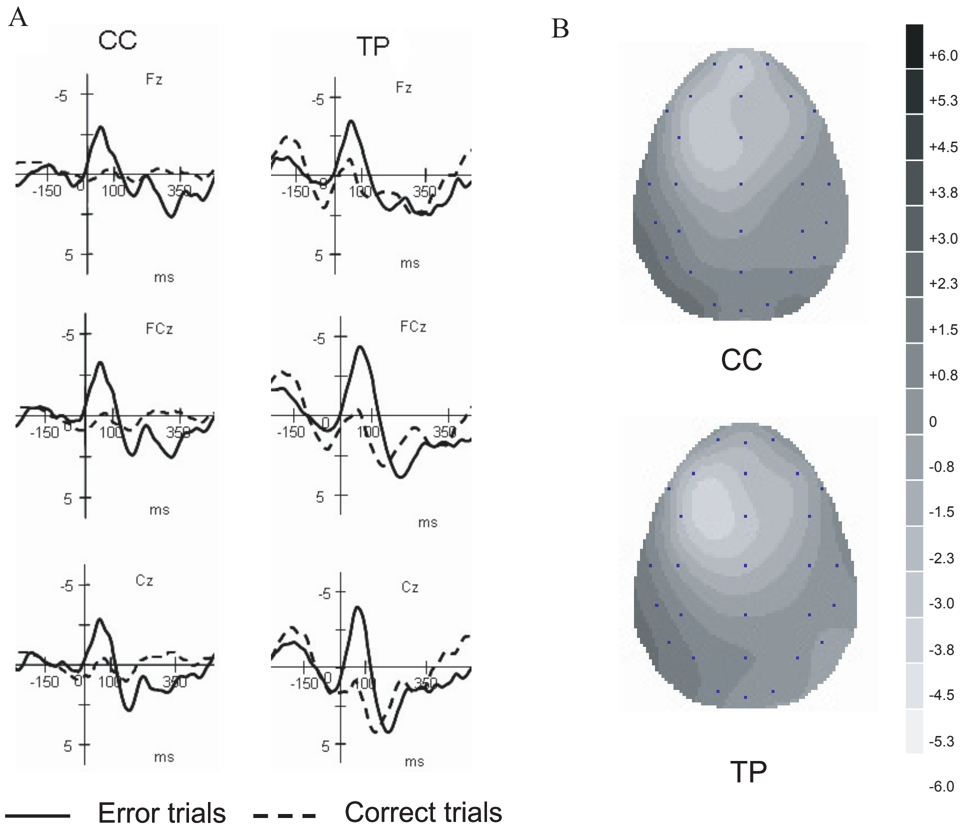

*Figure 2*. A: Averaged ERP waveforms for all incorrect versus correct trials across conditions and electrodes (CC – control condition, TP – time pressure condition). Correct and incorrect trials were matched on RTs and number of trials.
B: Topographic maps of the ERN amplitude between 0 and 100 ms after response onset. Negative regions depicted in light gray.

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 used a mean area analysis in the time window of 140 – 270 ms. A 2 (correct vs. error) by 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 ($F < 1$).

# 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 participants made more errors under time pressure. This is in accordance with previous 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 on-line 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.

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 language (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 on-line 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, i.e. Dutch. In order 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 when bilingual speakers became aphasic (e.g., Fabbro, Skrap, & Aglioti, 2000), undergo brain stimulation (e.g., Holtzheimer, Fawaz, Wilson, & Avery, 2005), or under psychological stress (e.g., Dornik, 1979, 1980; Grosjean, 1982). According to Levelt (1989), monitoring involves controlled processing that requires attentional control. In a second language, a considerable 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, i.e. 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, non-target language words (Rodríguez-Fornells et al., 2005). It is possible that under time pressure, inhibition of the non-target 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 et al., 2000; Colomé, 2001; Rodríguez-Fornells et al., 2005; but see Hermans et al., 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). 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). The possible generator of the ERN is the anterior cingulated cortex (ACC; Dehaene et al., 1994). In a previous literature, the ACC and the inferior frontal cortex are implicated in error processing (e.g., Menon et al., 2001). The ACC is thought to monitor competition between processes that conflict during task performance (Carter et al., 1998).

Suggestively, the increased amplitude of the ERN under time pressure in bilingual context 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, i.e.

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 performed the task more accurately and made fewer errors in the TP condition that the CC. The findings of the current study demonstrate the opposite, i.e. participants made more errors under time pressure than in the absence of time pressure. Thus, we believe that in the present study practice did not have a large influence on performance.

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, i.e., 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 since based on the previous literature, action slips did not lead to an enhancement of the ERN under time pressure (Falkenstein et al., 1991; Gehring et al., 1993). The reversal effect of time pressure on the ERN in multilingual context merits further research, 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). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

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.

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.

Costa, A., & Santesteban, M. (2006). The control of speech production by bilingual speakers: introductory remarks. *Bilingualism: Language and Cognition*, 9, 115-117.

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.

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.

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, 1-14.

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: 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.

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 of the 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, U.K.: 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.

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.

Menon, V., Adleman, N. E., White, C. D., & Reiss, A. L. (2001). Error-related brain activation during a go/nogo response inhibition task. *Human Brain Mapping*, 12, 131-143.

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, U.K.: 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., Cappa, S. F., Fazio, F., & Mehler, J. (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: 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.

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.

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.

# Appendix

The list of stimuli used in the experimental blocks. The approximate English translation is given in brackets. 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 (burrel), 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 (villege), 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)

# Chapter 4

# Brain error-monitoring activity is affected by semantic relatedness: An event-related brain potentials study[8]

## Abstract

Speakers continuously monitor what they say. Sometimes, self-monitoring malfunctions and errors pass undetected and uncorrected. In the field of action monitoring, an event-related brain potential, the Error-Related Negativity (ERN), is associated with error processing. The present study relates the ERN to *verbal* self-monitoring and investigates how the ERN is affected by auditory distractors during verbal monitoring. We found that the ERN was largest following errors that occurred after semantically related distractors had been presented, as compared to semantically unrelated ones. This result demonstrates that the ERN is not only sensitive to response conflict resulting from the incompatibility of motor responses but also to more abstract lexical retrieval conflict resulting from activation of multiple lexical entries. This in turn suggests that the functioning of the verbal self-monitoring system during speaking is comparable to other performance monitoring, such as action monitoring.

---

[8] This chapter is based on Ganushchak, L. Y., & Schiller, N. O. (2008). Brain error-monitoring activity is affected by semantic relatedness: an event-related brain potential study. *Journal of Cognitive Neuroscience*, 20, 1-14.

# Introduction

The neural basis of error monitoring has become a key issue in cognitive neuroscience. An electrophysiological index thought to be associated with error processing is the *Error-Related Negativity* (ERN; Falkenstein, Hohnsbein, Hoorman, & Blanke, 1991; Gehring, Goss, Coles, Meyer, & Donchin, 1993). The ERN is an 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). Originally, the ERN was thought to arise as a result of conscious 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 this view, the ERN reflects detection of response conflict and not detection of errors *per se* (Botvinick et al., 2001; Carter et al., 1998). 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. In the present study, we focused on a different type of monitoring, namely *verbal self-monitoring*. Verbal self-monitoring is a crucial part of language 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). Furthermore, malfunction of 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).

One prominent theory about verbal self-monitoring is the *perceptual-loop theory* proposed by Levelt (1983, 1989). According to this theory, a speech monitor checks the 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 self-monitoring is achieved via the speech comprehension system, i.e. the same system that is used for understanding the speech of others. The precise working of the monitor is not described in detail by theories of language production (e.g. Levelt, Roelofs, & Meyer, 1999 for an overview). However, the system must consist of at least two components, one of which attends to the output of the speech programming process, and another one that compares this output with some standard. If the output does not satisfy a particular criterion, the monitor initiates a self-correction process (Levelt et al., 1999; Postma, 2000).

A similar mechanism is implemented in a model of action monitoring (Desmurget & Grafton, 2000). According to this 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 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 central question we try to answer in this study is: Does *verbal monitoring* work 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. An action monitor, for example, may monitor for motor slips by checking for possible mismatches between representations of actual and desired motor behavior. A verbal monitor, on the other hand, may monitor some internal representation as it is generated during speech planning by checking potential mismatches between intended and actual verbal production.

If the ERN is associated with error processing in action monitoring, can it also be applied to error processing in verbal monitoring? Previous studies showed that an ERN can be elicited by verbal errors (e.g., Ganushchak & Schiller, 2006; Masaki, Tanaka, Takasawa, & Yamazaki, 2001; Sebastián-Gallés, Rodríguez-Fornells, Diego-Balaguer, & Díaz, 2006). For instance, Masaki and colleagues (2001) examined whether or not 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. 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.

Ganushchak and Schiller (2006) used a phoneme monitoring task to address the question whether or not an ERN occurs after verbal error detection and whether a potential ERN is affected by a time pressure manipulation. These authors obtained an ERN following verbal errors. Furthermore, their ERN showed a decrease in amplitude under conditions of severe time pressure.

Recently, Möller, Jansma, Rodríguez-Fornells, and Münte (2007) used a laboratory task known to elicit speech errors. In this task, participants are presented with inductor word pairs such as *ball doze*, *bash door*, and *bean deck*, which are followed by a target word pair such as *darn bore* (Motley et al., 1982). The reversal of initial phonemes in the target pair compared to the inductor pairs is supposed to lead to speech errors, e.g. *barn door*. Möller and colleagues asked their participants to covertly read the inductor word pairs and vocalize the target word pair immediately preceding a response cue. Möller and colleagues found a negative deflection on error trials, as compared to correct trials, preceding the response cue. They proposed that this activity reflects the simultaneous activation of competing speech plans. However, the authors do not make an explicit link between the negativity they found in their study and the ERN.

### *Current study*

The task employed in our present study was a phoneme monitoring go/nogo task, previously used in language production and verbal monitoring research (e.g., Morgan & Wheeldon, 2003; Schiller, 2005; Wheeldon & Levelt, 1995). In the current study, participants were required to internally name pictures and press a button if a particular target phoneme occurred in the name of the picture. For instance, if the target phoneme was /b/ and the target picture was *bear*, participants were required to press a response button. Thus, participants were asked to monitor their own internal speech production. Target pictures were presented with or without auditory distractors (SOA = 0 ms), which were semantically related or unrelated to the target pictures.

Previous studies have shown that participants are slower in naming pictures when distractors are semantically related to target pictures than when they are unrelated (e.g., Damian & Martin, 1999; Lupker, 1979; Schriefers, Meyer, & Levelt, 1990; Starreveld & La Heij, 1995, 1996). Current models of word production have proposed various accounts for this semantic interference effect (e.g., Levelt et al., 1999; Schriefers et al., 1990), but they share the core assumption that semantically related concepts, such as *train* and *bus*, co-activate each other through activation spreading in a semantic network (Collins & Loftus, 1975; Quillian, 1968). Co-activated concepts activate their corresponding lexical entries which compete for selection, and this competition affects selection latencies (see Levelt et al., 1999 for a review; but see Finkbeiner, Gollan, & Caramazza, 2006 as well as Mahon, Costa, Peterson, Vargas, & Caramazza, 2007 for an alternative view). The concept with the highest activation is selected for further processing. In light of these findings, we expected that in the present study semantically related distractors, as opposed to semantically unrelated distractors, would cause more lexical retrieval conflict thereby leading to slower and more error-prone functioning of the monitor (Damian & Martin, 1999; Lupker, 1979; Schriefers et al., 1990).

Moreover, the perceptual loop theory (Levelt, 1983, 1989) assumes that verbal self-monitoring occurs through the speech comprehension system. Thus, we expected that auditory distractors will cause longer button press latencies and higher error rates by impeding the functioning of the verbal self-monitor, compared to performance in the absence of the distractors.

Beside behavioral effects, we explored two ERP components in relation to verbal monitoring: the ERN and the N450. The N450 reflects negativity over fronto-central regions that peaks about 450 ms after stimulus onset on correct trials (Liotti, Woldorff, Perez, & Mayberg, 2000; West, 2003). The N450 is associated with lexical conflict in studies employing a Stroop task and is elicited by both response and non-response conflict (i.e. conflict that arises prior to response processing, e.g., at phonetic and/or semantic levels; Liotti et al., 2000; West, Bowry, & McConville, 2004). The amplitude of the N450 is larger when color and word information are incongruent than when they are congruent (West, 2003). In the present study, we expected to find an increase in the amplitude of the N450 in the presence of semantically related distractors, since they will cause higher lexical conflict than semantically unrelated distractors. The N450 amplitude should be smallest in the control condition (i.e. in the absence of distractors), since in the control condition there should be little conflict present. Note that West (2003) showed higher amplitudes of the N450 on incongruent trials, whereas we expected to find higher amplitudes of the N450 on congruent (i.e. semantically related) trials. This seeming contradiction is resolved when one considers the amount of conflict at the aforementioned trials. In a Stroop task, incongruent trials (e.g., the word RED presented in blue) lead to more conflict than congruent trials (e.g., RED presented in red). In our study, in contrast, congruent trials (e.g., a picture of a *nose* presented with word EAR) yield more conflict than incongruent trials (e.g., picture of a *nose* presented with word WINDOW).

Additionally, in response-locked EEG averages, the ERN was of special interest to us. Throughout the action monitoring literature, it has consistently been reported that the amplitude of the ERN increases when response conflict was increased (e.g., Yeung et al., 2004). However, there are no studies showing that the ERN can be affected by lexical retrieval conflict. As mentioned above, the presence of the semantically related distractors as compared to semantically unrelated ones may activate more entries that compete for response selection. For example, seeing a picture of a *nose* and hearing the distractor word *ear* will also activate other entries from the same semantic category, such as *lip*, *tongue*, *eye*, etc., thereby increasing the number of possibly correct, i.e. more or less appropriate, responses. This, in turn, may lead to increased response conflict and higher amplitudes of the ERN. Sensitivity of the ERN to the increase of lexical conflict might provide extra evidence for the hypothesis that verbal monitoring involves similar processes as non-verbal action monitoring. We expected to find no difference in the latency of the ERN, since latency seems to be invariant to the erroneous responses and largely independent from experimental manipulation (Falkenstein et al., 2000; Scheffers & Coles, 2000).

To summarize, we hypothesized that auditory distractors, by impeding the functioning of the verbal monitor and by increasing lexical conflict, lead to slower and more erroneous responses, as well as larger amplitudes of the ERN and the N450.

# Methods

## *Participants*

Twenty-two students of Maastricht University (19 women) took part in the experiment. All participants were right-handed, native speakers of Dutch, and had normal or corrected-to-normal vision. Participants gave written informed consent prior to participating in the study. They received a small financial reward for their participation in the experiment. Two participants were excluded from the analysis because they made no errors, which made it impossible to compute an ERN.

## *Materials*

Forty simple line drawings were used as target pictures in this experiment. The labels of all the pictures were monosyllabic Dutch words (e.g., *heks* 'witch', *brood* 'bread', etc.).

Table 1. Lexico-statistical characteristics of the target words.

| Target phoneme | Example | 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 |

All picture names had a moderate frequency of occurrence between 10 and 100 per million according to the CELEX database (see Table 1; CEnter for LEXical information, Nijmegen; Baayen, Piepenbrock, & Gulikers, 1995). Labels of the pictures started with consonants. The position of the target phoneme was equated across the stimuli.

For each picture, a word from the same semantic category was selected that served as semantically related distractor (e.g., for the target picture *neus* 'nose' the semantically related distractor was OOR 'ear'). An unrelated word without any obvious relationship to the target picture served as semantically unrelated distractor (e.g., the target picture *neus* 'nose' was paired with the word RAAM 'window'). Semantically related and unrelated distractors were selected such that they were phonologically unrelated to the picture labels. Target phonemes, for which participants were required to monitor during the task, were not contained in the distractors. Due to the difficulty of finding distractors which did not contain target phonemes, pictures for go and nogo trials were paired with different distractors. Distractors that were semantically related during go trials were used as semantically unrelated ones during the nogo trials and vice versa. For instance, on go trials the target picture *neus* ‘nose’ was paired with the semantically related distractor OOR ‘ear’; on nogo trials, this distractor was paired as semantically unrelated distractor with the picture *fles* ‘bottle’ (see the Appendix for the list of pictures and distractors). Distractor words for a particular target were matched for frequency of occurrence, number of syllables, and number of letters.

### *Design*

The experiment consisted of three experimental conditions: a control condition (CC), an auditory interference condition with semantically related distractors (SR+), and an auditory interference condition with semantically unrelated distractors (SR–). During the CC trials, participants only saw pictures, whereas during auditory interference conditions (SR+ and SR–), pictures were presented simultaneously with auditory distractors (SOA = 0 ms). Some of the SR+ and SR– trials (8.5% of all trials) were catch trials during which participants were asked to overtly name the last distractor word they heard.

There were twenty experimental blocks of on average 24 trials and one practice block of 19 trials. 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 practice trials. In all blocks, pictures were presented one by one on a computer screen. In each condition, each picture was repeated four times: twice as a non-target and twice as a target. Each time, participants were asked to monitor for a different phoneme. For instance, participants were asked to monitor once for the phoneme /n/ and once for the phoneme /s/ when *neus* ‘nose’ was a target. When *neus* was a non-target, participants had to monitor for /k/ and /r/ (see Figure 1 for an illustration of the task). A fixation point always preceded the pictures. The duration of the fixation point varied between 500 ms and 800 ms. In all conditions and after each trial, participants rated the accuracy of their response on a three-point Likert scale. The Likert scale was presented in the middle of the screen after a fixed time interval of 1,000 ms following disappearance of the visual stimulus or after a response to the target picture has been made. This scale included the following options: *surely correct*, *do not know*, and *surely incorrect*. Order of pictures was pseudo-randomized such that the same picture never appeared twice on subsequent trials. The block order was randomized for each participant.

During the learning phase, the names of the pictures were presented auditorily, in order to avoid priming for letters. The picture remained in view for 3,000 ms or until the response button was pressed. In the picture naming task, the pictures disappeared from the screen as soon as the voice key was activated or after the response deadline was reached, which was 550 ms (this response deadline is based on the outcome of a previous study; see Ganushchak & Schiller, 2006 for details).

### Procedure

Participants were tested individually while seated in a sound-proof room. They were presented with a familiarization phase, a picture naming task, a practice block, and the experimental blocks. During the familiarization 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 were familiarized with before. If errors were made, participants were told about their mistakes and correct responses were provided. Prior to practice and experimental blocks, participants received an auditory sample of the phoneme they were required to detect (e.g., *Reageer nu op de klank /l/ zo als in tafel, spelen, verhaal* 'React now to the sound /l/ as in table, play, tale'). Participants were asked to press a button if a target phoneme was present in the picture name (i.e. go trials). When a target phoneme was not present in the name of the picture, participants were required to refrain from button pressing (i.e. nogo trials). Participants were instructed to give all responses to go trials with their right hand. Responses to the Likert scale were given with the left hand. During the catch trials, participants were asked to say out loud the last distractor word they heard. 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 electro 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 mastoid. Eye movements were recorded to allow off-line rejection of contaminated trials. Lateral eye movements were measured using a bipolar montage of electrodes placed on the right and left external canthus. Eye blinks and vertical eye movements were measured using bipolar montage of 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 obtained including a 200 ms pre-response baseline. The EEG signal was corrected for vertical electrooculogram (EOG) artifacts, using the ocular reduction method described in Anderer, Satety, Kinsperger, and Semlitsch (1987). For the ERN, averaging was done across false alarm and miss trials. False alarm trials were compared with correct go trials and misses were compared with correct nogo trials. For false alarms, the amplitude and latency of the ERN was derived from each individual's response-locked average waveforms after filtering with a band pass, zero phase shift filter (frequency range was 1-12 Hz).

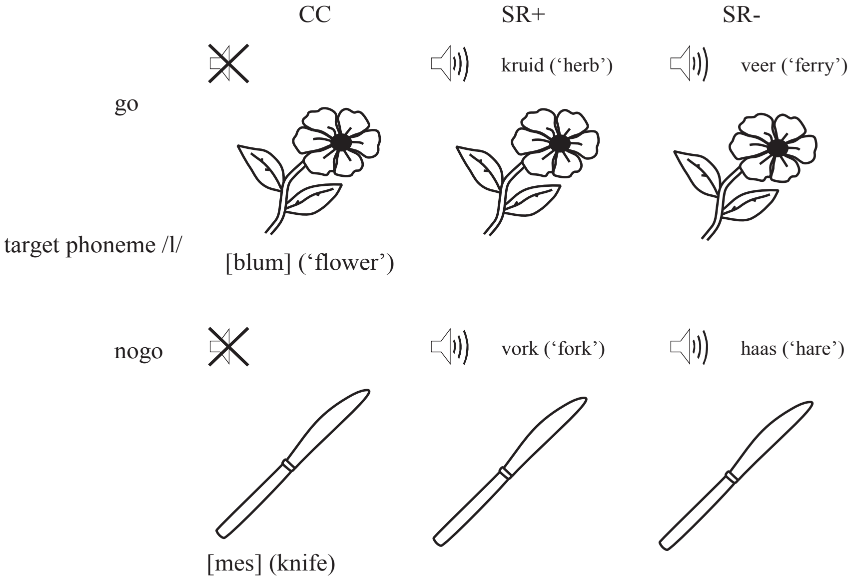

*Figure 1*. Example of the go and nogo trials for three auditory conditions, i.e. control condition (CC), semantically related distractors (SR+) and semantically unrelated distractors (SR–). In the figure, Dutch picture names written in phonetic transcription (taken from the CELEX database) and English translations are provided in brackets. Each picture depicted here represents separate trials. At the beginning of each block, participants heard for which phoneme they were asked to monitor.

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). More specifically, the amplitude of the ERN was defined as the difference between the most negative peak in a window from 0 to 150 ms after the response and the most positive peak from –50 to 0 ms preceding the overt response (Falkenstein et al., 2000). The latency of the ERN was defined as a point in time when the negative peak was at its maximum. Misses were computed from stimulus-locked averages and analyzed with a mean area amplitude analysis, as it was impossible to identify the peaks on a trial-by-trial basis. Time windows of interest were derived based on the visual inspection of the grand average waveforms with consideration of average button-press latencies in go trials. The selected time window of interest was 760 – 860 ms. The amplitude and latency of the ERN were recorded for each condition and type of error at electrode sites Fz, FCz, Cz, and Pz.

