Event-related brain potentials during the monitoring of speech errors

Niels O. Schiller [a,b,*], Iemke Horemans [b], Lesya Ganushchak [a], Dirk Koester [a,c]

[a] *Leiden Institute for Brain and Cognition, Leiden University, The Netherlands*
[b] *Department of Cognitive Neuroscience, Faculty of Psychology, Maastricht University, The Netherlands*
[c] *F. C. Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands*

ARTICLE INFO

ABSTRACT

When we perceive speech, our goal is to extract the meaning of the verbal message which includes semantic processing. However, how deeply do we process speech in different situations? In two experiments, native Dutch participants heard spoken sentences describing simultaneously presented pictures. Sentences either correctly described the pictures or contained an anomalous final word (i.e. a semantically or phonologically incongruent word). In the first experiment, spoken sentences were task-irrelevant and both anomalous conditions elicited similar centro-parietal N400s that were larger in amplitude than the N400 for the correct condition. In the second experiment, we ensured that participants processed the same stimuli semantically. In an early time window, we found similar phonological mismatch negativities for both anomalous conditions compared to the correct condition. These negativities were followed by an N400 that was larger for semantic than phonological errors. Together, these data suggest that we process speech semantically, even if the speech is task-irrelevant. Once listeners allocate more cognitive resources to the processing of speech, we suggest that they make predictions for upcoming words, presumably by means of the production system and an internal monitoring loop, to facilitate lexical processing of the perceived speech.

© 2008 Elsevier Inc. All rights reserved.

Introduction

When we communicate verbally during a conversation, the interplay between speaking and listening is crucial to understand each other efficiently, but has received relatively little attention in the psycholinguistic literature (but see Schiller and Meyer, 2003). Monitoring one's own and the speech of others is very important for fluent communication. Without monitoring, producing speech can potentially lead to embarrassment, for instance, when taboo words are uttered unintentionally (so-called *slips of the tongue*; Motley, Camden, and Baars, 1982), or speech output can result in awkward mishearing (so-called *slips of the ear*; Garnes and Bond, 1980).

When a speech planning error is made at the phonological processing level, it may be picked up by means of *internal self-monitoring*, i.e. the so-called *inner speech* is checked for errors (see Hartsuiker and Kolk, 2001; Postma, 2000; Levelt,1983, 1989; Levelt, Roelofs, and Meyer, 1999). *Overt speech*, in contrast, is evaluated through *external self-monitoring* (Christoffels, Formisano, and Schiller, 2007 for neurocognitive evidence). In internal and external self-

monitoring, information from several processing levels is first delivered to the speech comprehension system, where it is parsed and then transferred to the verbal monitor. The verbal monitor compares the parsed speech and the intentions of the speaker to the linguistic standards.

However, there are situations in which it is questionable whether listeners fully process the perceived speech signal (Chwilla, Brown, and Hagoort, 1995). As listeners, we are often presented with what may be called *irrelevant speech*. For example, when we wait at a bus stop