For the stimulus-locked analysis of correct trials a band pass, zero phase shift filter (frequency range: 1-30 Hz) was used. Stimulus-locked averaging was done for correct go responses. The mean amplitude analysis was performed for the time window from 350 ms to 550 ms after stimulus onset. This time window was derived based on the visual inspection of the grand average waveforms and previous studies employing the N450 (e.g., West, 2003).

# Results

## Behavioral data

Button-press latencies (RTs) shorter than 300 ms and longer than 1,500 ms were excluded from the analysis. During catch trials, participants almost perfectly named the last word they heard (error rate: 0.4%), indicating that participants processed the auditory input even though they were instructed to ignore distractors. The effect of the distractors on RTs was assessed by an ANOVA revealing a significant main effect of the Distractor Relatedness ($F(2, 34) = 33.62$, $MS_e = 291.93$, $p < .001$). Participants were significantly slower during the SR+ (743 ms, SD = 92; $F(1, 17) = 48.99$, $MS_e = 761.93$, $p < .001$) and SR– (729 ms, SD = 94; $F(1, 18) = 25.17$, $MS_e = 720.17$, $p < .001$) conditions than during the CC (699 ms, SD = 84). The difference between the conditions SR+ and SR– was also significant ($F(1, 17) = 12.73$, $MS_e = 269.46$, $p < .01$), i.e. there was a semantic interference effect.

Similar analyses were repeated for the number of errors as dependent variable. Participants made on average 4.0% alse alarms and 2.9% misses. An ANOVA revealed no effect of Distractor Relatedness ($F(2, 34) < 1$).

Participants' awareness about the correctness of their response was assessed by a 3 (Distractor Relatedness) by 3 (Certainty Level) ANOVA. The analysis showed a significant effect of Certainty Level ($F(2, 34) = 21.86$, $MS_e = 24.87$, $p < .001$). On 90% of all error trials participants knew that they made an error, as compared to 7% of all error trials on which participants were not aware of their errors, and 3% when they were not sure whether or not they made an error (for 'sure correct' vs. 'sure incorrect', $F(1, 17) = 20.81$, $MS_e = 24.78$, $p < .001$; for 'do not know' vs. 'sure incorrect', $F(1, 17) = 23.27$, $MS_e = 24.51$, $p < .001$; and for 'sure correct' vs. 'do not know', $F(1, 17) = 3.01$, $MS_e = 0.46$, n.s.). Neither the effect of Distractor Relatedness, nor the interaction between Distractor Relatedness and Certainty Level were significant (both F's < 1, respectively).

## Electrophysiological data

### *ERN*

The ERN was revealed in response-locked ERP averages for incorrect nogo trials (i.e.

false alarms). No negative deflection was observed in the ERP waveforms for correct trials during visual inspection of the EEG waves. Unfortunately, due to the very low amount of incorrect trials during which participants were unaware of their errors (on average, there were 0.2 errors during which participants were not aware of the incorrectness of their responses), it was not possible to analyze the subjective reliability data in relation to the ERN.

Figure 2 provides an overview of the response-locked averaged ERP waveforms for false alarms (incorrect nogo trials) and correct go trials across distractor conditions (CC, SR+, and SR–). Figure 3 displays the topographical representation of the ERN.

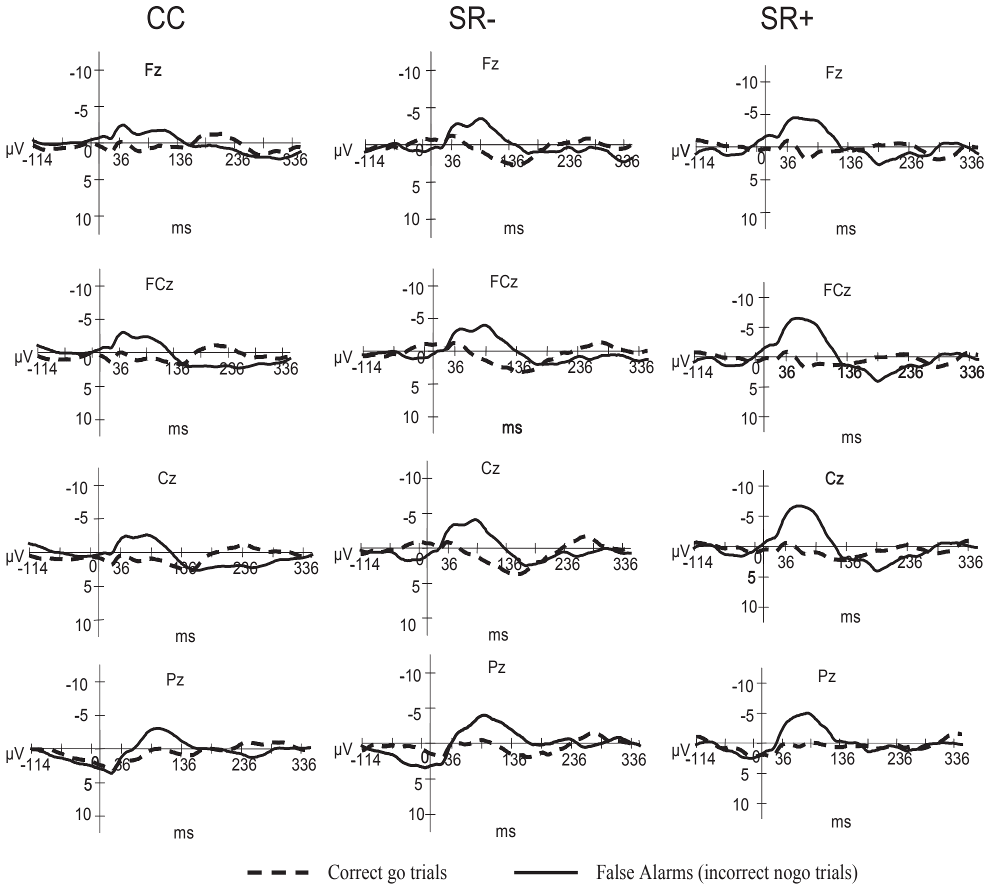

*Figure 2.* Averaged response-locked ERP waveforms for all false alarms (incorrect nogo trials; solid lines) versus correct trials (dashed lines) across conditions (control condition [CC], semantically related [SR+], and semantically unrelated [SR–] conditions). Correct and incorrect trials were matched on RTs and number of trials.

To investigate the effects of the auditory interference on the ERN, an ANOVA was run with Distractor Relatedness as independent variable and the amplitude of the ERN as dependent variable. This analysis revealed a significant effect of Distractor Relatedness (F(2, 34) = 5.36, $MS_e$ = 63.44, p < .05). A mean area analysis in the time window of interest (0 – 135 ms) was in accordance with our peak-to-peak analysis and revealed a significant effect of Distractor Relatedness (F(2, 34) = 9.47; $MS_e$ = 17.21, p < .01). The amplitude of the ERN was significantly larger in the SR+ condition (–5.73 µV, SD = 2.85) than in the CC (–3.06 µV, SD = 2.69; F(1, 17) = 16.74, $MS_e$ = 27.74, p < .01) and in the SR– condition (–3.19 µV, SD = 3.13; F(1, 17) = 13.56, $MS_e$ = 37.78, p < .01), again reflecting a semantic interference effect.

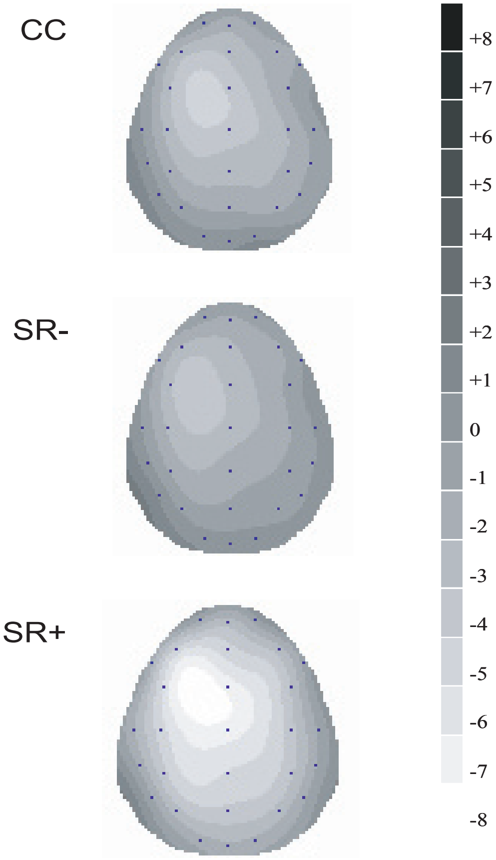

*Figure 3*. Topographic maps of the ERN amplitude between 0 and 120 ms after response onset. Negative regions depicted in light gray.

The difference in amplitude of the ERN between the SR– condition and the CC did not reach significance (F(1, 17) = 1.53, $MS_e$ = 37.75, n.s.). As expected, the analysis with the latency of the ERN as dependent variable showed no significant effect of Distractor Relatedness (F(2, 34) < 1). Figure 4 displays average false alarms waveforms of all experimental conditions superimposed across four midline electrode sites.

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).

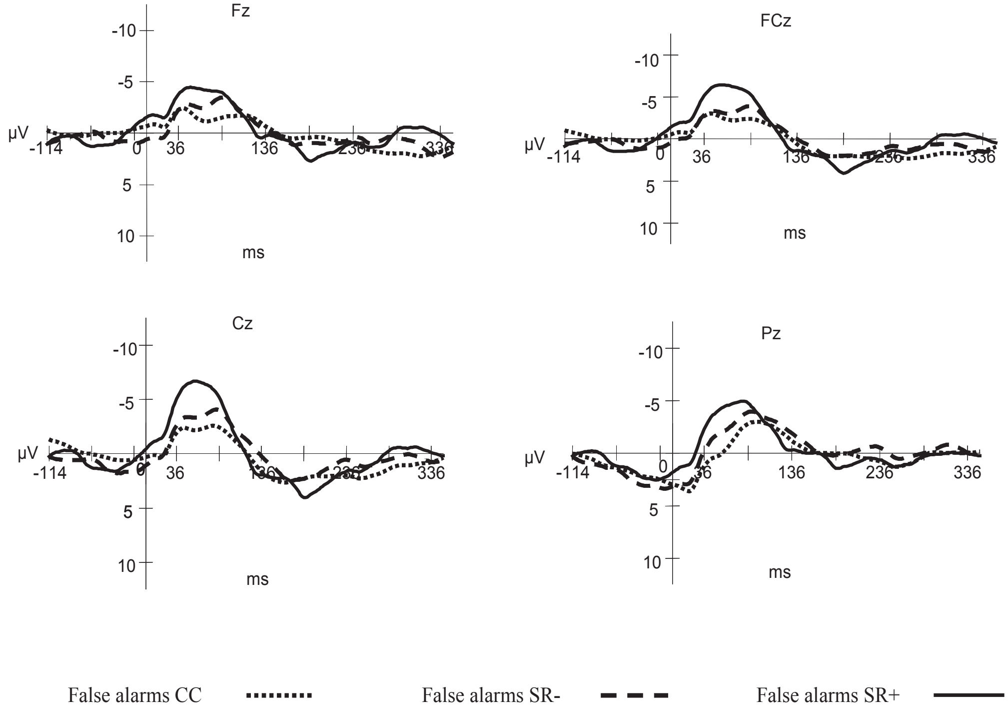

*Figure 4*. Averaged response-locked ERPs for false alarms (incorrect nogo trials). Dotted lines depict the control condition (CC), solid lines depict the semantically related condition (SR+), and dashed lines depict the semantically unrelated condition (SR–).

A mean area analysis was used to test whether there was a significant difference between misses and correct rejections. An ANOVA was run with Correctness of Response as independent variable and the amplitude of the ERN-like response as dependent variable. The analysis showed that erroneous responses had more negative amplitudes than correct responses (F(1, 17) = 10.13, $MS_e$ = 14.14, p < .01). Unfortunately, we did not have sufficient trials to compare erroneous trials across conditions.

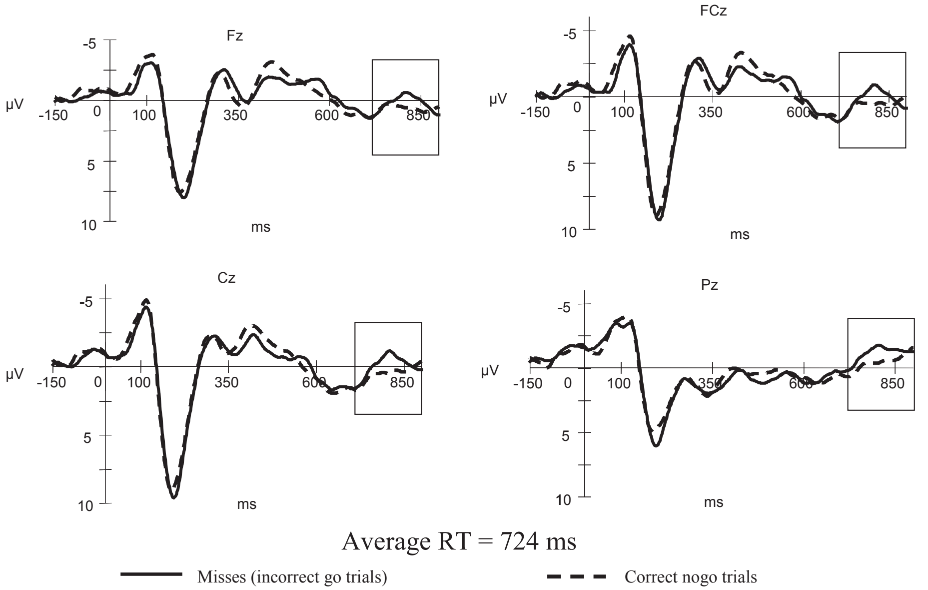

*Figure 5*. Averaged stimulus-locked ERP waveforms for all misses (solid lines) versus correct rejections (dashed lines) collapsed across conditions. Correct and incorrect trials were matched on RTs and number of trials. Areas selected by rectangles depict the time window of average button-press latencies for correct responses. Mean RTs for overt responses in go trials are provided.

## Stimulus-locked averages

### *Early components*

In stimulus-locked averages, there is an apparent visual difference between distractor conditions in the N1-P2 complex. An ANOVA revealed a significant main effect of Distractor Relatedness (F(2, 34) = 9.16, $MS_e$ = 14.53, p < .001; see Figure 6). More detailed analyses showed that in the CC the amplitude of N1 was lower than in the SR– and SR+ conditions (F(1, 17) = 9.88, $MS_e$ = 31.22, p < .01 and F(1, 17) = 11.74, $MS_e$ = 40.23, p < .01, respectively). Importantly, however, the difference between the SR– and the SR+ conditions was not significant (F(1, 17) = 1.16, $MS_e$ = 15.73, n.s.).

Similar analyses with the latency of the N1 as dependent variable showed that the N1 peaked earlier during the CC than during the SR– and SR+ conditions (F(1, 17) = 18.22, $MS_e$ = 867.61, p < .001 and F(1, 17) = 22.19, $MS_e$ = 899.02, p < .001, respectively). Again, there was no significant difference between the SR– and SR+ conditions (F(1, 17) < 1). These early differences between the CC and the SR+ and SR− conditions are not surprising given that auditory distractors were presented in the SR– and SR+ conditions, whereas no such distractor was present in the CC. Importantly, however, no differences were obtained between the SR+ and the SR− conditions at these early moments in the signal.

### *N450*

In stimulus-locked averages in the time window of interest (350 – 550 ms) there was a significant main effect of Distractor Relatedness ($F(2, 34) = 24.59$, $MS_e = 3.83$, $p < .001$; see Figure 6). A higher amplitude of the N450 was visible in the SR+ condition (–1.30 µV, SD = 0.31) as compared to the SR– condition (–1.01 µV, SD = 0.29) and the CC (–0.41, SD = 0.29). These differences in the amplitude of the N450 were significant (SR+ vs. CC: $F(1, 17) = 14.57$, $MS_e = 2.41$, $p < .001$; SR– vs. CC: $F(1, 17) = 6.77$, $MS_e = 1.80$, $p < .05$; and SR+ vs. SR–: $F(1, 17) = 6.26$, $MS_e = 0.95$, $p < .05$).

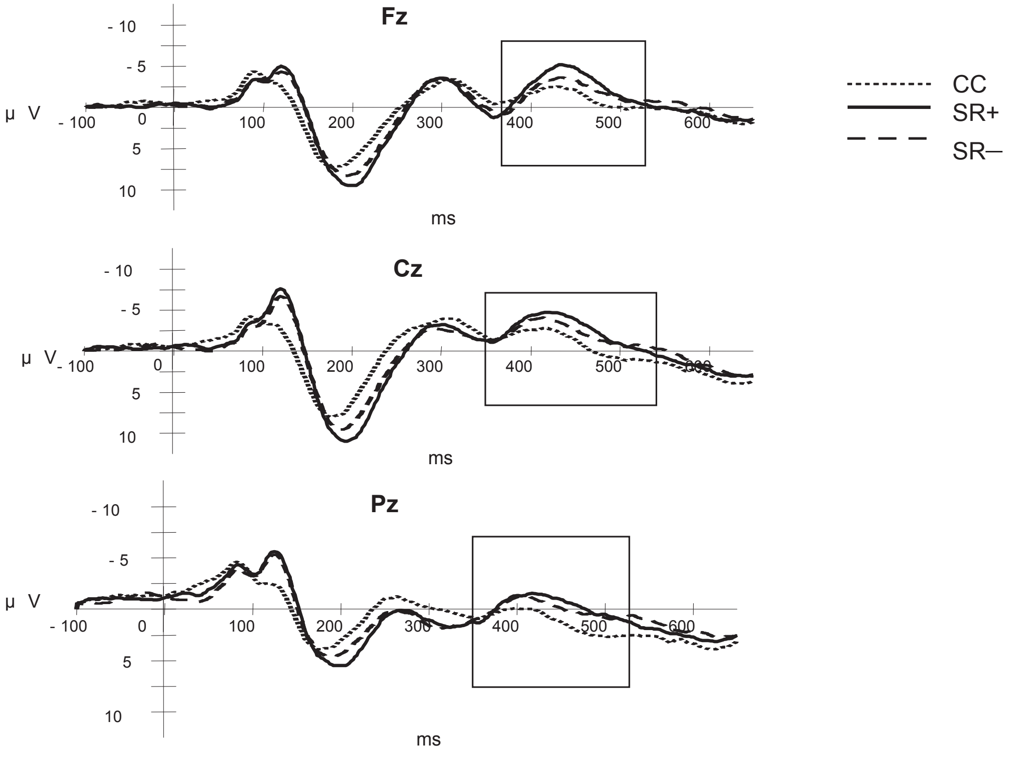

*Figure 6*. Grand average ERPs time-locked to the onset of the stimulus. Dotted lines depict the control condition (CC), solid lines depict the semantically related condition (SR+), and dashed lines depict the semantically unrelated condition (SR–). The time-window 350 – 550 ms that was used for statistical analysis is framed. Areas showing statistically significant differences between conditions are indicated by rectangles.

To localize this effect, we ran two separate ANOVAs. First, a 3 (conditions) by 2 (anterior vs. posterior position) by 10 (electrode sites) ANOVA was run to investigate whether the N450 had a more frontal or posterior distribution.

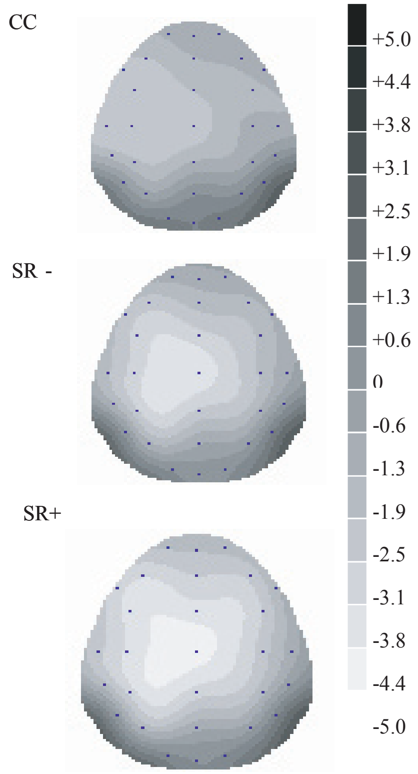

*Figure 7.* Topographic maps of the N450 amplitude between 400 and 500 ms after target picture onset. Negative regions depicted in light gray.

Second, a 3 (conditions) by 2 (left vs. right hemisphere) by 10 (electrode sites) ANOVA was run to investigate whether the amplitude of the N450 was larger over the left or right hemisphere. Both of these ANOVAs revealed significant effects of Position (F(1, 17) = 15.22, $MS_e = 27.74$, p < .01, for anterior vs. posterior; and F(1, 17) = 6.45, $MS_e = 10.39$, p < .05, for left vs. right). To refine this analysis, we ran a separate 3 (conditions) by 3 (anterior vs. central vs. posterior position) by 6 (electrode sites) ANOVA. This analysis yielded a significant effect of Position (F(2, 34) = 31.23, $MS_e = 12.83$, p < .01), with the highest amplitude of the N450 over anterior sites (–1.63 µV, SD = 1.80), followed by central and posterior sites (–1.06 µV, SD = 1.66; 0.51 µV, SD = 1.33, respectively). To be more precise, the N450 in our study had a left-fronto-central distribution (see Figure 7).

## Discussion

The goal of the present study was to investigate how verbal self-monitoring is affected by auditory distractors and whether the ERN is sensitive to verbal stimulus manipulation. As expected, semantically related distractors caused a larger interference effect than unrelated distractors. Surprisingly, the presence of distractors had no influence on error rate: Participants did not commit more errors in the semantically related than in the semantically unrelated condition. However, as hypothesized, the amplitude of the ERN was significantly larger when distractors were semantically related to the target picture than when they were unrelated. Interestingly, we also demonstrated an ERN after misses, in the absence of overt motor responses (see also Ganushchak & Schiller, 2006). During misses, participants failed to detect the target phoneme in the covertly generated picture name. It is possible that after the decision not to respond, further processing of the stimulus revealed that there actually was a target phoneme in the picture name. This, in turn, resulted in the mismatch between actual and desired response, which led to a higher conflict in the misses as compared to the correct nogo responses. Hence, an ERN was generated. Unfortunately, we did not have enough trials to investigate whether or not the ERN after misses was affected by distractor relatedness. For correct go trials, we obtained a larger negative deflection, peaking around 450 ms after stimulus onset, for semantically related distractors than semantically unrelated ones. The N450 was smallest for the control condition, where auditory distractors were absent. These findings will be discussed in more detail below.

In accordance with a number of previous studies (e.g., Damian & Martin, 1999; Glaser & Düngelhof, 1984; Lupker, 1979; Roelofs, 1992; Schriefers et al., 1990) we demonstrated that semantically related auditory distractors induced an interference effect relative to semantically unrelated distractors. Semantically related distractors presumably co-activate, through the spreading of activation, multiple concepts that are semantically related to one another (Maess, Friederici, Damian, Meyer, & Levelt, 2002). Hence, there are multiple plausible entries that are simultaneously active and compete for lexical selection (but see Finkbeiner et al., 2006 and Mahon et al., 2007). The verbal self-monitor needs to verify on-line whether or not the correct lexical item was chosen from the pool of competing items, which in turn leads to slower responses. In the unrelated condition, however, such verification might not be as relevant, since unrelated distractors do not activate related concepts, and therefore less competition may be present at the time of the response.

Surprisingly, distractors slowed down the functioning of the monitor in our study, but did not compromise its working completely. According to the perceptual loop theory, the verbal monitor proceeds through the speech comprehension system (Levelt et al., 1999). The same system is used for processing auditory input from other speakers. Thus, in presence of auditory distractors, the comprehension system must process auditory input and simultaneously monitor inner speech, which inevitably should lead to a higher error rate. However, in the present study, auditory distractors did not lead to more errors. Previous studies showed inconsistent changes of error rate in relation to distractors. Interestingly, it seems that studies that employed visual distractors report increases in error rate in the presence of semantically related distractors, as compared to semantically unrelated ones (e.g., Bloem et al., 2004; Damian & Bowers, 2003). However, studies using auditory distractors seem to find no effect of distractor relatedness on error rate (e.g., Jescheniak et al., 2002). The fact that auditory distractors do not cause higher error rates may be due to more independence of the verbal monitor from the comprehension system than the perceptual loop theory presumes. This possibility is also implied by research on aphasic patients demonstrating a double dissociation between the comprehension and the monitoring system (e.g., intact comprehension but impaired monitoring or vice versa). For example, Marshall, Robson, Pring, and Chiat (1998) described patients with jargon aphasia who failed to detect their neologisms despite their relatively preserved comprehension, while Marshall, Rappaport, and Garcia-Bunuel (1985) reported aphasic patients with severely impaired auditory comprehension who nevertheless exhibited well-preserved self-monitoring skills.

Our behavioral results are supported by the EEG data. First, in stimulus-locked averages we found that the amplitude of the N450 was modulated by the relatedness of distractors. Semantically related distractors led to higher N450 amplitudes than semantically unrelated distractors. In previous studies, the N450 was associated with the amount of conflict; the larger the conflict, the larger the amplitude of the N450 (West & Alain, 1999). In the present study, semantically related distractors presumably increased conflict by activating multiple concepts thus increasing the amplitude of the N450. Second, we showed that the amplitude of the response-locked ERN was also contingent on the relatedness of distractors. The amplitude of the ERN was smallest when distractors were absent and largest when distractors were semantically related to the target picture. In the present study, the semantically unrelated distractors did not lead to a significant increase of the amplitude of the ERN, as compared to the control condition. It might be that semantically unrelated distractors did not elicit enough conflict to be detected by changes in the ERN. It is also possible, however, that this difference would become significant at a higher error rate than in the present study.

The ERN has been associated with the amount of conflict between plausible responses (Botvinick et al., 2001). Our data are in accordance with this assumption. As mentioned above, semantically related distractors activate multiple concepts that are related to the target, thereby leading to greater competition between them. Hence, in the semantically related condition at the time of a response there may be more conflict between competing responses than in the semantically unrelated condition or in the absence of distractor words, thereby leading to a larger ERN amplitude. Due to the simultaneous activation of competing lexical items, the verbal self-monitor presumably needs to be more alert during the semantically related condition than during unrelated and control conditions in order to validate whether the given response was correct or erroneous.

The design of the present experiment makes it difficult to disentangle if the increased amplitude of the ERN was due to increased response conflict or other factors (e.g., awareness of errors). The presence of the semantically related distractors presumably increases conflict at the level of lexical retrieval but does not necessarily lead to response conflict. We argued that the semantically related distractors activate multiple lexical entries, which in turn lead to multiple potential responses active at the same time and therefore, albeit indirectly, lead to higher response conflict. However, we cannot exclude the possibility that there was only lexical conflict present. Interestingly, West and colleagues (2004) demonstrated that the anterior cingulated cortex (ACC) as well as N450 are sensitive not only to response conflict, as generally accepted, but also to non-response conflict (i.e. conflict that arises prior to response processing, e.g., at phonetic and/or semantic levels). The dipole model used to fit the neural generators of the N450 was similar to that used to fit the ERN. These findings suggest that the ERN might be sensitive not only to response conflict but also to non-response conflict. In light of these findings, it is possible that response conflict is not a necessary condition to elicit the ERN, and that the increase in amplitude of the ERN found in the present study could have been due to the increase of lexical conflict.

Alternatively, previous research showed that the ERN can be modulated by the perceived awareness of errors (e.g., Scheffers & Coles, 2000; for an alternative view see Nieuwenhuis et al., 2001; O'Connell et al., 2007). In the present study, one may argue that the difference in ERN amplitude between conditions was due to differences in awareness of the errors that were made. For example, participants could have been more aware of their errors in the semantically related condition than in semantically unrelated and control conditions, thereby causing higher ERN amplitudes. However, the subjective confidence data, collected in the present study, showed that participants were equally aware of their errors in all experimental conditions. Therefore, we conclude that the increase in ERN amplitude during the semantically related condition was most likely due to the simultaneous activation of competing lexical items and not because participants could more easily recognize their errors in the semantically related than in the semantically unrelated condition.

To conclude, there is increasing evidence that verbal errors are similar to action errors of a more general performance monitor. As stated above, in the current study, we showed that the ERN is sensitive to the presence of lexical conflict. Similarly, in previous studies (i.e., Ganushchak & Schiller, 2006) we reported a typical decrease of the ERN amplitude after errors of verbal monitoring under conditions of time pressure. Furthermore, Möller and colleagues (2007) – employing the inverse source localization method – identified a medial frontal generator in the supplementary motor area (SMA) as the main source of the negativity preceding erroneous vocalizations. A typical ERN has been located within the ACC/SMA regions (e.g., Dehaene, Posner, & Tucker, 1994). Based on these findings, we suggest here that the verbal monitor works in a similar way as a general performance monitor. It is possible that during verbal monitoring, as well as during executive action monitoring, a copy of the on-line response is created and compared to the representation of the correct response (see Levelt et al., 1999). If there is a mismatch between them, an error signal is generated and corrective processes can be started – similar to action monitoring (see Desmurget & Grafton, 2000; Rodríguez-Fornells et al., 2002). We would like propose here that verbal monitoring might be a special case of general performance monitoring instead of a completely separate process.

In summary, we showed that the ERN can successfully be elicited by errors of verbal monitoring and is sensitive to verbal stimulus manipulations. This provides further proof that the ERN could be used as an electrophysiological marker of error processing in language research. However, in the present study the required responses were button presses. We believe that the majority of errors observed in the current study were errors of the verbal monitor and based on the incorrect decision about the target phoneme (as demonstrated by the ERN-like responses on misses). However, we cannot completely rule out the possibility that some of the errors could have been due to action slips and not slips of verbal monitoring *per se*. Therefore, more research is needed to be able to make a clearer dissociation between errors of verbal and other action monitoring.

# 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), Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

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.

Bloem, I., Van den Boogaard, S., & La Heij, W. (2004). Semantic facilitation and semantic interference in language production: further evidence for the conceptual selection model of lexical access. *Journal of Memory and Language, 51*, 307-323.

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.

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.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407-428.

Damian, M. F. & Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychonomic Bulletin and Review, 10*, 111-117.

Damian, M. F. & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 345-361.

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., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: a tutorial. *Biological Psychology, 51*, 87-107.

Falkenstein, M., 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.

Finkbeiner, M., Gollan, T. H., & Caramazza, A. (2006). Lexical access in bilingual speakers:

What's the (hard) problem? *Bilingualism: Language and Cognition, 9*, 153-166.

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

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.

Glaser, W. R., & Düngelhof, F. J. (1984). The time course of picture-interference. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 640-654.

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.

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

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. S. (1999). A theory of lexical access in speech production. *Behavioral and 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.

Liotti, M., Woldorff, M. G., Perez, R. 3rd, & Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia, 38*, 701-711.

Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory & Cognition, 7*, 485-495.

Maess, B., Friederici, A. D., Damian, M., Meyers, A. S., & Levelt, W. J. M. (2002). Semantic category interference in overt picture naming: sharpening current density localization by PCA. *Journal of Cognitive Neuroscience, 14*, 455-462.

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical Selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 33, 503-535.

Marshall, J., Robson, J., Pring, T., & Chiat, S. (1998). Why does monitoring fail in jargon aphasia? Comprehension, judgment, and therapy evidence. *Brain and Language, 63*, 79-107.

Marshall, R.C., Rappaport, B.Z., & Garcia-Bunuel, L. (1985). Self-monitoring behavior in a case of severe auditory agnosia with aphasia. *Brain and Language*, *24*, 297-313.

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

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

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.

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P. H., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology, 38*, 752-760.

O'Connell, R. G., Dockree, P. M., Bellgrove, M. A., Kelly, S. P., Hester, R., Garavan, H., Robertson, I. H., & Foxe, J. J. (2007). The role of cingulate cortex in the detection of errors with and without awareness: a high-density electrical mapping study. *European Journal of Neuroscience*, 25, 2571-2579.

Oomen, C. C. E., Postma, A., & Kolk, H. (2001). Prearticulatory and postarticulatory self-monitoring in Broca's aphasia. *Cortex, 37*, 627-641.

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.

Quillian, M. R. (1968). Semantic memory. In M. Minsky (ed.), *Semantic information processing* (pp. 227-270). Cambridge, MA: MIT Press.

Rodrígues-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.

Scheffers, M. K., & Coles, M. G. H. (2000). Performance monitoring in a confusing world: error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology, 26*, 141-151.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: picture-word interference studies. *Journal of Memory and Language, 26*, 86-102.

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.

Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 686-698.

Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 896-918.

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.

West, R., & Alain, C. (1999). Event-related neural activity associated with the Stroop task. *Cognitive Brain Research, 8*, 157-164.

West, R., Bowry, R., & McConville, C. (2004). Sensitivity of medial frontal cortex to response and nonresponse conflict. *Psychophysiology, 41*, 739-748.

West, R. (2003). Neural correlates of cognitive control and conflict detection in the Stroop and digit-location tasks. *Neuropsychologia, 41*, 1122-1135.

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.

# Appendix

*List of stimuli used in the experiment*

| Pictures | Distractors | | | |
|---|---|---|---|---|
| | go trials | | nogo trials | |
| | semantically related | semantically unrelated | semantically related | semantically unrelated |
| baard ('beard') | wenkbrauw ('brow') | wagen ('wagon') | sik ('goatee') | wok ('wok') |
| bloem ('flower') | kruid ('herb') | veer ('ferry') | plant ('plant') | bus ('bus') |
| broek ('trousers') | vest ('waistcoat') | stal ('stable') | trui ('sweater') | kruid ('herb') |
| fiets ('bicycle') | brommer ('moped') | medaillon ('medallion') | step ('scooter') | tap ('tap') |
| film ('film') | dia ('slide') | sik ('goatee') | rol ('spool') | lip ('lip') |
| fles ('bottle') | kan ('pitcher') | fee ('fairy') | glas ('glass') | oor ('ear') |
| heks ('witch') | elf ('elf') | berg ('mountain') | fee ('fairy') | hes ('smock') |
| hemd ('shirt') | rok ('skirt') | wang ('cheek') | shirt ('shirt') | schors ('bark') |
| hoorn ('horn') | fluit ('flute') | muts ('bonnet') | viool ('violin') | kruik ('jar') |
| jurk ('dress') | bloes ('blouse') | been ('leg') | kous ('stocking') | beurs ('fair') |
| kaart ('card') | fiche ('token') | bazaar ('bazar') | roulette ('roulette') | collier ('necklace') |
| kaas ('cheese') | paté ('paté') | metro ('metro') | ham ('ham') | hol ('den') |
| kar ('cart') | slee ('sled') | muil ('slipper') | wagen ('wagon') | wenkbrauw ('brow') |
| knie ('knee') | voet ('foot') | rol ('spool') | been ('leg') | bank ('couch') |
| kraan ('faucet') | tap ('tap') | step ('scooter') | douche ('shower') | fiche ('token') |
| kroon ('crown') | scepter ('scepter') | roulette ('roulette') | baret ('beret') | kano ('canoe') |
| maan ('moon') | heelal ('cosmos') | douche ('shower') | hemel ('sky') | vinger ('finger') |
| markt ('market') | beurs ('fair') | viool ('violin') | bazaar ('bazaar') | gondel ('gondola) |
| mes ('knife') | lepel ('spoon') | ladder ('ladder') | vork ('fork') | haas ('hare') |

| | | | | |
|---|---|---|---|---|
| muur (‘wall’) | dak (‘roof’) | plant (‘plant’) | raam (‘window’) | voet (‘foot’) |
| nest (‘nest’) | hol (‘den’) | trui (‘sweater’) | stal (‘stable’) | rok (‘skirt’) |
| neus (‘nose’) | oor (‘ear’) | raam (‘window’) | wang (‘cheek’) | bloes (‘blouse’) |
| pan (‘pan’) | wok (‘wok’) | das (‘badger’) | schotel (‘dish’) | scepter (‘scepter’) |
| pet (‘cap’) | hes (‘smock’) | huig (‘uvula’) | muts (‘bonnet’) | fluit (‘flute’) |
| plank (‘board) | staaf (‘rod’) | kous (‘stocking’) | balk (‘beam’) | sjaal (‘scarf’) |
| pot (‘pot’) | kruik (‘jar’) | kruk (‘stool’) | vaas (‘vase’) | staaf (‘rod’) |
| riem (‘belt’) | sjaal (‘scarf’) | vaas (‘vase’) | das (‘badger’) | dia (‘slide’) |
| ring (‘ring’) | collier (‘necklace’) | schotel (‘dish’) | medaillon (‘medallion’) | brommer (‘moped’) |
| rots (‘cliff’) | leem (‘loam’) | ham (‘ham’) | berg (‘mountain’) | elf (‘elf’) |
| schip (‘ship’) | kano (‘canoe’) | baret (‘beret’) | boot (‘boat’) | kaak (‘jaw’) |
| schoen (‘shoe’) | klomp (‘clog’) | balk (‘beam’) | muil (‘slipper’) | slee (‘sled’) |
| slang (‘snake’) | haas (‘hare’) | vork (‘fork’) | hagedis (‘lizard’) | heelal (‘cosmos’) |
| snor (‘mustache’) | kaak (‘jaw’) | stam (‘trunk’) | haar (‘hare’) | dak (‘roof’) |
| tak (‘branch’) | schors (‘bark’) | shirt (‘shirt’) | stam (‘trunk’) | vest (‘waistcoat’) |
| tong (‘tong’) | lip (‘lip’) | haar (‘hair’) | huig (‘uvula’) | leem (‘loam’) |
| tram (‘tram’) | gondel (‘cable car’) | hagedis (‘lizard’) | veer (‘ferry’) | klomp (‘clog’) |
| trap (‘stairs’) | lift (‘lift’) | boot (‘boat’) | ladder (‘ladder’) | lepel (‘spoon’) |
| trein (‘train’) | bus (‘bus’) | duim (‘thumb’) | metro (‘metro’) | pate (‘pate’) |
| troon (‘throne’) | bank (‘couch’) | glas (‘glass’) | kruk (‘stool’) | kan (‘pitcher’) |
| vuist (‘fist’) | vinger (‘finger’) | hemel (‘sky’) | duim (‘thumb’) | lift (‘lift’) |

# Chapter 5

# Motivation and semantic context affect brain error-monitoring activity: An event-related brain potentials study[9]

## Abstract

During speech production, we continuously monitor what we say. In situations in which speech errors potentially have more severe consequences, e.g. during a public presentation, our verbal self-monitoring system may pay special attention to prevent errors than in situations in which speech errors are more acceptable, such as a casual conversation. In an event-related potential study, we investigated whether or not motivation affected participants' performance using a picture naming task in a semantic blocking paradigm. Semantic context of to-be-named pictures was manipulated; blocks were semantically related (e.g., *cat*, *dog*, *horse*, etc.) or semantically unrelated (e.g., *cat*, *table*, *flute*, etc.). Motivation was manipulated independently by monetary reward. The motivation manipulation did not affect error rate during picture naming. However, the high-motivation condition yielded increased amplitude and latency values of the error-related negativity (ERN) compared to the low-motivation condition, presumably indicating higher monitoring activity. Participants showed semantic interference effects in reaction times, error rates and the ERN amplitude, presumably indicating that semantic relatedness induces more conflict between possible verbal responses.

---

[9] This chapter is based on Ganushchak, L. Y. & Schiller, N. O. (2008). Motivation and semantic context affect brain error-monitoring activity: An event-related brain potentials study. *NeuroImage*, 39, 395-405.

# Introduction

Speaking is a very fast and seemingly effortless process. In overt speaking, we produce up to 150 words per minute. However, the speech error rate in normal individuals is not more than one error in every 1,000 words (Levelt, 1989). Such low error rates may be the result of a verbal self-monitor that detects and corrects errors. The most prominent theory of verbal monitoring is the *perceptual-loop theory* proposed by Levelt (1983, 1989). According to this theory, there is a single, central verbal monitor that checks the message for its appropriateness, inspects the speech plan, and detects errors prior to its articulation via the speech comprehension system (Postma & Noordanus, 1996; Schiller, 2005, 2006; Schiller, Jansma, Peters, & Levelt, 2006; Wheeldon & Levelt, 1995; Wheeldon & Morgan, 2002), as well as after speech has become overt (Postma, 2000).

As stated above, the error rate under normal circumstances is very low indicating that verbal monitoring generally has low susceptibility to interference. However, there may be specific circumstances that produce interference with the working of the monitor. For instance, it is possible that in situations in which speech errors potentially have more significance because they are less acceptable, e.g. during giving an interview vs. having a casual conversation, the verbal self-monitoring system works harder in order to prevent errors. One question to ask is about the role of the verbal context in which a conversation takes place. If we hear or see information that is related to what we are planning to say, does that information interfere with verbal monitoring, thereby leading to more erroneous speech output? We will try to answer this question in the present study.

One way to study monitoring is by looking at error monitoring. An electrophysiological measure related to error processing is the so-called *Error-Related Negativity* (ERN; Falkenstein et al., 1991; Gehring et al., 1993), 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 et al., 1995; Holroyd & Yeung, 2003; Scheffers et al., 1996). Originally, the ERN was thought to reflect conscious error detection (Bernstein et al., 1995). However, according to the conflict hypothesis, the ERN arises not due to error detection *per se* but rather as a result of response conflict that arises when multiple responses compete for selection (Botvinick et al., 2001; Carter et al., 1998). Presence of conflicting responses reflects situations in which 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).

Interestingly, a number of studies demonstrated the influence of emotional/motivational factors on the ERN (e.g., Boksem et al., 2006; Luu et al., 2000; Pailing & Segalowitz, 2004; Ullsperger & Von Cramon, 2004). The general finding is that the ERN increases when monetary incentives are offered for accuracy (Gehring et al., 1993; Hajack et al., 2005; Pailing & Segalowitz, 2004). For instance, Pailing and Segalowitz (2004) manipulated value of response error by selectively financially rewarding one type of response over another in a four-choice letter task. Pailing and Segalowitz found that more costly types of errors were associated with higher amplitude of the ERN. However, this dependency was only present for participants who scored high on neuroticism. Hajack and colleagues (2005) also investigated whether the ERN is sensitive to value of errors. They manipulated motivational significance or errors by administrating monetary punishment for them. Consistent with previous studies, these authors showed that the ERN was significantly larger on high-value errors than low-value errors. Consistently with the EEG studies, Ullsperger and Von Cramon (2004) performed an fMRI study in which they also modulated the relevance of errors by a financial reward manipulation. Ullsperger and Von Cramon found that error-related activation in posterior fronto-medial cortex, previously shown to be involved in performance monitoring, was modulated by error relevance.

Most studies on the ERN investigate the working of action monitoring. In the present study, however, we use the ERN to explore the workings of the *verbal* monitoring system. There are only few studies that looked at the ERN after verbal errors (see Ganushchak & Schiller, 2006, in press; Masaki et al., 2001; Möller, Jansma, Rodríguez-Fornells, & Münte, 2007; Sebastián-Gallés, Rodríguez-Fornells, Diego-Balaguer, & Díaz, 2006), which we will briefly review below.

Masaki and colleagues (2001) examined whether or not the ERN occurs in relation to speech errors in the Stroop color-word task. Participants in their study were instructed to overtly name the color of each stimulus as quickly and accurately as possible. Masaki and colleagues found an ERN-like response after speech errors, e.g. when participants named the wrong color.

Sebastián-Gallés and colleagues (2006) assessed Spanish-dominant and Catalan-dominant bilinguals using an auditory lexical decision task in Catalan. The authors showed that Spanish-dominant bilinguals had great difficulty in rejecting experimental non-words, and did not show an ERN in their erroneous non-word decisions either. According to Sebastián-Gallés et al. this suggests that Spanish-dominant bilinguals activated the same lexical entry from experimental words and non-words (in the experimental stimuli, the vowel change involved a Catalan-specific /e – ɛ/ contrast) and therefore showed no differences between correct and erroneous responses. In contrast, Catalan-dominant bilinguals demonstrated a clear ERN.

Recently, Möller et al. (2007) employed a laboratory task known to elicit speech errors to investigate verbal monitoring. In this task, participants are presented with inductor word pairs such as ‘*ball doze*’, ‘*bash door*’, and ‘*bean deck*’, which are followed by a target word pair such as ‘*darn bore*’ (see Motley et al., 1982). The reversal of initial phonemes in the target pair compared to the inductor pairs is supposed to lead to speech errors such as '*barn door*'. Möller and colleagues asked their participants to covertly read the inductor word pairs and vocalize the target word pair immediately preceding a response cue. They found a negative deflection on error trials, as compared to correct trials, preceding the response cue. Möller et al. proposed that this activity reflects the simultaneous activation of competing speech plans. However, these authors do not make an explicit link between the negativity they found in their study and the ERN.

Ganushchak and Schiller (2006) used a phoneme-monitoring task to investigate the effects of verbal monitoring under time pressure. Participants were presented with pictures and had to indicate whether the target phoneme was present in the name of the picture. For example, if the presented picture was *table* and target phoneme was /t/, then participants had to press a button; however, if the target phoneme was /m/, they had to withhold their response. Ganushchak and Schiller obtained an ERN following verbal errors which showed a typical decrease in its amplitude under severe time pressure.

In more recent study by the same authors (Ganushchak & Schiller, in press), a similar phoneme-monitoring task was employed to investigate the effect of auditory distractors on verbal monitoring. Participants were requested to press a button when a target phoneme was present in the pictures' name. However, simultaneously with the picture participants heard a semantically related distractor, a semantically unrelated distractor, or no distractor at all. Ganushchak and Schiller observed a larger ERN when auditory distractors were semantically related to the picture than when distractors were unrelated or no distractors were present at all. Presence of distractors, by activating more related concepts, presumably increased conflict at the time of response and therefore led to higher amplitudes of the ERN. This result may indicate that the ERN after verbal errors, as well as after general performance errors, is sensitive to conflict present at the time of response (see Botvinick et al., 2001). The goal of the present study was to further investigate the relationship between the ERN and verbal monitoring.

In the study described above, Ganushchak and Schiller (in press) used a phoneme-monitoring task in which button-press responses were required, and not pure verbal responses. In contrast, in the current study, we employed a blocked picture naming task in which recorded responses were overt verbal responses. The blocked naming paradigm manipulates the context in which to-be-named pictures appear. In semantically related blocks, pictures from the same semantic category appear on successive trials, for example *table, chair, couch*, and *closet*. In contrast, in semantically unrelated, mixed blocks, pictures from different semantic categories appear one at a time, for instance *table, snake, apple*, and *car*. Speakers take longer to name pictures from the same semantic category than from different categories. This increase in naming latencies is attributed to the increased competition for lexical selection from semantically related competitors (Belke, Meyer, & Damian, 2005; Damian, Vigliocco, & Levelt, 2001; Levelt et al., 1999; Schnur, Schwartz, Brecher, & Hodgson, 2006).

In our own study, we employed this semantic blocking picture naming paradigm to investigate the effects of the semantic context on verbal self-monitoring and the ERN. How does semantic blocking relate to the verbal self-monitor? According to the Levelt's perceptual loop theory (1983, 1989), the verbal self-monitoring system not only monitors for errors, but also for semantic appropriateness. In semantically related blocks, the monitor presumably checks whether or not the correct alternative has indeed been chosen as the target response from the set of competing items (Maess et al., 2002). This check is less urgent in the mixed blocks, where co-activation of competing items supposedly occurs less than in the semantically related blocks. In the present study, we expected to find more errors and slower reaction times while naming pictures in semantically related blocks, compared to naming pictures in mixed blocks.

Moreover, we were interested to examine what happens when people commit errors. Will error signals be different when an error occurred during semantically related blocks compared to error signals during mixed blocks? In semantically related blocks, as opposed to mixed blocks, there are multiple semantically related entries active which compete for the lexical selection, thus leading to a higher conflict between various semantically related competitors. The ERN is sensitive to the amount of conflict present at the time of the response (Botvinick et al., 2001). Therefore, one may hypothesize that the amplitude of the ERN will be larger following errors in the semantically related blocks than following errors in the mixed blocks.

In the present study, we investigated not only the effects of the semantic context, but also effects of motivation on verbal monitoring and the ERN. Pictures were presented in two colors, i.e. orange and purple, and participants were told that the more errors they make while naming orange pictures, the smaller their financial reward for participation would be, i.e. the high-motivation condition. If naming errors were made on purple pictures, participants received neither financial punishment, nor financial reward: the low-motivation condition. In the high-motivation condition, making errors had more consequences for participants than making errors during the low-motivation condition. Therefore, we expected to find higher amplitudes of the ERN during the high-motivation than the low-motivation condition. In previous research, it has been shown that the amplitude of the ERN is increased when response accuracy is emphasized over speed (e.g., Falkenstein et al., 2000; Gehring et al., 1993). Furthermore, participants were instructed to name pictures as quickly and as accurately as possible in both high-motivation and low-motivation conditions in the current study. Therefore, we did not expect to find differences in reaction times in naming pictures in the high-motivation and low-motivation conditions.

Note that in semantically related and mixed blocks participants were instructed to name pictures as fast and as accurately possible. However, between these two types of blocks, we expected to find reaction time differences because the semantic interference effect is a robust effect which occurs automatically without participants' awareness. In contrast, the behavioral differences in the high-motivation versus the low-motivation conditions most likely do not occur due to automatic processes but due to strategies participants applied, for instance, in order to gain accuracy in the high-motivation condition by slowing the responses down, compared to the low-motivation condition. Therefore, by stressing the importance of reacting as fast and as accurately as possible we hoped to counteract the potential differences in response strategies in the high-motivation versus the low-motivation conditions, and therefore expected no behavioral differences between these conditions.

To summarize, we predicted that participants will be slower and make more errors naming pictures in semantically related blocks than in mixed blocks. Moreover, we expected to obtain an ERN after erroneous trials across all conditions. However, the amplitude of the ERN should increase while naming pictures in semantically related blocks compared to mixed blocks. Finally, the amplitude of the ERN should also be larger during the high-motivation condition than low-motivation condition.

# Methods

## *Participants*

Twenty-two students of Maastricht University (19 female) took part in the experiment. All participants were right-handed, native Dutch speakers, and had normal or corrected-to-normal vision. Participants gave written informed consent prior to participating in the study. They received a small financial reward for their participation in the experiment. Due to technical problems, the data of three participants were lost.

## *Materials*

Seventy-five simple line drawings were selected from fifteen semantic categories of five exemplars each (see Appendix). Items for categories were selected minimizing within-category visual similarity. In a control study, we asked another 20 participants to ignore the semantic similarity of the pictures and judge all 150 pairs of pictures in terms of visual similarity. Participants were required to rate visual similarity on a five-point scale (1 = not similar at all, 5 = very similar). Within-category similarity was only slightly higher (mean: 2.4) than between-category similarity (mean: 1.6). These judgments are similar to the ones reported in Damian et al. (2001), who obtained a within-category similarity of 2.4 and a between-category similarity of 1.9.

Picture names were on average 1.3 syllables long (range: 1 – 3) and 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, each picture was presented once in orange and once in purple (i.e. 255, 127, 0 and 158, 73, 161, respectively, on the RGB scale), and degraded with 9 pt dashes and 16 pt spaces, and the weight of the lines was 3 pt. Pictures were dashed to make the task a bit more difficult and provoke participants to commit more speech errors. In the simple picture naming task, participants could recognize and name all dashed pictures correctly (for examples of stimuli, see Figure 1 on the back side of the cover).

### *Design*

The experiment consisted of learning, practice, and main task. During the learning and practice phases, participants saw all pictures in black-on-white in the middle of the screen. In the learning phase, each picture was presented simultaneously with its label written underneath it. Pictures stayed on the screen for 2,000 ms. In the practice task, participants saw the same pictures without the picture names written underneath. In practice and main tasks, 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, i.e. a picture. Pictures disappeared from the screen as soon as the voice key was activated or after 500 ms maximally. The inter-trial interval was variable, depending on the response latency, such that each trial had a total duration of 2,000 ms.

For the main task, five-item sets were formed. In semantically related blocks, five exemplars from the same semantic category were presented together in a set (e.g., *giraffe* ‘giraffe’, *kameel* ‘camel’, *hert* ‘deer’, *olifant* ‘elephant’, *zebra* ‘zebra’). In mixed blocks, the five-item sets comprised exemplars from different semantic categories (e.g., *giraffe* ‘giraffe’, *bank* ‘couch’, *arm* ‘arm’, *piano* ‘piano’, *citroen* ‘lemon’). Each block consisted of five pictures that were repeated four times in different order, resulting in blocks of 20 trials each. In total, there were 15 semantically related and 15 mixed blocks. Each block was presented twice: once with all pictures in a block colored orange and once with all pictures colored purple. In such a way, motivation was manipulated on a block basis. Each participant saw 15 semantically related and 15 mixed blocks of 20 trials each and repeated in both colors, i.e. 30 blocks x 20 trials x 2 colors = 1,200 trials altogether. The order of pictures was pseudo-randomly varied for each participant in such a way that identical pictures did not appear on consecutive trials. The order of blocks followed a Latin square design.

### *Procedure*

Participants were tested individually while seated in a sound-proof room. They were presented with the learning phase, the practice task, and, finally, the main task. During the learning task, participants were familiarized with the pictures and their corresponding names. In the practice task, participants were asked to overtly name the pictures with the labels they learned during the learning phase. If errors were made, participants were told about their mistakes and correct responses were provided by the experimenter. During the main task, participants were asked to overtly name the same pictures as in the practice task. Participants were told that their total financial reward for participation would depend on their performance during the task. If they made errors during naming pictures presented in orange, they would receive €12.50, €10, €7.50, or €5 depending on amount of errors made. For errors made on pictures presented in purple, there was neither financial gain nor punishment. Participants were asked to overtly name all orange and purple pictures as fast and as accurately possible. At the end of the experiment, all participants received the full financial reward independent of their performance.

### 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 electro 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 mastoid. Eye movements were recorded to allow off-line rejection of contaminated trials. Lateral eye movements were measured using a bipolar montage of electrodes placed on the right and left external canthus. Eye blinks and vertical eye movements were measured using bipolar montage of electrodes placed above and below the left eye. Impedance level for all electrodes was kept below 5kΩ.

### Data analysis

Epochs of 1,300 ms (from –400 ms to +900 ms) were computed. A 100 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 at the voice-key onset of the erroneous response. For the correct trials, averaging was done for the voice-key onset of the correct responses. To compute the difference between correct and error trials, a mean area amplitude analysis was used in a time window between 0 and 100 ms after response onset. For this analysis, we used a mean area amplitude analysis since it was impossible to identify peaks on correct trials. The amplitude 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 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 after the response and the most positive peak from –50 to 0 ms preceding the ERN (Falkenstein et al., 2000). The latency of the ERN was defined as a point in time when the negative peak was at its maximum. The amplitude and the latency of the ERN were recorded for each condition at electrode sites Fz, FCz, Cz, and Pz.

## Results

### Behavioral data

Latencies shorter than 300 ms and longer than 1,500 ms were excluded from the analysis. Effects of Motivation and Semantic Relatedness on naming latencies were assessed by repeated measures ANOVAs. These analyses revealed a significant effect of Semantic Relatedness (F(1,

18) = 10.43, $MS_e = 457.42$, $p < .001$), but no effect of Motivation (F(1, 18) = 1.08, $MS_e = 553.29$, n.s.) and no interaction between the two factors (F(1, 18) = 2.71, $MS_e = 373.42$, n.s.). Naming latencies were 15 ms longer during semantically related blocks than during mixed blocks (see Table 1 for mean naming latencies).

Similar analyses are reported for number of errors. Participants made on average 3.0% phonological and semantic errors. Eighty percent of these errors were phonological errors and only 20% were semantic errors. An example of a semantic error is when a participant incorrectly named the picture of a *giraffe* as *zebra*. An example of a phonological error is the non-word *drood* instead of *brood* 'bread'. Trials on which participants failed to give a response or the voice key was triggered by an inappropriate response (e.g., sneezing or coughing) were not included in the analysis (0.8 % of all trials). A repeated measures ANOVA with number of errors as dependent variable revealed a significant effect of Semantic Relatedness (F(1, 18) = 7.77, $MS_e = 17.55$, $p < .01$). Participants made more errors during semantically related blocks than during mixed blocks (see Table 1 for error rates). There was no effect of Motivation nor was there an interaction between Motivation and Semantic Relatedness (both Fs < 1).

Table 1. Overview of the behavioral data. Mean (± standard deviation) reaction times (in ms) and error rates (% relative to the number of trials per condition) as a function of motivation and semantic context manipulation.

| | *Semantically related blocks* | *Mixed blocks* | *High-motivation* | *Low-motivation* |
|---|---|---|---|---|
| Reaction times | 683 (65) | 668 (63) | 673 (64) | 678 (65) |
| Error rates | 3.3 (7) | 2.8 (6) | 2.9 (6) | 3.0 (6) |

### *Electrophysiological data*

Inspection of the grand averages for error trials revealed a clear negative deflection on error trials but not on correct trials (see Figures 2 and 3). Figure 4 shows the scalp distribution of the ERN by means of topographic maps. A repeated measures ANOVA with mean ERN amplitude as dependent variable revealed a significant effect of Condition (correct vs. incorrect; F(1, 18) = 6.79, $MS_e = 40.42$, $p < .05$). As expected, the ERN appeared only on erroneous trials and not on correct trials. First, we looked at errors made during semantically related and semantically unrelated blocks separately in high-motivation and low-motivation conditions. This analysis revealed no significant interaction between Motivation and Semantic Relatedness (F(1, 18) < 1). Therefore, for all analyses described below, errors made during semantically related and mixed blocks in the high-motivation condition were collapsed. Errors made during semantically related and mixed blocks in the low-motivation condition were also pulled together. Similar, we collapsed errors made in semantically related and mixed blocks across high- and low-motivation conditions.

In the remaining analyses, amplitudes and latencies of the ERN were submitted to a repeated-measures General Linear Model (GLM) analysis. All ANOVAs included two factors, i.e. Electrode Site (Fz, FCz, Cz, and Pz) and either Motivation (high vs. low) or Semantic Relatedness (semantically related vs. mixed).

To investigate the effect of motivation on the ERN, repeated measures ANOVAs were run with Motivation as independent variable and amplitude of the ERN as dependent variable. This analysis demonstrated that the amplitude of the ERN was modulated by the motivation manipulation ($F(1, 18) = 5.44$, $MS_e = 26.26$, $p < .05$). The amplitude of the ERN was significantly larger in the high-motivation condition (–5.44 μV, SD = 3.50) compared to the low-motivation condition (–3.51 μV, SD = 1.94). There was no significant effect of Electrode Site ($F(3, 54) = 1.97$, $MS_e = 1.98$, n.s.), nor was there an interaction between Electrode Site and Motivation ($F < 1$).

Interestingly, similar analyses with latency of the ERN as dependent variable revealed a significant effect of Motivation ($F(1, 18) = 18.38$, $MS_e = 2986.56$, $p < .001$). The ERN peaked significantly later in the high-motivation condition (92 ms, SD = 36) compared to the low-motivation condition (54 ms, SD = 28). There was no effect of Electrode Site ($F(3, 54) = 1.49$, $MS_e = 218.73$, n.s.), nor was there an interaction between Electrode Site and Motivation ($F(3, 54) = 2.32$, $MS_e = 332.22$, n.s.).

The corresponding analysis with Semantic Relatedness as independent variable and amplitude of the ERN as dependent variable demonstrated a significant effect of Semantic Relatedness ($F(1, 18) = 8.63$, $MS_e = 68.42$, $p < .01$). The amplitude of the ERN was significantly larger in the semantically related blocks (–0.73 μV, SD = 2.75) as opposed to the mixed blocks (1.18 μV, SD = 2.86). The analysis of Electrode Site ($F < 1$) and the interaction between Electrode Site and Semantic Relatedness ($F < 1$) revealed no significant effects. Similar analyses with latency of the ERN as dependent variable showed no significant results (all Fs < 1).

Figure 3 provides an overview of the response-locked averaged ERP waveforms for errors and correct trials across semantically related and mixed blocks. There was a significant effect of Electrode Site ($F(3, 54) = 4.01$, $MS_e = 218.06$, $p < .05$). Further investigation of this effect revealed that the ERN peaked significantly later at electrode site Pz than at electrode site Fz ($F(1, 18) = 5.43$, $MS_e = 381.80$, $p < .05$). However, there was no interaction between Electrode Site and Semantic Relatedness ($F(3, 54) = 2.51$, $MS_e = 305.53$, n.s.)

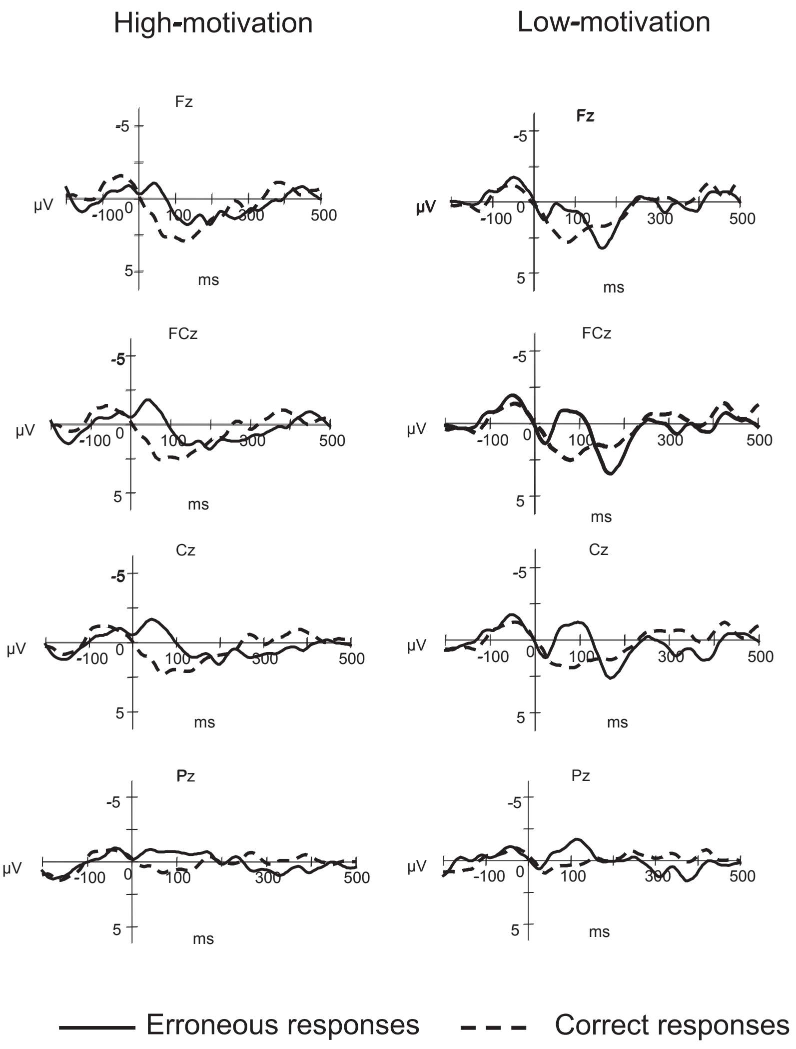

*Figure 2*. Averaged response-locked ERP waveforms for all error trials (solid lines) versus correct trials (dashed lines) across high-motivation and low-motivation conditions. Correct and incorrect trials were matched on RTs and number of trials. For graphical representation, waveforms where filtered with a high-pass filter; all analyses were done prior to the filtering.

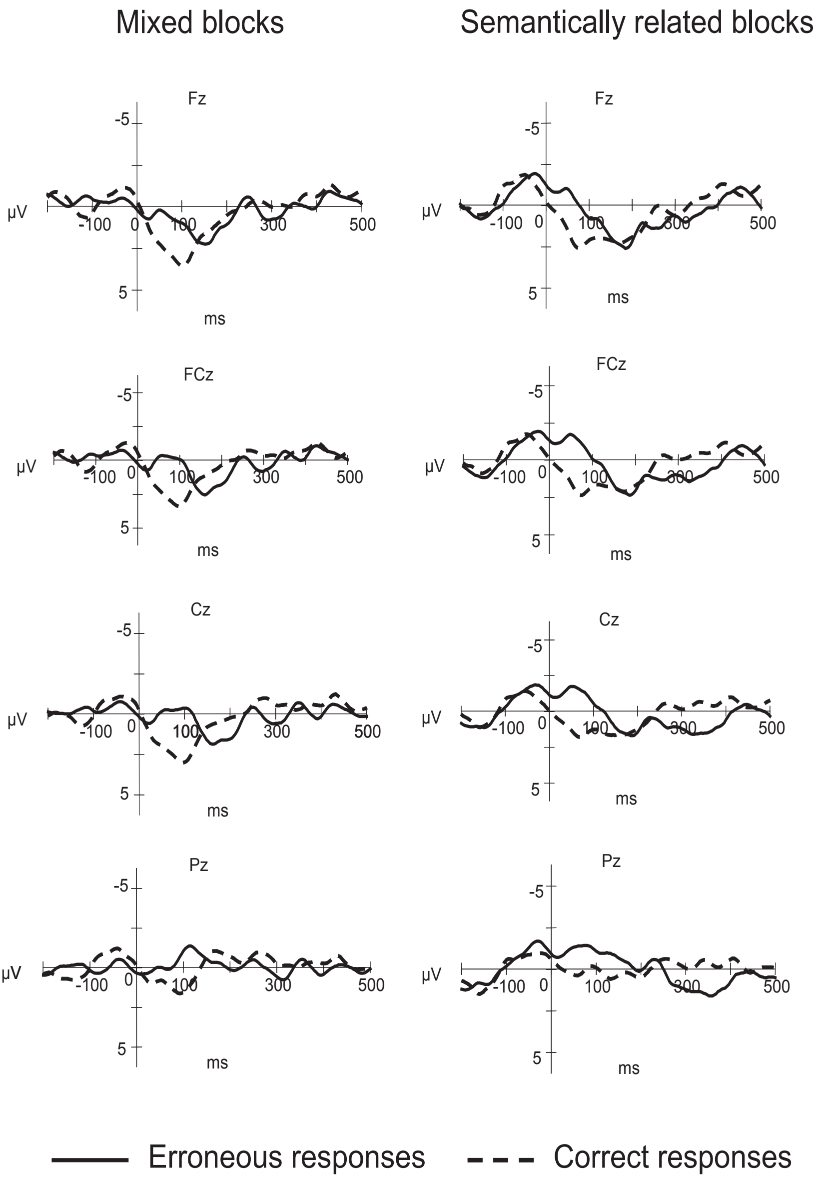

*Figure 3*. Averaged response-locked ERP waveforms for all error trials (solid lines) versus correct trials (dashed lines) across mixed and semantically related blocks. Correct and incorrect trials were matched on RTs and number of trials. For graphical representation, waveforms where filtered with a high-pass filter; all analyses were done prior to the filtering.

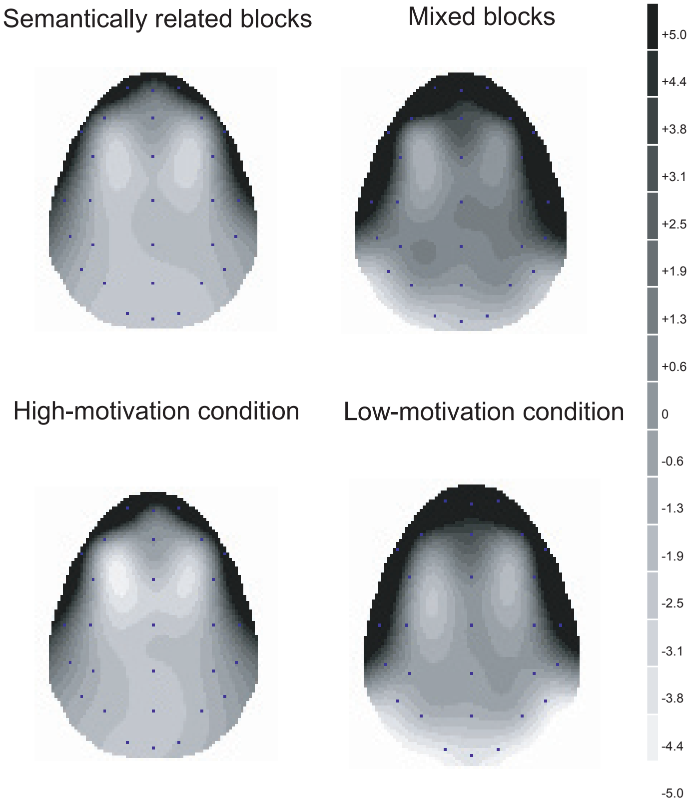

*Figure 4*. Topographic maps of the ERN amplitude between 0 and 100 ms after response onset. Negative regions depicted in light gray.

# Discussion

The goal of the present study was to investigate how verbal self-monitoring and the ERN are affected by motivation and semantic context. As expected, we obtained a typical semantic interference effect (for instance, Lupker, 1979; Schriefers, Meyer, & Levelt, 1990; etc.). Participants were slower and made more errors during picture naming in semantically related blocks as compared to mixed blocks. Our motivation manipulation had no effect on naming latencies or error rates indicating that participants did not employ different response strategies in the high- and low-motivation conditions. In the electrophysiological data, we saw a clear negative deflection on error trials compared to correct trials. Investigation of topographic maps revealed that this negative deflection has a fronto-central distribution, typical to the classic ERN (e.g., Gehring et al., 1993; Falkentein et al., 1991). However, in our statistical analysis, there was no significant effect of Electrode Site suggesting that the negative deflection found in our study was more evenly distributed across the scalp. Note, however, that in our analysis only central electrodes were included. Thus, while statistically the ERN was no different at frontal and posterior electrode sites, it is statistically unclear whether or not the ERN was largest at central versus lateral electrode sites. One possible reason for non-significant effect of Electrode Site is that we measured non-masked overt speech as opposed to button presses. It is possible that the effect we have demonstrated in the present study has a more wide-spread distribution than the effect previously shown with motor tasks. However, all other characteristics of the negative deflection in our study are in correspondence with the classic ERN. The negative deflection in our study is present only on erroneous responses and absent from correct trials, it peaks within 100 ms of the overt erroneous response, and is descriptively largest at fronto-central electrode sites. Therefore, we would like to propose here that the negative deflection found in our study is an ERN-like response. In the remaining part of the discussion, we will refer to the ERN-like response in our study as the ERN.

The electrophysiological data obtained in the present study agree with our predictions. The amplitude of the ERN was significantly larger in semantically related blocks than in mixed blocks. Further, the amplitude of the ERN was also significantly larger in high- than low-motivation conditions. Interestingly, the latency of the ERN was also affected by the motivation manipulation. The ERN peaked significantly later in the high-motivation condition than in the low-motivation condition. Neither in the behavioral nor in the electrophysiological data was there an interaction between the context in which pictures were named and the motivation manipulation. These findings will be discussed in more detail below.

In accordance with previous studies (e.g., Damian et al., 2001, Schnur at al., 2006; Vitkovitch & Humphreys, 1991), we showed that pictures were named slower and responses were more erroneous in the context of the same-category items than in the context of items from different semantic categories. This semantic interference can be accounted for by competition between co-activated lexical entries in the same semantic context. This competition, in turn, affects selection latencies (see Levelt et al., 1999 for a review; but see Finkbeiner, Gollan, & Caramazza, 2006 as well as Mahon, Costa, Peterson, Vargas, & Caramazza, 2007 for an alternative view). Co-activation of multiple lexical entries may lead to multiple potential responses active at the same time and therefore increased conflict present at the time of response. The verbal self-monitor has to verify on-line whether or not the correct entry was chosen from the set of competing candidates, presumably resulting in slower naming responses. In the mixed context, however, such verification may be faster, since unrelated words do not lead to the activation of related concepts, and therefore less competition may be present at the time of the response.

Our EEG data are in accordance with this assumption. We showed that the amplitude of the ERN increases in the semantically related context as opposed to the mixed context. In previous research, Ganushchak and Schiller (in press) demonstrated an increase in the amplitude of the

ERN in the presence of semantically related distractors to target pictures as compared to unrelated distractors. They concluded that this increase in ERN amplitude during the semantically related condition was most likely due to the simultaneous activation of competing lexical items. The present study replicates and extends this finding. Ganushchak and Schiller (in press) employed a phoneme-monitoring task, in which participants were instructed to press a button when a target phoneme was present in the name of the picture. Even though a phoneme monitoring task is verbal in nature and involves monitoring of internal speech production, the authors could not completely exclude the possibility that at least some of the errors observed in their study were motor slips (of the hand) and not verbal errors *per se*. In contrast, in the present study, we employed a more natural picture naming task in which all responses given were verbal responses, and we demonstrated an enhancement of the amplitude of the ERN in the semantically related context compared to the mixed context.

A reinforcement-learning theory cannot fully account for the increase in the amplitude of the ERN in the semantically related blocks compared to the mixed blocks. According to this theory, errors induce a phasic decrease in mesencephalic dopaminergic activity when ongoing events are determined to be worse than expected (Holroyd & Coles, 2002). However, there is no reason to suggest that the monitoring system could not make an optimal evaluation of current events and events that were predicted in the mixed blocks compared to the semantically related blocks.

Alternatively, it is possible the difference in the ERN amplitude between semantically related and mixed blocks was due to the differences in naming latencies. As stated above, participants were slower in naming pictures in the semantically related blocks than mixed blocks. However, to our knowledge, differences in the response latencies only have an effect on the ERN when accuracy is emphasized over speed (e.g., Gehring et al., 1993). This, however, is not applicable to our case, since we did not observe a speed-accuracy trade-off: Participants were not only slower but also made more errors in the semantically related blocks than in the mixed blocks. There is a whole range of studies that demonstrate an increase in amplitude of the ERN on incongruent as compared to congruent trials, despite differences in the behavioral responses on these trials. Participants are generally slower and make more errors on incongruent compared to congruent trials (e.g., Hajcak et al., 2005; Fiehler et al., 2005; Yeung et al., 2004). This increase in amplitude of the ERN is attributed to the increased amount of response conflict during incongruent trials compared to congruent trials and not due to behavioral differences. Therefore, we think that in our case the higher amplitude of the ERN during semantically related blocks relative to mixed blocks is also due to the increased response conflict. This, in turn, provides stronger evidence that the ERN is sensitive to the conflict that arises during lexical competition.

Interestingly, the motivation manipulation also yielded an effect of ERN amplitude. The high-motivation condition was associated with higher amplitudes and longer latencies of the ERN compared to the low-motivation condition. However, in our behavioral data, we found no effect of the motivation manipulation. Participants were equally fast and accurate in the high- and low-motivation condition. Due to the fact that the high- and low-motivation condition did not differ with respect to any behavioral measure, it is unlikely that the increased ERN in the high-motivation condition could be driven by performance-related differences. This is supported by Gehring and colleagues (1993) who showed in their original paper that the ERN was increased when accuracy was emphasized over speed. However, no such trade-off was observed in our behavioral data.

It is likely that in the high-motivation condition there was more conflict present than in the low-motivation condition, which consequently led to an enhanced ERN. According to the conflict monitoring theory, this could be explained by assuming that participants striving for accurate responses tended to focus more effectively on the task at hand in the high-motivation than in the low-motivation condition and therefore an increased tendency to correct errors in the former condition. This increased error-correcting activity, in turn, may have led to increased conflict with the error just produced, resulting in an increased ERN (Yeung, Botvinick, & Cohen, 2004).

In terms of the reinforcement learning theory (Holroyd & Coles, 2002), participants could have had a higher expectation of a good outcome in the high-motivation condition than in the low-motivation condition, since making errors in the high-motivation condition was associated with financial punishment. Therefore, errors in the high-motivation condition lead to larger violations of the prediction than errors in the low-motivation condition, and thus a more pronounced ERN was generated.

Alternatively, in the high-motivation condition, errors had a higher significance for participants than in the low-motivation condition. It is plausible that the ERN is sensitive to the motivational significance of errors. For instance, Hajcak and colleagues (2005) used an arrowhead version of a flanker task and varied monetary value on trial-by-trial basis. Interestingly, they found that errors committed in a condition with higher monetary value gave rise to a higher ERN. Note that Hajcak and colleagues demonstrated the effect of monetary value on the ERN despite it having no systematic effect on participants' behavioral performance. The authors concluded that the ERN reflects motivation significance of errors. Our own results are in agreement with their findings. It is possible that in circumstances when errors are less acceptable and potentially have more severe consequences, e.g. when giving a public speech, the verbal monitoring system has to be more alert in order to verify whether or not the selected response was correct.

Besides the increased amplitude of the ERN in the high-motivation condition, we also showed prolonged ERN latency, compared to the low-motivation condition. We expected to find no difference in ERN latency, since it seemed to be invariant with respect to the erroneous response and largely independent of experimental manipulation (Falkenstein et al., 2000; Scheffers & Coles, 2000). Recently, however, more studies showed a modulation of ERN latency depending on the experimental manipulations (e.g., Fiehler, Ullsperger, & Von Cramon, 2005; Johannes et al., 2001; Krigolson & Holroyd, 2007). It is possible that the latency of the ERN reflects prolonged error-monitoring mechanisms (Johannes et al., 2001). In the high-motivation condition, errors had more severe consequences for participants than in the low-motivation condition. Therefore, it is plausible to assume that in the high-motivation condition error monitoring was slowed down in order to verify whether or not the selected response was indeed the correct response. In the low-motivation condition, such verification could be carried out faster and was less important, since errors did not have any consequences.

In semantically related and mixed blocks, there was no difference in ERN latency. According to the logic applied above, this lack of difference in ERN latency is not surprising. Errors made in semantically related blocks did not have more consequences for participants than errors made in mixed blocks. Alternatively, shift in ERN latency has been attributed to the corrective processes (e.g., Fiehler et al., 2005; Falkenstein et al., 1996). It has been proposed that slow error corrections seem to be based on a delayed correct response tendency resulting in a later peak of the ERN (Fiehler et al., 2005). However, this finding is not robust and there are studies that failed to show a relationship between ERN latency and error correction (Falkenstein et al., 1994; Rodríguez-Fornells et al., 2002). It is difficult to say whether or not our data can be interpreted with the help of this account. Possibly, participants internally corrected more errors during the high-motivation condition than the low-motivation condition and therefore showed a delayed ERN latency. However, in the behavioral data there was no significant difference in error rate between the high- and low-motivation conditions. If participants corrected more errors in the high-motivation condition than in the low-motivation condition, one would expect to see fewer errors in the former compared to the latter condition.

There is a seeming contradiction in our results. On the one hand, we showed that the amplitude of the ERN increased in the semantically related context as opposed to the mixed context, indicating that the ERN is sensitive to the presence of conflict. On the other hand, we demonstrated that the ERN was higher in the high-motivation condition than low-motivation condition. In the high-motivation condition, errors had a higher significance than in the low-motivation condition, indicating that the ERN is sensitive to the motivational manipulation. It is unlikely that errors in the semantically related context had higher significance than in the mixed context since the financial reward was independent of semantic context (see above). Furthermore, after the experimental session participants reported that they attempted to name pictures as accurately as possible independently of the context in which pictures were presented. To our knowledge, conflict and motivational accounts of the ERN are two mutually exclusive hypotheses in the existing literature, and it has not been shown so far that the ERN can be affected by both factors. We would like to propose here that possibly the conflict and motivational theories are closer related than previously thought. The detection of conflict or errors is likely to have direct affective consequences (Yeung, 2004). Ullsperger and Von Cramon (2004) also suggest that there might be a close interplay with emotional and motivational functions and performance monitoring. Therefore, a clear-cut distinction between theories that associate the ERN with a process of error/conflict detection and theories that associate it with a process giving rise to affective/motivational changes related to error or conflict detection may not be possible, since both may refer to one and the same process (Yeung, 2004).

One potential methodological problem of the current study is the vocalization-related cortical potential (VRCP). The VRCP consists of a movement-related potential preceding vocalization and an auditory evoked potential which follows vocalization (Gunji, Hoshiyama, & Kakigi, 2000). The auditory evoked potential peaks around 100 ms after vocalization and therefore has a similar time course as the ERN. However, unlike the ERN, an auditory evoked potential is independent of response correctness. Masaki et al. (2001) used loud pink noise to suppress the vocalization-elicited components. In the current study, we did not use any masking procedure. Participants perceived their own voice as feedback to monitor their own speech. Removing such feedback, by masking participants' voices with pink noise, might disrupt and interfere with the normal working of the monitor. For instance, Christoffels, Formisano, and Schiller (2007) asked participants to name pictures when participants could hear their own voices and when they could not due to the presentation of masking pink noise during their responses. Christoffels and colleagues demonstrated that the masking of feedback was associated with a reduction of activity in areas found in overt speech production in comparison to the normal feedback condition. Therefore, in order to keep our task as natural as possible, we choose not to administer white or pink noise to our participants. We argue that, since auditory evoked potentials are independent of response correctness and are present during correct as well as erroneous trials, the changes between correct and erroneous responses found in the present study are due to the changes in the ERN and not auditory evoked potentials.

Finally, we would like to comment on the issue of individual differences since motivation is somewhat dependent on personality or mood characteristics of individual participants. For instance, Boksem et al. (2006) reported that individuals who score high on a measure of punishment sensitivity have larger ERN amplitudes than individuals who score low on such a measure. Pailing & Segalowitz (2004) demonstrated that individuals who score high on conscientiousness have smaller changes in ERN amplitudes with manipulations of motivation compared to individuals who score low on conscientiousness. However, our study had a within-subject design and participants were not pre-selected on their personality or mood characteristics but were randomly selected. Therefore, it is rather unlikely that our results could be accounted for by individual differences between participants.

To conclude, we argue that due to the simultaneous activation of competing items, the verbal self-monitor presumably needs to be more alert in the semantically related context than in the mixed context in order to validate whether a given response was correct or erroneous. Additionally, in circumstances when errors have more severe consequences for the speaker, e.g. when giving a public speech or during an interview, the monitor needs to work harder to prevent errors and to correct errors already made. Further, we showed that the ERN is sensitive to the presence of lexical conflict. Similarly, in previous studies (i.e., Ganushchak & Schiller, in press) we reported an enhanced ERN after errors of verbal monitoring in the presence of semantically related distractors as opposed to semantically unrelated ones. Möller and colleagues (2007) – employing the inverse source localization method – identified a medial frontal generator in the supplementary motor area (SMA) as the main source of the negativity preceding erroneous vocalizations. A typical ERN has been located within the ACC/SMA regions (e.g., Dehaene, Posner, & Tucker, 1994). These findings provide converging evidence that the ERN could be used as an electrophysiological marker of error processing in language research. However, a note of caution may be in place here. To our knowledge, we are the first to use unmasked overt speech to investigate the ERN. As stated above, we did not find an unambiguous localization of the ERN in fronto-central electrodes. It is possible that the distribution of the ERN during overt speech is somewhat different than the classic ERN. The topography of the ERN during overt speech production clearly deserves further experimental investigation.

# 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. and Gulikers, L. (1995). *The CELEX lexical database* (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *The Quarterly Journal of Experimental Psychology*, 58A, 667-692.

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.

Boksem, M. A. S., Tops, M., Wester, A. E., Meijman, T. F., Lorist, M. M. (2006). Error-related ERP components and individual differences in punishment and reward sensitivity. *Brain Research*, 1101, 92-101.

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.

Christoffels, I. K., Formisano, E., & Schiller, N. O. (2007). Neural correlates of verbal feedback processing: an fMRI study employing overt speech. *Human Brain Mapping*, 28, 868-879.

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

Damian, M. F., Vogliocco, G., & Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition,* 81, B77-B86.

Falkenstein, M., Hohnsbein, J., & Hoorman, J. (1994). Event-related potential correlates of errors in reaction tasks. In G. Karmos, M. Molnar, V. Csepe, I. Czigler, & J. E. Desmedt (Eds.), *Perspectives of event-related potentials research* (pp. 287-296). Amsterdam: Elsevier Science.

Falkenstein, M., Hohnsbein, J., & Hoorman, J. (1996). Differential processing of motor errors. In C. Ogura, Y. Koga, & M. Shimokochi (Eds.), *Recent advances in event-related brain potential research* (pp. 192-195). Tilburg: Tilburg University Press.

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.

Fiehler, K., Ullsperger, M., & Von Cramon, D. Y. (2005). Electrophysiological correlates of error correction. *Psychophysiology*, 42, 72-82.

Finkbeiner, M., Gollan, T. H., & Caramazza, A. (2006). Lexical access in bilingual speakers: What's the (hard) problem? *Bilingualism: Language and Cognition*, 9, 153-166.

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). Effects of auditory distractors on verbal self-monitoring. *Journal of Cognitive Neuroscience*, 20, 1-14.

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.

Gunji, A., Hoshiyama, M., & Kakigi, R. (2000). Identification of auditory evoked potential of one's own voice. *Clinical Neurophysiology*, 111, 214-219.

Hajcak, G., Moser, J. S., Yeung, N., & Simons, R. F. (2005). On the ERN and the significance of errors. *Psychophysiology*, 42, 151-160.

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.

Johannes, S., Wieringa, B. M., Nager, W., Rada, D., Dengler, R., Emrich, H. M., Münte, T. F., & Dietrich, D. E. (2001). Discrepant target detection and action monitoring in obsessive-compulsive disorder. *Psychiatry Research: Neuroimaging Science*, 108, 101-110.

Krigolson, O. E., & Holroyd, C. B. (2007). Predictive information and error processing: the role of medial-frontal cortex during motor control. *Psychophysiology*, 44, 586-595.

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*. MIT Press, Cambridge, MA.

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

Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory & Cognition, 7*, 485-495.

Luu, P., Collins, P., & Tucker, D. (2000). Mood, personality, and self-monitoring: negative affect and emotionality in relation to frontal lobe mechanisms of error-monitoring. *Journal of Experimental Psychology: General*, 129, 43-60.

Maess, B., Friederici, A. D., Damian, M., Meyer, A. S., & Levelt, W. J. M. (2002). Semantic category interference in overt picture naming: sharpening current density localization by PCA. *Journal of Cognitive Neuroscience,* 14, 455-462.

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical Selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 33, 503-535.

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., Rodriguez-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.

Pailing, P. E., & Segalowitz, S. J. (2004). The error-related negativity as a state and trait measure: motivation, personality, and the ERPs in response to errors. *Psychophysiology*, 41, 84-95.

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.

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.

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.

Scheffers, M. K., & Coles, M. G. H. (2000). Performance monitoring in a confusing world: error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology, 26*, 141-151.

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.

Schnur, T. T., Schwartz, M. F., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language*, 54, 199-227.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: picture-word interference studies. *Journal of Memory and Language, 26*, 86-102.

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.

Vitkovitch, M., & Humphreys, G. W. (1991). Perseverant responding in speeded naming of pictures: it's in the links. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 17, 664-680.

Ullsperger, M., & Von Cramon, D. Y. (2004). Neuroimaging of performance monitoring: error detection and beyond. *Cortex*, 40, 593-604.

Yeung, N. (2004). Relating cognitive and affective theories of the error-related negativity. In M. Ullsperger & M. Falkenstein (Eds.), *Errors, conflicts, and the brain. Current opinions on performance monitoring* (pp. 63-70). Leipzig, Germany: Max Planck Institute of Cognitive Neuroscience.

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.

# Appendix

A list of the 15 semantic categories and 75 targets used in the current experiment. The approximate English translation is given in brackets.

FURNITURE: bank (couch), tafel (table), stoel (chair), bureau (desk), kruk (stool)
ANIMALS_1: giraffe (giraffe), kameel (camel), hert (deer), olifant (elephant), zebra (zebra)
ANIMALS_2: muis (mouse), slang (snake), vis (fish), eend (duck), libel (dragonfly)
BODYPARTS: been (leg), hand (hand), vinger (finger), voet (foot), arm (arm)
MUSICAL INSTRUMENTS: gitaar (guitar), fluit (flute), trompet (trumpet), harp (harp), piano (paino)
FRUITS: appel (apple), banana (banana), kers (cherry), peer (peer), citroen (lemon)
BIRDS: uil (owl), duif (pigeon), zwaan (swan), ooievaar (stork), pinguin (penguin)
HEADWARE: pet (pet), muts (bonnet), kroon (crown), helm (helm), hoed (hat)
TOOLS: hamer (hammer), zaag (saw), tang (tongs), vijl (file), boor (drill)
VEGETABLES: prei (leek), sla (lettuce), wortel (carrot), tomaat (tomato), ui (onion)
CLOTHING: broek (trousers), rok (skirt), trui (sweater), hemd (shirt), jurk (dress)
FOOD: brood (bread), donut (donut), ei (egg), worst (sausage), kaas (cheese)
VEHICLES: trein (train), auto (car), fiets (bicycle), schip (ship), bus (bus)
KITCHEN UTENSILS: pan (pan), vergiet (colander), rasp (grater), ketel (kettle), wok (wok)
UTENSILS: glas (glass), kop (cup), fles (bottle), schaal (dish), bord (plate)

# Chapter 6

# When chair acquires gender: An ERP study on gender transfer from L1 to L2[10]

## Abstract

This study addressed how Error-Related Negativity (ERN) is affected by conflict in bilingual context. Dutch-English bilinguals saw Dutch words in white print that needed to be classified (right or left button-press) according to their grammatical gender and colored words that were to be classified on the basis of their color. Colored words included Dutch common and neuter gender words, and English translations of those words. Performance was more erroneous on incongruent trials, in which there was a mismatch between color and gender response mapping, than on congruent trials, in which no such discrepancy was present. We obtained an ERN following incorrect classifications for colored words which was larger for incongruent than congruent trials. Higher error rates and enhanced amplitude of the ERN on incongruent trials were independent of the language in which target words were presented. This may suggest that when multiple languages are active, the verbal monitor has more difficulty to keep languages separated and therefore suffers more from intrusions from a second language, resulting in more response conflict and more error-prone performance. These results also provide evidence that under certain circumstances people can transfer some grammatical characteristics of their first language (e.g. gender) to their second language, even if such characteristics are absent from the latter.

---

[10] This chapter is based on Ganushchak, L. Y. & Schiller, N. O. (submitted). When chair acquires gender: an ERP study on gender transfer from L1 to L2. .

# Introduction

Errors are part of everyday life and are often the basis of learning and developing new strategies. Therefore, error processing forms a major part of research on human performance monitoring. 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 conscious *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*, which states that the ERN reflects detection of response conflict and not detection of errors *per se* (Botvinick et al., 2001). Response conflict arises when multiple responses compete for selection. Presence of conflicting responses reflects situations in which 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. In the present study, we are interested in a different kind of monitoring, namely *verbal self-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 (see Postma, 2000 as well as Hartsuiker & Kolk, 2001 for overviews). Verbal monitoring is achieved via the speech comprehension system.

Previous studies showed that an ERN can be elicited by verbal errors (e.g., Ganushchak & Schiller, 2006, in press; Masaki, Tanaka, Takasawa, & Yamazaki, 2001; Möller, Jansma, Rodríguez-Fornells, & Münte, 2007; Sebastián-Gallés, Rodríguez-Fornells, Diego-Balaguer, & Díaz, 2006). In the present study, we investigated the relationship between the ERN and verbal-monitoring in a non-native language. Nowadays, bilingualism is the rule rather than an exception (Costa & Santesteban, 2006), 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. The goal of the present study is to investigate whether or not the ERN is influenced by conflict in bilingual context. There are indications that the ERN is sensitive to lexical conflict in monolingual tasks. For instance, Ganushchak and Schiller (in press) used a phoneme-monitoring task to investigate how the ERN is affected by auditory distractors during verbal monitoring. The authors found that the ERN was largest following errors that occurred after semantically related distractors had been presented, as compared to semantically unrelated ones. This result demonstrates that the ERN is not only sensitive to response conflict resulting from the incompatibility of motor responses but also to more abstract lexical retrieval conflict resulting from activation of multiple lexical entries.

There is some evidence that the ERN can also be observed in bilingual context. For instance, Sebastián-Gallés and colleagues (2006) assessed Spanish-dominant and Catalan-dominant bilinguals using an auditory lexical decision task in Catalan. The authors showed that Spanish-dominant bilinguals had great difficulty in rejecting experimental non-words (e.g. deciding "no" to /gəλedə/; correct /gəλɛdə/, phonological transcription of the Catalan word "galleda", meaning *bucket*), and did not show an ERN in their erroneous non-word decisions either (i.e. false alarms, e.g. deciding "yes" to the non-word /gəλedə/). According to Sebastián-Gallés et al. this suggests that Spanish-dominant bilinguals activated the same lexical entry from experimental words and non-words (in the experimental stimuli, the vowel change involved a Catalan-specific /e – ɛ/ contrast) and therefore showed no difference between correct and erroneous responses. In contrast, Catalan-dominant bilinguals showed a clear ERN.

In the present study, we investigated whether or not the ERN is affected by conflict resulting from simultaneous activation of two languages: Dutch and English. To address this question, we used a modified version of the Extrinsic Affective Simon Task (EAST; De Houwer, 2003). The EAST was developed as an indirect measure of attitudes. For instance, the EAST was used to asses affective evaluation of phobia-relevant stimuli (Huijding & De Jong, 2006), the valence of normatively positive and negative nouns (De Houwer, 2003), the affective evaluation of unipolar concepts (e.g., *flower*; De Houwer, 2003), and implicit alcohol-related cognition in heavy drinkers (De Houwer, Crombez, Koster, & De Beul, 2004; De Jong, Wiers, Van de Braak, & Huijding, 2007). The EAST relies on the principle that it is easier to give a response that is associated with positive valence to positive items than to negative items and easier to give a response that is associated with negative valence to negative items than to positive items, even when the valence of the items is irrelevant (De Houwer et al., 2004).

In a typical EAST task, participants are presented with white, blue, and green words. Participants are instructed that in a case of the white words, the valance/meaning of the word is important. In a case of the colored words, only color is important and valance/meaning of the words is irrelevant. For example, participants might be asked to press a right key for positive and green words and press a left key for negative and blue words (De Houwer, 2003; De Houwer et al., 2004). By assigning one response to positive white words and the other response to negative white words, responses become associated with positive or negative valence (De Houwer, 2003). Crucial trials are trials on which colored words are presented. The typical finding is that participants are faster and more accurate in responding when the valence of the response matches with the valence of the colored words. In the example above, positive and green words are assigned to a right key and negative and blue words are assigned to a left key. Thus, it is easier to respond to positive green words than negative green words because of the match between the valance of the stimulus and the response in the former, but not in the latter case. Similarly, for the same reason, it is easier to respond to negative blue words than positive blue words (De Houwer, 2003; De Houwer et al., 2004).

In the present study, we proposed to use a modified version of the EAST in language research. In the literature, there is some evidence that the non-response language is nonetheless activated in bilinguals (e.g., Colomé, 2001; Costa et al., 2000; Kroll & Dijkstra, 2002; Rodríguez-Fornells et al., 2005) and that native and foreign languages are based on the similar neural substrate (e.g., Klein et al., 1995, 1999; Perani et al., 1998). Therefore, it is plausible to assume that bilingual speakers attribute some characteristics of their native language to their second language, even if that characteristic is absent in the second language. Specifically, we explored whether native Dutch speakers attribute grammatical gender to their second language, i.e. English.

The Dutch language has a grammatical gender system that differentiates between two genders: common and neuter. In English, however, no such grammatical gender system exists. In the present study, we presented Dutch common and neuter gender words in white, and participants were instructed to press one key for common nouns and another key for neuter nouns. The purpose of these white words was twofold: On the one hand, they induced the association between a response key and a particular grammatical gender. On the other hand, white words assured that participants were in a gender monitoring mode, i.e. participants should be inclined to determine the grammatical gender of words even if it was not relevant for the task. On colored trials, we presented Dutch common and neuter gender words as well as the English translations of these words. Each word was presented both in green and blue. The instruction for colored words was to categorize them based on color and not their gender. For instance, participants were asked to press a right key for common gender and green words and a left key for neuter gender and blue words. We expected to find that participants were more accurate in responding to green common gender words than green neuter gender words because of the match between grammatical gender of response and stimulus. Likewise, responses should be faster and more accurate for blue neuter gender words than blue common gender words. Critically, we hypothesized to see this pattern in results for both Dutch *and* English words.

More specifically, we predict superior performance for green English words that are translations of Dutch common gender words than green English words that are translations of Dutch neuter gender words. Similarly, performance for blue English words that are translations of Dutch neuter gender words should be better than performance for blue English words that are translations of Dutch common gender words. If we find this pattern of results, this would indicate that participants transferred grammatical gender features from one language (i.e. Dutch) to another (i.e. English). Further, we predict that the amplitude of the ERN will be larger for incongruent trials, in which response mapping for gender does not correspond to a response mapping for color, than to congruent trials, in which no such discrepancy between response mapping for gender and color exists.

# Methods

## *Participants*

Twenty-two undergraduate students of Maastricht University participated in the experiment. Participants received a financial reward for their participation in the experiment and gave written informed consent prior to participating in the study. Due to technical problems, the data of one participant were lost. All participants were right-handed native Dutch speakers with good knowledge of English. Most teaching materials at the university are in English and some classes at the undergraduate level are also given in English.

Participants' level of proficiency was assessed with a vocabulary test based on an English non-speeded lexical decision task that was originally developed by Meara (1996; 2005). Participants were required to indicate whether or not they knew the meaning of an English letter string. The test consisted of words selected from five different word frequency bins. The test also contained non-words, which were used to check how reliable the claims by the participants were. The score of the test ranges from 0 to 5,000 and is corrected for misattribution of non-words. The higher the score, the better is knowledge of English. This test discriminates reliably between native and intermediate level speakers. The corrected mean score in the present study was 3930, which means that participants correctly recognized words and correctly rejected non-words in 80% of the trials.

## *Materials*

Eighty Dutch (40 common and 40 neuter gender) words and eighty English words (translations of the Dutch words) were presented on the colored trials, whereas 80 common and 80 neuter gender Dutch words were presented on the white trials (see Appendix). Additionally, 40

Dutch (20 common and 20 neuter gender) words were presented in a practice task as white words. Another 40 Dutch (20 common and 20 neuter gender) with their corresponding English translation were used for a second practice task as colored words. None of the words used in either of the practice tasks occurred in the experimental task. Dutch and English words were matched on word length and frequency (see Table 1). All words had a moderate frequency of occurrence between 10 and 200 per million according to the CELEX database (CEnter for LEXical information, Nijmegen; Baayen, Piepenbrock, & Gulikers, 1995).

Table 1. Lexico-statistical characteristics of the target words.

| | | Example | Mean CELEX frequency (per one million words) | Mean length in syllables |
|---|---|---|---|---|
| Dutch words | Common | poes | 65.6 | 1.2 |
| | Neuter | blad | 69.5 | 1.3 |
| English words | “Common” | cat | 65.1 | 1.3 |
| | “Neuter” | leaf | 70.0 | 1.3 |

*Note*. English translations of the Dutch words were used in the experiment for English words.

The colored words were presented either in green or blue. The green color was created by setting the red, green, and blue values to 0, 200, and 150, respectively. The red, green, and blue values were set to 0, 150, and 200 in order to create the blue color. As a result, the blue and green colors were very similar, which made the discrimination of color more difficult, thus assuring that participants would have enough time to process the meaning of the words (De Houwer, 2003). The red, green, and blue values were set to 0, 0, and 0 for white color. All words were presented on a black background.

### *Design and Procedure*

Participants were tested individually while seated in a sound-proof room. The experiment started with two practice blocks and one experimental block. During the first practice block, white words were presented in a random order. Participants were instructed to press the left control key for common gender Dutch words and the right control key for neuter gender Dutch words. In the second practice block, participants were presented with colored words (Dutch and English words), once in green and once in blue. Half of the participants were instructed to press the left control key for green words and the right control key for blue words. The other half of participants received the reverse color-response assignment (i.e. green was coupled with a right key and blue with a left key). In the experimental block, both white and colored words were presented in a random order. White words included only the Dutch common and Dutch neuter gender words. The colored words were Dutch common and neuter gender words and English translations of those words. Note that none of the colored words were presented as white words. The words in the experimental block were not previously presented in the practice tasks. Participants were informed that if the color of the word was white, then classification should be made based on the grammatical gender of the word. However, if words were colored, then classification should be made based on the color of the word. All practice and experimental blocks consisted of the following sequence of events: each trial started with a fixation point in the center of the screen for a duration varying between 500 and 800 ms, followed by a blank screen for 500 ms, and finally followed by a word which remained in view until a response was given or the maximum response time of 2,500 ms was reached (the maximum time was based on the pilot study). Upon completion of the task, participants were asked to perform an English proficiency test. Additionally, participants were asked to translate a list of English words, used in the experiment, into Dutch including the corresponding definite article. This was done in order to verify that participants were familiar with the presented English words and that translations were consistent with translations used in the experimental block.

### *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 and band-pass filtered 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 (from –300 ms to +800 ms) were obtained including a 100 ms pre-response baseline. The EEG signal was corrected for vertical electrooculogram (EOG) artifacts, using the ocular reduction method described in Anderer, Satety, Kinsperger, and Semlitsch (1987). For the ERN, averaging was done across error trials. The amplitude of the ERN was derived from each individual's response-locked average waveforms after filtering with a band pass, zero phase shift filter (frequency range was 1-12 Hz). The mean amplitude analysis was performed for the time window from 0 ms to 100 ms after response onset. This time window was derived based on the visual inspection of the grand average waveforms and previous studies employing the ERN (e.g., Rodríguez-Fornells, Kurzbuch, & Münte, 2002). The amplitude of the ERN was recorded for each condition at electrode sites Fz, FCz, and Cz.

All analyses were performed on colored trials, i.e. words presented in green and blue. Mean reaction times (RTs), error rates, and amplitudes of the ERN from each participant were submitted to a repeated-measures General Linear Model (GLM) analysis. The analysis always involved planned comparisons with Language (Dutch vs. English) and Conflict (congruent vs. incongruent) as independent variables. On incongruent trials, there was a discrepancy between grammatical gender of response and stimulus. For instance, participants were asked to press a right key for common gender and green words and a left key for neuter gender and blue words. Thus, neuter gender green and common gender blue words might lead to increased response conflict because of the mismatch between grammatical gender and color. On congruent trials, however, no such mismatch was present. Based on the previous example, for common gender green and neuter gender blue words there is a match between grammatical gender of response and color of stimulus (see Figure 1; see back side of the oover).

# Results

## *Behavioral data*

Trials on which English words were presented which participants did not know or in case they misattributed grammatical gender of a word were removed from the analysis. Latencies shorter than 300 ms and longer than 1,500 ms were excluded from the analysis as well. A 2 (language) by 2 (conflict) ANOVA revealed no significant effect of Language (F(1, 20) = 1.08, $MS_e$ = 4,077.85, n.s.) or Conflict (F < 1), and no interaction between these two factors (F < 1; see Table 2 for an overview of the behavioral data).

Table 2. Overview of the behavioral data. Mean (± standard deviation) reaction times (in ms) and error rates (% relative to the number of trials per condition) as a function of conflict and language.

| | *Dutch incongruent* | *Dutch congruent* | *English incongruent* | *English congruent* |
|---|---|---|---|---|
| Reaction times | 670 (92) | 664 (104) | 649 (60) | 656 (57) |
| Error rates | 5.2 (4) | 2.0 (1) | 3.9 (1) | 2.7 (1) |

A similar analysis was performed with proportion of errors as dependent variable. Participants made on average 3.4% errors. A 2 (language) by 2 (conflict) ANOVA revealed a significant effect of Conflict (F(1, 20) = 17.99, $MS_e$ = 3.62, p < .01), but no effect of Language (F < 1) and no interaction between these two factors (F(1, 20) = 4.03, $MS_e$ = 3.84, n.s.). Participants made more errors on incongruent trials than on congruent trials (see Table 2).

### *ERN*

The ERN was revealed in response-locked ERP averages for incorrect trials. No negative deflection was observed in the ERP waveforms for correct trials during visual inspection of the EEG waves. Figure 2 provides an overview of the response-locked averaged ERP waveforms for error and correct trials across conditions (Dutch incongruent, Dutch congruent, English incongruent, and English congruent). Figure 3 displays the topographical representation of the ERN.

An ANOVA was run with Conflict and Language as independent variables and the amplitude of the ERN as dependent variable. This analysis revealed a significant effect of Conflict (F(1, 20) = 4.71, $MS_e$ = 8.22, p < .05). The amplitude of the ERN was larger for incongruent trials (2.79µV, SD = 1.73) than congruent trials (2.11µV, SD = 1.76; see Figure 4). There was neither an effect of Language (F(1, 20) = 1.35, $MS_e$ = 6.81, n.s.) nor an interaction between Language and Conflict (F < 1).

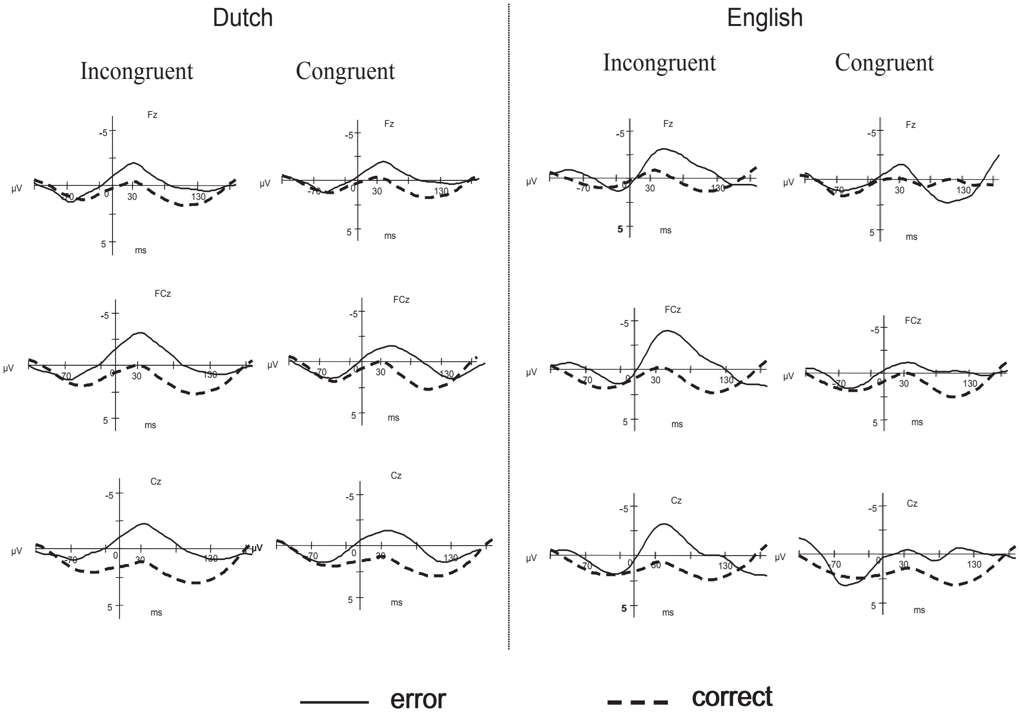

*Figure 2*. Averaged response-locked ERP waveforms for all error trials (solid lines) versus correct trials (dashed lines) across conditions (Dutch incongruent, Dutch congruent, English incongruent, and English congruent). Correct and incorrect trials were matched on RTs and number of trials.

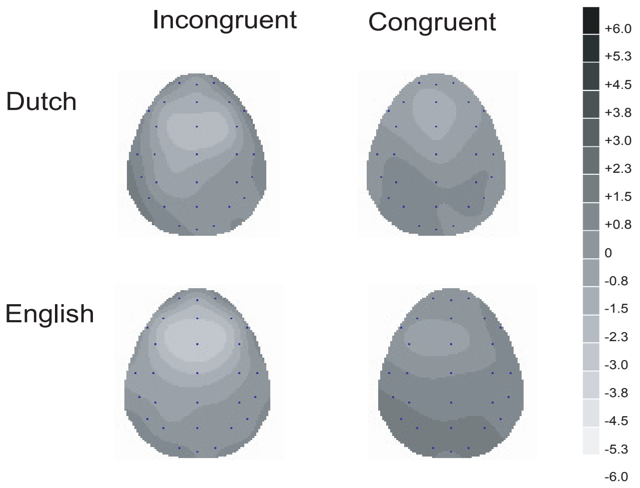

*Figure 3*. Topographic maps of the ERN amplitude between 0 and 100 ms after response onset. Negative regions depicted in light gray.

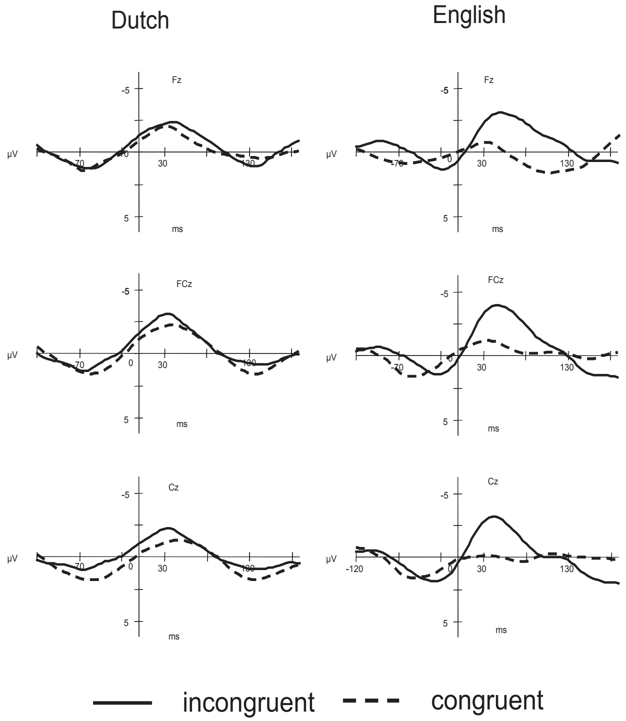

*Figure 4*. Averaged response-locked ERPs for incorrect trials. Dotted lines depict the congruent trials and solid lines depict incongruent trials for Dutch and English.

# Discussion

The goal of the present study was to investigate whether and how the ERN is affected by conflict in bilingual context. We found no difference in RTs between incongruent and congruent trials. This is not unusual since the effects of the task employed here are most prominent in the error data (e.g., De Houwer, 2003; Huijding & De Jong, 2005). As expected, we found enhancement of the error rate and the amplitude of the ERN on incongruent trials compared to congruent trials. This effect was independent of whether target words were presented in Dutch or in English. Note that participants were instructed to make a color classification on target words; e.g., they had to press left for green words and right for blue words. Therefore, based on the purely erroneous color decision, there should not be any difference between incongruent and congruent trials, since both types of trials include an erroneous decision for both green and blue colors. What makes a difference between incongruent and congruent trials is the discrepancy between response mappings for grammatical gender, created by responses to white words, and color. Thus, on incongruent trials there was a mismatch between color and gender response mapping, and on congruent trials no such discrepancy was present. Presumably, even when performing a color judgment, participants not only processed the presented word, but also retrieved its grammatical gender. Critically, participants also transferred Dutch grammatical gender to English where grammatical gender is absent. Note, that all participants had good knowledge of English and were fully aware of the fact that English has no grammatical gender.

Our results might be accounted by the fact that English words activated their Dutch equivalents, and those in turn activated their grammatical gender. We cannot rule out the possibility that Dutch equivalents of English words were activated when English words were presented. However, if one is to assume this, then it is likely that participants would react somewhat slower for English words, since they would have to process their activated Dutch translations. We found no significant difference between button presses for Dutch and English words. Additionally, participants in our study were overall fast, responding on average in 660 ms. In the existing literature on word translation, reported response times lie between 900 ms and 1,000 ms (e.g., De Groot & Poot, 1997; Kroll et al., 2002; La Heij et al., 1996).

Interestingly, in a recent study, Midgley and colleagues (2007) found that at an early stage of English language acquisition, L1 French speakers showed sensitivity to L1 gender during L2 sentence processing. In French, possessive determiners are gender-marked and must agree with the gender of the noun they modify and not with the gender of the referent, as is the case in English. Midgley and colleagues asked French speakers in their second year of English studies to read grammatically correct English sentences for comprehension. Critical nouns were preceded by possessive determiners that either agreed or disagreed with the gender of the translation equivalents in French (e.g., “Barbara saw her shoe under the bed” [congruent condition] and “Peter saw his shoes under the bed” [incongruent condition, since in French *shoe* is feminine]).

Midgley and colleagues found a difference in EEG averages between congruent and incongruent conditions, which suggests that grammatical gender associated with nouns from L1 can influence how nouns are processed in L2.

We argue that a similar transfer of grammatical gender from L1 to L2 occurred in the present study as well. Assuming that participants implicitly attributed grammatical gender to English words, it is plausible that at the time of response there was competition between an inappropriate response (e.g., response to grammatical gender for both languages) and a correct response (e.g., response to color). It is possible that under circumstances when multiple languages are active, the verbal monitor has more difficulty to keep languages separated and therefore suffers more from intrusions from a second language, resulting in less accurate performance. Activation of both English and Dutch words could have resulted in more response conflict and thus higher amplitudes of the ERN (e.g., Botvinick et al., 2001; Yeung et al., 2004).

Alternatively, it is possible that response conflict in this study did not arise from co-activation of Dutch and English and transfer of L1 grammatical gender to L2, but was due to the conflict in response mapping. In the task, white and colored words were presented in a random order and not in a blocked fashion. Thus, participants had to constantly switch between grammatical gender and color discrimination. It is possible that, at least on some error trials, participants confused instructions that led to a conflict between intended (i.e. color decision) and actual response (i.e. grammatical gender decision). However, if participants confused instructions and for some of the colored words executed the gender discrimination task instead of color discrimination, then this fails to explain why we found an effect for English words as well, since for English words gender discrimination is impossible, due to the absence of grammatical gender in nouns. Hence, we believe that the conflict in the present task resulted from co-activation of Dutch and English, and from transferring L1 grammatical gender characteristics to L2.

In summary, we showed that the ERN is sensitive to the linguistic context. Participants made more errors and showed an enhanced ERN on incongruent trials compared to congruent trials. 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. However, more research is needed to make a clear-cut separation between possible explanations for the effect of grammatical gender obtained in English. We believe that the effect found in the current study is a result of gender transfer from one language to another. However, other accounts cannot be completely excluded.

# 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). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

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.

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.

Costa, A., & Santesteban, M. (2006). The control of speech production by bilingual speakers: introductory remarks. *Bilingualism: Language and Cognition*, 9, 115-117.

De Groot, A. M. B., & Poot, R. (1997). Word translation at three levels of proficiency in a second language: the ubiquitous involvement of conceptual memory. *Language Learning*, 47, 215-264.

De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology*, 50, 77-85.

De Houwer, J., Crombez, G., Koster, E. H. W., & De Beul, N. (2004). Implicit alcohol-related cognitions in a clinical sample of heavy-drinkers. *Journal of Behavioral Therapy and Experimental Psychiatry*, 35, 275-286.

De Jong, P. J., Wiers, R. W., Van de Braak, M., & Huijding, J. (2007). Using the Extrinsic Affective Simon Test as a measure of implicit attitudes towards alcohol: relationship with drinking behavior and alcohol problems. *Addictive Behaviors*, 32, 881-887.

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.

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). Effects of auditory distractors on verbal self-monitoring. *Journal of Cognitive Neuroscience,* 20, 1-14.

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.

Hartsuiker, R. J., & Kolk, H. H. J. (2001). Error monitoring I speech production: a computational test of the perceptual loop theory. *Cognitive Psychology*, 42, 113-157.

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.

Huijding, J., & De Jong, P. J. (2006). Specific predictive power of automatic spider-related affective associations for controllable and uncontrollable fear responses toward spiders. *Behaviour Research and Therapy*, 44, 161-176.

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 of the 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.

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

Kroll, J. F., Michael, E., Tokowicz, N., & Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research*, 18, 137-171.

La Heij, W., Hooglander, A., Kerling, R., & Van der Velden, E. (1996). Nonverbal context effects in forward and backward word translation: evidence for concept mediation. *Journal of Memory and Language*, 35, 648-665.

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.

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

Meara, P. M. (2005). *X_Lex: the Swansea Vocabulary Levels Test. v2.05*. Swansea: Lognostics.

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

Midgley, K. J., Wicha, N. Y. Y., Holcomb, P. J., & Grainger, J. (2007). Gender agreement transfer during sentence processing in early language learners: an electrophysiological study (abstract). In *Proceedings of the annual meeting of Cognitive Neuroscience Society* (pp. 170).

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.

Perani, D., Paulesu, E., Sebastián-Gallés, N., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S. F., Fazio, F., & Mehler, J. (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.

Rodrígues-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.

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.

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.

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.

# Appendix

*List of the 80 common gender and 80 neuter gender Dutch words, which were presented in white in the current experiment. The approximate English translations are given in brackets.*
WHITE COMMON GENDER WORDS: bril (‘glasses’), eend (‘duck’), hoed (‘hat’), schaar (‘scissors’), vis (‘fish’), hand (‘hand’), kop (‘head’), tent (‘tent’), trompet (‘trumpet’), gitaar (‘guitar’), peer (‘pear’), haan (‘rooster’), kwast (‘brush’), tand (‘tooth’), bal (‘ball’), tijger (‘tiger’), pen (‘pen’), zon (‘sun’), schaats (‘skate’), maan (‘moon’), stoep (‘pavement’), kaart (‘map’), emmer (‘bucket’), ladder (‘ladder’), nest (‘nest’), glas (‘glass’), trein (‘train’), hamster (‘hamster’), asperge (‘asparagus’), cello (‘cello’), hammer (‘hammer’), bus (‘bus’), worm (‘worm’), vonk (‘spark’), giraffe (‘giraffe’), jacht (‘hunt’), broer (‘brother’), nacht (‘night’), zomer (‘summer’), wind (‘wind’), wereld (‘world’), muziek (‘music’), pijn (‘pain’), grond (‘ground’), appel (‘apple’), bezem (‘broom’), boom (‘tree’), dolfijn (‘dolphin’), draak (‘dragon’), fakkel (‘torch’), klomp (‘wooden shoe’), koffer (‘suitcase’), muts (‘bonnet’), pijp (‘pipe’), raket (‘rocket’), kat (‘cat’), magneet (‘magent’), sleutel (‘key’), muur (‘wall’), helm (‘helmet’), zaak (‘business’), lamp (‘lamp’), hengel (‘pole’), pleister (‘plaster’), cactus (‘cactus’), das (‘tie’), ooievaar (‘stork’), hark (‘rake’), plant (‘plant’), haai (‘shark’), heks (‘witch’), gieter (‘watering can’), brommer (‘moped’), voet (‘foot’), viool (‘violin’), temple (‘temple’), laars (‘boot’), tas (‘bag’), ster (‘star’), pet (‘pet’).

WHITE NEUTER GENDER WORDS: kleed (‘carpet’), masker (‘mask’), net (‘net’), oog (‘eye’), penseel (‘paint brush’), geweer (‘gun’), beleg (‘sandwich filling’), harp (‘harp’), bot (‘bone’), slot (‘lock’), fornuis (‘furnace’), potlood (‘pencil’), plein (‘square’), hart (‘heart’), circus (‘circus’), hoofd (‘head’), rek (‘rack’), blik (‘can’), web (‘web’), luik (‘hatch’), fruit (‘fruit’), blok (‘block’), robot (‘robot’), laken (‘sheets’), schort (‘apron’), kalf (‘calf’), vest (‘vest’), monster (‘sample’), zadel (‘saddle’), orgel (‘organ’), stuur (‘steering wheel’), schip (‘ship’), bed (‘bed’), lam (‘lam’), schild (‘sheeld’), brein (‘brain’), kind (‘child’), rag (‘cobweb’), jong (‘young’), land (‘country’), uur (‘hour’), woord (‘word’), licht (‘light’), lied (‘song’), park (‘park’), lint (‘ribbon’), podium (‘stage’), gewei (‘antkers’), graf (‘grave’), gras (‘grass’), gezicht (‘face’), loket (‘locket’), kuiken (‘kitchen’), vuur (‘fire’), wapen (‘weapon’), rooster (‘roster’), anker (‘anchor’), gordijn (‘curtain’), bad (‘bath’), harnas (‘armour’), kompas (‘compas’), vergiet (‘colander’), vlot (‘fleet’), zwaard (‘sword’), toilet (‘toilet’), balkon (‘balcony’), dorp (‘village’), boeket (‘bouguet’), bos (‘forest’), bont (‘fur’), fossiel (‘fossil’), kado (‘present’), orakel (‘oracle’), object (‘object’), palet (‘pallete’), ravijn (‘ravine’), spoor (‘rail’), vat (‘barrel’), verkeer (‘traffic’), vlies (‘film’).

*List of the 40 common gender and 40 neuter gender target Dutch words, which were presented in green and blue in the current experiment. The English translations are given in brackets and were used as English targets in the current experiment.*

COMMON GENDER WORDS: poes (‘cat ’), fabriek (‘factory’), stoel (‘chair’), vork (‘fork’), mond (‘mouth’), tafel (‘table’), aap (‘monkey’), kerk (‘church’), muis (‘mouse’), jas (‘jacket’), zoon (‘son’), zeep (‘soap’), kers (‘cherry’), zaag (‘saw’), taart (‘cake’), fles (‘bottle’), klap (‘bang’), schoen (‘shoe’), auto (‘car’), staart (‘tail’), wortel (‘carrot’), bank (‘couch’), lepel (‘spoon’), neus (‘nose’), kast (‘closet’), kan (‘jug’), noot (‘nut’), monnik (‘monk’), tabak (‘tobacco’), zak (‘bag’), trede (‘step’), taal (‘language’), fluit (‘flute’), bijl (‘axe’), fiets (‘bike’), bloem (‘flower’), kaars (‘candle’), wolk (‘cloud’), jurk (‘dress’), trui (‘sweater’).
NEUTER GENDER WORDS: blad (‘leaf’), pak (‘suit’), kasteel (‘castle’), touw (‘rope’), strand (‘beach’), konijn (‘rabbit’), bord (‘plate’), hemd (‘shirt’), been (‘leg’), spook (‘ghost’), schaap (‘sheep’), wiel (‘wheel’), boek (‘book’), paard (‘horse’), raam (‘window’), bureau (‘desk’), brood (‘bread’), papier (‘paper’), schaak (‘chess’), vlees (‘meat’), tapijt (‘carpet’), koren (‘corn’), oor (‘ear’), kanon (‘cannon’), hert (‘deer’), veer (‘ferry’), ei (‘egg’), varken (‘pig’), dak (‘roof’), geld (‘money’), katoen (‘cotton’), hek (‘fence’), beeld (‘statue’), gewicht (‘weight’), paleis (‘palace’), beest (‘animal’), ijs (‘ice’), mes (‘knife’), kruis (‘cross’), spel (‘game’).

# Chapter 7

# General Discussion

This thesis investigated the neurocognitive correlates of verbal self-monitor in healthy adults. It further focused on various factors that can interfere with normal functioning of the verbal self-monitoring system, e.g., time pressure, auditory interference, motivation of speakers, semantic context, and multilingual context. Brain activity was measured using electroencephalography (EEG). This is a non-invasive technique that allows the investigation of on-line information processing in the brain. Particularly, in this thesis the Error-Related Negativity (ERN) was of interest. The ERN is an event-related potential (ERP), which is associated with the working of a general performance monitor.

The studies in this thesis provide empirical evidence that the ERN can be observed after verbal errors and can be affected by linguistic manipulations (e.g., lexical conflict, Chapters 3 and 5). Presence of the ERN after verbal errors might suggest that the verbal monitor works in a similar way as a more general performance monitor. We have shown that the verbal self monitor is affected by time pressure. However, the effect of time pressure on verbal self-monitoring is dependent on whether a task is performed in a native language (Chapter 2) or in a second language (Chapter 3). Further, the verbal monitor is sensitive to auditory interference (Chapter 4), motivation and semantic context (Chapter 5), and bilingual context (Chapter 6). The main findings of this thesis will be discussed below.

In Chapter 2, we addressed the question whether or not an ERN occurs after verbal error detection and whether or not it is affected by time pressure. We found that participants made more errors and showed a decrease in amplitude of the ERN under severe time pressure. Why is verbal monitoring affected by time pressure? According to Levelt (1989), verbal self-monitoring is a controlled process, and therefore resource-limited. 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. Thus, more errors are to be expected at higher speech rates (Levelt et al., 1999). It is also 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 (Levelt, 1989). This is supported by our EEG data. The amplitude of the ERN decreased under time pressure. It is likely that the verbal monitor checks an abstract phonological representation by screening 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 ERN is generated.

Interestingly, in the study described in Chapter 3, we asked participants to perform the same task as in Chapter 2 (i.e., a phoneme-monitoring task with and without time pressure manipulation). The difference between these studies is that in Chapter 2 native Dutch speakers performed task in Dutch, whereas in Chapter 3 native German speakers performed the task in their second language, i.e., Dutch. As native Dutch speakers, German native speakers made more errors under time pressure compared to no time pressure. 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; see Chapter 2). Surprisingly, we did not find the typical decrease of amplitude of the ERN under time pressure when the task was performed in a second language. On the contrary, we observed enhanced ERN under time pressure, as compared to the control condition.

How can we explain this reversed effect of time pressure on the ERN? Participants, in the present study were bilingual German-Dutch students, who were requested to perform a phoneme-monitoring task in their second language, i.e. Dutch. In order 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, when bilingual speakers are aphasic (e.g., Fabbro, Skrap, & Aglioti, 2000), under stress (e.g., Dornik, 1979, 1980; Grosjean, 1982), or undergo brain stimulation (e.g., Holtzheimer, Fawaz, Wilson, & Avery, 2005). It is possible that under time pressure participants had more difficulty inhibiting their dominant native language and experienced more intrusions from it. 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 et al., 2000; Colomé, 2001; Rodríguez-Fornells et al., 2005; but see Hermans et al., 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, which made it more difficult for the monitor to verify which response was correct and which was erroneous. This in turn may have led to more response conflict and higher amplitudes of the ERN (e.g., Botvinick et al., 2001, Yeung et al., 2004).

Presumably, the increased amplitude of the ERN under time pressure when the task is performed in a second language 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). Bilingual participants in the present study completed a Dutch course 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.

In the Chapter 3, we observed that performing the task in a second language reversed the effect of time pressure on the ERN, compared to the effects of time pressure when the same task was performed in the native language. This suggests that the ERN was sensitive to the verbal manipulations. In the following chapter, Chapter 4, we have manipulated the presence and relatedness of auditory distractors to further investigate the effects of verbal manipulation on the ERN. Participants were required to perform a phoneme-monitoring task with semantically related distractors, semantically unrelated ones, or in the absence of distractors. Error rate was independent of distractors to a target picture. This was somewhat unexpected. According to the perceptual loop theory, the verbal monitor proceeds through the speech comprehension system (Levelt et al., 1999). The same system is used for processing auditory input from others. Thus, in presence of auditory distractors, the comprehension system must process auditory input and simultaneously monitor inner speech, which inevitably should lead to a higher error rate. This suggests that the verbal monitoring may be more independent of the comprehension system than the perceptual loop theory presumes. Research on aphasic patients, for instance, demonstrated a double dissociation between the comprehension and the monitoring system (e.g., intact comprehension but impaired monitoring or vice versa; Marshall, Robson, Pring, & Chiat, 1998; Marshall, Rappaport, & Garcia-Bunuel, 1985).

However, we did observe a typical increase in response latencies for semantically related relative to semantically unrelated distractors (Damian et al., 2001; Schriefers et al., 1990). Additionally, the amplitude of the ERN was larger on trials when distractors were semantically related to the target picture compared to semantically unrelated ones. Semantically related distractors presumably co-activate, through the spreading of activation, multiple concepts that are semantically related to one another (Maess, Friederici, Damian, Meyer, & Levelt, 2002). Hence, there are multiple plausible entries that are simultaneously active and compete for lexical selection (but see Finkbeiner et al., 2006 and Mahon et al., 2007). Hence, in the semantically related condition at the time of a response there may be more conflict between competing responses than in the semantically unrelated condition or in the absence of distractor words, thereby leading to a larger ERN amplitude. The verbal self-monitor needs to verify on-line whether the correct entry was chosen from the pool of competing items, which in turn leads to slower responses. In the unrelated condition, however, such verification might not be as relevant, since unrelated distractors do not lead to activation of related concepts, and therefore less competition may be present at the time of the response.

This finding was replicated and extended in a study described in the Chapter 5. Contrary to the previous studies, where we used a phoneme-monitoring task (Chapters 2, 3, and 4), in the Chapter 5 we employed a picture naming task. Even though a phoneme monitoring task is verbal in nature and involves monitoring of internal speech production, it is possible that at least some errors found in phoneme-monitoring were motor slips and not verbal errors *per se*. In contrast, in the task described in the Chapter 5, all given responses were verbal. Participants were required to name pictures in a semantic context, in which all pictures were from the same semantic category, and in a mixed context, in which all pictures were from different semantic categories. Additionally, we manipulated participants' motivation, i.e. for errors in the high-motivation condition, participants were told to be financially punished, whereas for errors in the low-motivation condition, participants received neither financial punishment nor reward.

We found a typical semantic interference effect (Damian et al., 2001; Lupker, 1979; Schriefers, Meyer, & Levelt, 1990; Schnur et al., 2006; Vitkovitch & Humphreys, 1991). Participants were slower and made more errors in naming pictures in semantically related blocks as opposed to mixed blocks. Additionally, we observed enhanced ERN in semantically related blocks than in mixed blocks. The enhancement of the ERN and selection latencies in semantically related blocks was most likely due to the simultaneous activation of competing lexical items, which in turn led to a higher conflict at a time of response. This is in accordance with the results that we showed in the previous study, where amplitude of the ERN was larger in presence of semantically related distractors than semantically unrelated ones (Chapter 4). We argue that due to the simultaneous activation of competing items, the verbal self-monitor is presumably more alert in the semantically related context than in the mixed context in order to validate whether the given response was correct or not.

The motivation manipulation had no effect on naming latencies or error rates, which indicates that participants did not develop different response strategies in the high-motivation and the low-motivation conditions. In the electrophysiological data, we saw a clear ERN on error trials compared to correct trials. However, the amplitude and the latency of the ERN were affected by the motivation manipulation. We observed enhanced and delayed ERN in the high-motivation condition, compared to the low-motivation condition. In the high-motivation condition, errors had more severe consequences for participants than in the low-motivation condition. It is plausible that the ERN is sensitive to the motivational significance of errors (Hajcak et al., 2005). Under circumstances when errors have more severe consequences, e.g. when giving a speech, the verbal monitor has to be more alert in order to verify whether or not the selected response was correct. It is possible that a delayed ERN latency reflects prolonged error-monitoring mechanism (Johannes et al., 2001). In the high-motivation condition, error monitoring may have been slowed in order to verify that the selected response is indeed a correct response. In the low-motivation conditions, such verification was faster and of less significance, since errors did not have severe consequences.

The goal of the study described in Chapter 6 was to investigate how the ERN is affected by conflict in a bilingual situation. Dutch-English bilinguals saw Dutch words in white print that needed to be classified (right or left button-press) according to their grammatical gender. Furthermore, they were presented with colored words that were to be classified on the basis of their color. Colored words consisted of Dutch common and neuter gender words and their English translations. Performance was more erroneous on incongruent trials, in which there was a mismatch between color and gender response mapping, compared to congruent trials, i.e. in case no such discrepancy was present. We obtained an ERN following incorrect classifications for colored words which was larger for incongruent than congruent trials. Higher error rates and enhanced amplitude of the ERN on incongruent trials were independent of the language in which target words were presented. This may suggest that when multiple languages are active, the verbal monitor has more difficulty to keep languages separated. Therefore, it may suffer more from intrusions from a second language, resulting in more response conflict and more error-prone performance. These results also provide evidence that under certain circumstances people can transfer some grammatical characteristics of their first language (e.g. gender) to their second language, even if such characteristics are absent from the latter.

Taking all results of our studies together, it seems that there is a conceptual overlap between verbal monitoring and general performance monitoring theories. Both types of 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. It is possible that during verbal monitoring, as well as during executive action monitoring, a copy of the on-line response is created and compared to the representation of the correct response (see Levelt et al., 1999). If there is a mismatch between them, an error signal is generated and corrective processes can be started – similar to action monitoring (see Desmurget & Grafton, 2000; Rodríguez-Fornells et al., 2002). In a series of experiments we showed a typical ERN in a various tasks in which performance was dependent on a verbal judgment (Chapters 2, 3, 4, 5, and 6). More specifically, in Chapters 3 and 5, we showed that the ERN was not only present after verbal errors, but was also affected by lexical conflict, which presumably was the result of simultaneous activation of multiple concepts from the same semantic category. Additionally, there is further recent evidence that verbal errors activate the anterior cingulated cortex (ACC) and medial frontal cortex (SMA; Möller, Jansma, Rodríguez-Fornells, & Münte, 2007). A typical ERN has been located within the ACC/SMA region (Dehaene et al., 1994). Based on this evidence, we suggest that verbal monitoring is not a process separate from but rather a special case of general performance monitoring.

Interestingly, in Chapters 2 and 4 we also demonstrated an ERN-like response on incorrect go-trials (i.e. misses). This type of 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 was initiated at the time of mean response latency. Additionally, the negative deflection after misses was affected by time pressure (Chapter 2) in a similar way as the ERN following false alarms. The amplitude of the negativity after misses decreased under time pressure, as compared to its amplitude during the control condition (Chapter 2). Unfortunately, in the Chapter 4, we did not have enough trials to investigate whether or not the auditory distractors had an effect on the ERN-like response after misses. 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 which are characterized by the absence of such overt motor responses.

To conclude, our findings provide converging evidence that verbal monitoring might be a special case of general performance monitoring instead of a completely separate process. Furthermore, our results suggest that the ERN can be used as an electrophysiological marker of error processing in psycholinguistic research.

## References

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.

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.

Damian, M. F., , G., & Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition,* 81, B77-B86.

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.

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.

Finkbeiner, M., Gollan, T. H., & Caramazza, A. (2006). Lexical access in bilingual speakers: What's the (hard) problem? *Bilingualism: Language and Cognition*, 9, 153-166.

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

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: Harvard University Press.

Hajcak, G., Moser, J. S., Yeung, N., & Simons, R. F. (2005). On the ERN and the significance of errors. *Psychophysiology*, 42, 151-160.

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.

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.

Johannes, S., Wieringa, B. M., Nager, W., Rada, D., Dengler, R., Emrich, H. M., Münte, T. F., & Dietrich, D. E. (2001). Discrepant target detection and action monitoring in obsessive-compulsive disorder. *Psychiatry Research: Neuroimaging Science*, 108, 101-110.

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

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

Lupker, S. J. (1979). The semantic nature of response competition in the picture-word interference task. *Memory & Cognition, 7*, 485-495.

Marshall, J., Robson, J., Pring, T., & Chiat, S. (1998). Why does monitoring fail in jargon aphasia? Comprehension, judgment, and therapy evidence. *Brain and Language*, 63, 79-107.

Marshall, R.C., Rappaport, B.Z., & Garcia-Bunuel, L. (1985). Self-monitoring behavior in a case of severe auditory agnosia with aphasia. *Brain and Language*, 24, 297-313.

Maess, B., Friederici, A. D., Damian, M., Meyers, A. S., & Levelt, W. J. M. (2002). Semantic category interference in overt picture naming: sharpening current density localization by PCA. *Journal of Cognitive Neuroscience,* 14, 455-462.

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 33, 503-535.

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, U.K.: Oxford University Press.

Möller, J., Jansma, B. M., Rodriguez-Fornells, A., & Münte, T. F. (2007). What the brain does before the tongue slips. *Cerebral Cortex*, 17, 1173-1178.

Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture–word interference studies. *Journal of Memory and Language*, 29*,* 86–102.

Schnur, T. T., Schwartz, M. F., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language*, 54, 199-227.

Rodriguéz-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.

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

Vitkovitch, M., & Humphreys, G. W. (1991). Perseverant responding in speeded naming of pictures: it's in the links. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 17, 664-680.

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.

# Chapter 8

# Summary/Samenvatting

## 8.1 Summary

This thesis investigated the neurocognitive correlates of verbal self-monitoring in healthy adults. Particularly, in this thesis the Error-Related Negativity (ERN) was of interest. The ERN is a brain potential, which is associated with the working of a general performance monitor. The central questions addressed in the thesis are: If the ERN is associated with error processing in action monitoring, can it also be applied to error processing in verbal monitoring? Does *verbal monitoring* work in a similar way as *action monitoring*?

Chapter 2 addressed the question whether or not an ERN occurs after verbal error detection and whether or not it is affected by time pressure. In the perceptual loop theory (Levelt, 1983), the monitor is assumed to be resource-limited. In this case, the monitor should be affected by a time-pressure manipulation. Participants had to perform a phoneme-monitoring task with and without a time pressure manipulation. The results show that participants made more errors and showed a decrease in ERN amplitude under severe time pressure. It is likely that the verbal monitor checks an abstract phonological representation by screening 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. 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 ERN is generated and more errors pass undetected or corrective processes are not activated fast enough.

Interestingly, in the study described in Chapter 3, German-Dutch bilinguals performed the same task as in Chapter 2 (i.e., a phoneme-monitoring task with and without time pressure manipulation) in their second language, i.e. Dutch. As native Dutch speakers, German native speakers made more errors under time pressure compared to no time pressure. Surprisingly, however, there were enhanced ERN amplitudes under time pressure, as compared to the control condition. It is possible that under time pressure participants had more difficulty inhibiting their dominant native language and experienced more intrusions from it. Hence, it is possible that at the time of response, there was not only the Dutch name of the picture active but also the German name, which made it more difficult for the monitor to verify which response was correct and which was erroneous. This, in turn, may have led to more response conflict and higher amplitudes of the ERN.

In the following chapter, Chapter 4, the presence and relatedness of auditory distractors was manipulated to further investigate the effects of verbal manipulation on the ERN. The ERN is sensitive to the response conflict. However, there are no studies showing that the ERN can be affected by lexical retrieval conflict. Sensitivity of the ERN to the increase of lexical conflict might provide additional support for the hypothesis that verbal monitoring involves similar processes as non-verbal action monitoring. Participants were required to perform a phoneme-monitoring task with semantically related distractors, semantically unrelated ones, or in the absence of any distractors. Error rate was independent of distractors to a target picture. However, there was an increase in response latencies and amplitude of the ERN for semantically related relative to semantically unrelated distractors. Semantically related distractors presumably co-activate, through the spreading of activation, multiple concepts that are semantically related to one another competing for lexical selection. Hence, in the semantically related condition at the time of response there was more conflict between competing responses than in the semantically unrelated condition or in the absence of distractor words, thereby leading to larger ERN amplitudes. The verbal self-monitor needs to verify on-line whether the correct entry was chosen from the pool of competing items, which in turn leads to slower responses. In the unrelated condition, however, such verification might not be as relevant, since unrelated distractors do not lead to activation of related concepts, and therefore less competition may be present at the time of response.

This finding was replicated and extended in the study described in Chapter 5. Contrary to the previous studies, in which a phoneme-monitoring task was used (Chapters 2, 3, and 4), in the Chapter 5 a picture naming task was employed. Participants were required to name pictures in a semantic context, in which all pictures were from the same semantic category, and in an unrelated context, in which all pictures were from different semantic categories. Additionally, participants' motivation was manipulated, i.e. for errors in the high-motivation condition, participants were told to be financially punished, whereas for errors in the low-motivation condition, participants received neither financial punishment nor reward.

Participants were slower, made more errors, and showed enhanced ERN amplitudes in naming pictures in semantically related blocks as opposed to unrelated blocks. The enhancement of the ERN and selection latencies in semantically related blocks was most likely due to the simultaneous activation of competing lexical items, which in turn led to a higher conflict at a time of response. It is possible that due to the simultaneous activation of competing items, the verbal self-monitor is presumably more alert in the semantically related context than in the mixed context in order to validate whether or not the given response was correct.

The motivation manipulation had no effect on naming latencies or error rates. However, in the electrophysiological data, the amplitude and the latency of the ERN were affected by the motivation manipulation. Enhanced and delayed ERN was observed in the high-motivation condition, compared to the low-motivation condition. In the high-motivation condition, errors had more severe consequences for participants than in the low-motivation condition. It is plausible that under circumstances in which errors have more severe consequences the verbal monitor has to be more alert in order to verify whether or not the selected response was correct. In the high-motivation condition, error monitoring may have been slowed in order to verify that the selected response is indeed the correct response. In the low-motivation condition, such verification was faster and of less significance, since errors did not have severe consequences.

The goal of the study described in Chapter 6 was to investigate whether and how the ERN is affected by conflict in a bilingual situation. Dutch-English bilinguals saw Dutch words in white print that needed to be classified (right or left button-press) according to their grammatical gender and colored words that were to be classified on the basis of their color. Colored words included Dutch common and neuter gender words, and English translations of these words. Performance was more error-prone on incongruent trials, in which there was a mismatch between color and gender response mapping, as compared to congruent trials, in which no such discrepancy was present. There was an ERN following incorrect classifications for colored words which was larger for incongruent than congruent trials. Higher error rates and enhanced amplitude of the ERN on incongruent trials were independent of the language in which target words were presented. This may suggest that when multiple languages are active, the verbal monitor has more difficulty to keep languages separated and therefore suffers more from intrusions from a second language, resulting in more response conflict and more error-prone performance (see also Chapter 3). These results also provide evidence that under certain circumstances speakers can transfer grammatical characteristics of their native language (e.g. gender) to their second language, even if such characteristics are absent from the latter.

Taking all results of studies in the present thesis together, it seems that there is a conceptual overlap between verbal monitoring and general performance monitoring theories. Both types of 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. It is possible that during verbal monitoring, as well as during executive action monitoring, a copy of the on-line response is created and compared to the representation of the correct response. If there is a mismatch between them, an error signal is generated and corrective processes can be started. In a series of experiments, a typical ERN was shown in a various tasks in which performance was dependent on a verbal judgment (Chapters 2, 3, 4, 5, and 6). More specifically, in Chapters 3 and 5, the ERN was not only present after verbal errors, but was also affected by lexical conflict, which presumably was the result of simultaneous activation of multiple concepts from the same semantic category.

To conclude, these findings provide converging evidence that verbal monitoring might be a special case of general performance monitoring instead of a completely separate process. Furthermore, these results suggest that the ERN can be used as an electrophysiological marker of error processing in psycholinguistic research.

# 8.2 Samenvatting

In dit proefschrift worden de neurologische maten voor verbale zelfmonitoring in gezonde volwassenen onderzocht. In het bijzonder ging mijn interesse uit naar de error-related negativity (ERN). De ERN is een hersenpotentiaal, die te maken heeft met de werking van een algemene uitvoeringsmonitor, een cognitieve module die de correctheid van ons gedrag controleert. De centrale vragen in het proefschrift zijn: Als de ERN te maken heeft met het verwerken van fouten bij processen van de uitvoeringsmonitor, kan deze dan ook gebruikt worden bij processen van de verbale monitor? Werkt de verbale monitor op een vergelijkbare manier als een actie-monitor?

In Hoofdstuk 2 wordt een ERP-studie beschreven, waarin is onderzocht of de ERN gebruikt kan worden om verbale fouten op te sporen en of de ERN wordt beïnvloedt als proefpersonen onder tijdsdruk worden gezet. In de *perceptual-loop* theorie (Levelt, 1983), wordt de monitor gezien als een bronafhankelijk proces. Als dat het geval is zou de monitor beïnvloedt moeten worden door manipulatie van de tijdsdruk. Proefpersonen moesten een foneemdetectie taak zowel met als zonder tijdsdruk uitvoeren. Het blijkt dat dat proefpersonen meer fouten maken en een lagere ERN hebben onder flinke tijdsdruk dan zonder tijdsdruk. Het is dus waarschijnlijk dat de verbale monitor een abstracte fonologische representatie controleert op verschillen tussen de bedoelde en daadwerkelijke respons. De verbale respons wordt met de intentionele response vergeleken, en als er overeenkomst is, dan wordt er een fout gesignaleerd.

Het is mogelijk dat er onder tijdsdruk niet genoeg tijd beschikbaar is om een optimale vergelijking tussen de correcte en de daadwerkelijke respons te maken. Als een gevolg hiervan wordt er een zwakkere ERN gegenereerd en worden meer fouten niet opgemerkt of worden corrigerende processen niet snel genoeg geactiveerd.

Hoofdstuk 3 beschrijft een studie, waarin Duitse tweetalige proefpersonen dezelfde taak als in Hoofdstuk 2 (een foneem-detectie taak met en zonder tijdsdruk manipulatie) in het Nederlands, hun tweede taal, moesten uitvoeren. Net als mensen met Nederlands als moedertaal, maakten de Duitse proefpersonen met Nederlands als tweede taal ook meer fouten in de situatie met tijdsdruk in vergelijking met de situatie zonder tijdsdruk. Echter, tegen de verwachting in, vond ik hogere ERN scores in de situatie met tijdsdruk dan zonder tijdsdruk als de taak in de tweede taal wordt uitgevoerd. Het is mogelijk dat onder tijdsdruk proefpersonen meer moeilijkheden hadden om hun dominante moedertaal te onderdrukken en daardoor meer last hadden van intrusies van de moedertaal. Ook is het mogelijk dat op het moment van de respons er niet alleen de Nederlandse naam van het plaatje beschikbaar was maar ook de Duitse naam. Het was daardoor voor de monitor moeilijker is om te bepalen welke respons correct en welke respons foutief was. Dit zou geleid kunnen hebben tot meer conflict tussen strijdende responsen en hogere amplitude van de ERN.

Vervolgens heb ik in Hoofdstuk 4 de aanwezigheid en de relatie tussen auditieve distractoren gemanipuleerd om verder te onderzoeken wat de effecten van verbale manipulaties op ERN zijn. De ERN is gevoelig voor grotere conflicten tussen strijdende responsen. Echter, tot op heden was er nooit onderzocht of de ERN ook gevoelig is voor een lexicaal conflict. Gevoeligheid van de ERN voor een toename van een lexicaal conflict zou extra evidentie kunnen zijn voor de hypothese dat verbale monitoring vergelijkbare processen gebruikt als non-verbale actie monitoring. Proefpersonen moesten een foneem-detectie taak uitvoeren met semantisch gerelateerde-, semantisch ongerelateerde distractoren, en zonder distractoren. Het foutenpercentage was niet afhankelijk van de distractoren. Maar ik vond wel een verhoging in de reactietijden en de amplitude van de ERN voor semantisch gerelateerde distractoren in vergelijking met semantisch ongerelateerde distractoren. Semantische distractoren activeren kennelijk, door de spreiding van activatie, meerdere concepten die semantisch aan elkaar gerelateerd zijn en die met elkaar strijden voor lexicale selectie. Daardoor was er in de semantisch gerelateerde conditie meer conflict tussen strijdende responsen dan in de semantisch ongerelateerde conditie of bij afwezigheid van afleidende woorden. Dit heeft als gevolg hogere amplitudes van de ERN en tragere reactietijden. De verbale monitor moet verifiëren dat het daadwerkelijke antwoord de correcte respons is en dit zorgt voor een vertraagde reactie. In de ongerelateerde conditie is deze verificatie van minder belang. Dit omdat ongerelateerde distractoren niet leiden tot co-activatie van gerelateerde lexicale concepten en dus tot minder competitie tussen de alternatieve antwoorden.

In Hoofdstuk 5 worden bovenstaande resultaten herhaald en verder uitgebreid. In tegenstelling tot de vorige studies, waar ik een foneem-detectie taak heb gebruikt (Hoofdstukken 2, 3, en 4), heb ik in Hoofdstuk 5 een plaatjesbenoeming taak gebruikt. Proefpersonen werd gevraagd plaatjes te benoemen in een semantische context (b.v., hand, mond, arm, been, en voet) en een ongerelateerd context (b.v., hand, giraffe, stoel, jurk, en bus). Daarnaast heb ik ook de motivatie van proefpersonen beïnvloedt. Voor gemaakte fouten in de zogenaamde hoog-motivatie conditie kregen proefpersonen te horen dat er een financiële straf zou volgen. In de laag-motivatie conditie kregen de proefpersonen geen straf maar ook geen beloning.

Proefpersonen waren langzamer, maakten meer fouten, en de amplitude van de ERN was groter in de benoeming van plaatjes bij semantisch gerelateerde blokken in vergelijking met gemengde blokken. De verhoging van de amplitude van de ERN en reactietijden in semantisch gerelateerde blokken is hoogstwaarschijnlijk een gevolg van de gelijktijdige activering van concurrerende lexicale items waardoor er meer conflict plaatsvondt op het moment van de respons. Als gevolg van de gelijktijdige activering van concurrerende items, is de verbale zelfmonitor meer actief in de semantische gerelateerde context dan in de ongerelateerd context, om te verifiëren dat de daadwerkelijke respons de correcte respons was.

Manipulatie van de motivatie had geen effect op reactietijden of foutenscores. Echter, de amplitude van de ERN bleek hoger en vertraagd in de hoog-motivatie conditie in vergelijking met de laag-motivatie conditie. In de hoog-motivatie conditie hadden fouten meer consequenties voor sprekers dan in de laag-motivatie conditie, dus moest de verbale monitor alerter zijn om te verifiëren of de geselecteerde respons correct was of niet. In de hoog-motivatie conditie was foutendetectie waarschijnlijk vertraagd om te verifiëren dat de geselecteerde respons daadwerkelijk de correcte respons was. In de laag-motivatie conditie was deze verificatie sneller en minder belangrijk, omdat fouten geen erge gevolgen hadden.

In Hoofdstuk 6 wordt onderzocht hoe de ERN door een meertalige context wordt beïnvloed. Nederlands-Engels tweetaligen zagen witte Nederlandse woorden die ze moesten classificeren. Hun taak was om een beslissing te nemen over het grammaticale geslacht van het woord. Proefpersonen kregen ook gekleurde woorden te zien die ze moesten classificeren op basis van hun kleur. De gekleurde woorden waren Nederlandse mannelijke en onzijdige woorden en de Engelse vertaling van deze woorden. De proefpersonen presteerden slechter bij incongruente trials wanneer er geen overeenkomst tussen kleur en grammaticale geslacht was. Op congruente trials was hier van geen sprake. Uit het onderzoek bleek dat proefpersonen meer fouten maakten en een hogere ERN hebben op incongruente dan congruente trials. Dit effect was onafhankelijk van de taal waarin de gekleurde woorden werden gepresenteerd. Dit kan erop wijzen dat wanneer meerdere talen tegelijk actief zijn, de monitor meer moeilijkheden heeft om de talen apart te houden en daarom meer last heeft van intrusies van de andere taal, wat leidt tot meer conflict tussen strijdende responsen en meer fouten (zie ook Hoofdstuk 3). Deze resultaten laten ook zien dat sprekers onder bepaalde omstandigen sommige grammaticale karakteristieken van hun eerste taal naar hun tweede taal kunnen overdragen (b.v. grammaticaal geslacht), ook als zulke karakteristieken eigenlijk afwezig zijn in de tweede taal.

Samenvattend blijkt uit dit onderzoek dat er een conceptuele overlap is tussen theorieën over de verbale monitor en een algemene uitvoeringsmonitor. Beide theorieën claimen onafhankelijk van elkaar dat, om een fout op te kunnen sporen, de monitor een vergelijking moet maken tussen de correcte respons en een kopie van een daadwerkelijke respons. Als er geen overeenkomst is, dan is er sprake van een fout en beginnen corrigerende processen te werken. In een reeks van experimenten heb ik aangetoond dat er een typische ERN voor verbale fouten in verschillende verbale taken was en dat de ERN afhankelijk was van een verbale beslissing (Hoofdstukken 2, 3, 4, 5, en 6). Sterker nog, in de Hoofdstukken 4 en 5, heb ik laten zien dat de ERN niet alleen aanwezig was voor verbale fouten maar ook gevoelig was voor lexicale conflicten.

Mijn bevindingen leveren convergerend bewijs dat de verbale monitoring niet een apart proces is, maar eerder een specificatie van de algemene monitor. Bovendien tonen huidige bevindingen aan dat de ERN gebruikt kan worden als elektrofysiologisch middel om de verwerking van fouten in taalonderzoek te meten.

# Acknowledgements

It's a special feeling to bring a huge story to the end. However, a PhD thesis is not written by itself, but is the result of the cooperation between a lot of people. I would like to thank all the people who have helped me travel through my PhD years and made writing this PhD thesis possible.

First of all, my deepest thanks to my parents. Thanks for giving me the chance of becoming who I am. Pap, thanks for making me question everything and doubt even the most solid facts. Thanks for all the evening battles, in which we conquered and defeated 'nice and friendly' creatures and of course hunt for yellow dots. Mam, thanks for keeping me motivated and stuffing me with all the nice food to keep me going. :)

Jeroen thanks for being there for me and for believing in me. I think you got the hardest job of needing to handle me on an everyday basis. And dealing with my frustrations with statistics, uncooperative data, and software packages that refused to do what I wanted them to do. And of course, accepting the fact that one day you might wake up in a jungle with ferocious tigers. ;) Thank you. Dette and Will thank you for your support.

Marjolein, Nele, and Bianca thank you all for keeping me company on a daily basis via MSN. Thank you for listening, helping, and just having fun :) Mar, I still remember how many years ago we were filling in our university applications  and you firmly stated that psychology is not a real science and you wanted to do something more scientific and ventured to study physics, chemistry, and biology. And I went off studying psychology. Both of us are now promoted in psycholinguistics :) Thanks for helping me with selecting countless semantically related words, and helping constructing sentences that made sense. Thanks for the discussions about life, universe, and everything (and believing that answer is 42). I'm not sure if I'm that grateful for introducing me to multiple series of books resulting in me getting addicted to all of them and wanting to have them all! ;)  Nele thanks for all the good times we had :) And I think you are the only person whom I know, who thinks that statistics are fun and trying with full enthusiasm to make other people believe it. I must admit, that you did not manage to convince me that statistics are fun, but you definitely killed my fear of diving into the world of numbers and formulas. After my statement that I hate InDesign, screaming out that it is "A great program. I love it!" and immediately providing me with a handbook and a number of links to websites. This thesis would not turn out to be like this without you. Bianca thanks for all the help with correcting my Dutch, being there for me, for all the coffee we had, for all the conversations, and generally for all the fun.

Thomas thank you for believing in me, fueling my interest in criminology, having stimulating discussions, and just relaxing evenings :) Mart thanks a lot! You were a great help at work as well as outside of work. Thanks a lot for helping with battles with Presentation and the Zerg. Thanks for all the wine-, good food-, and game-evenings.

Heidi, Iemke, Ingrid, and Dirk without you this thesis would not have been possible. Thank you for the brainstorms about experiment ideas, design, and making sense of the data. Iemke and Heidi thank you for making me focus on other things in life and thanks for making me consider making changes. Rinus thank you for e-prime, lab talks and passionate stories about Japan :)

And last but not least, my gratitude to Vivian. Thank you, Vivian for struggling through my Dutch summary and making it readable and understandable.

Thank you all!

# Curriculum vitae

A. (Lesya) Y. Ganushchak was born in Nikolaev (Ukraine), on February 18th 1978. She attended the Jeanne D'Arc College in Maastricht, where she obtained an International Bachelorate in 1998. She commenced to study psychology at Maastricht University in 1998, where she became a teaching assistant in 2001. She graduated in 2003, with specializations in Neuropsychology and Educational Psychology. She was employed as a Ph.D. student in 2003 at the department of Cognitive Neuroscience at Maastricht University, and in 2006 transferred to the section of Cognitive Psychology at Leiden University.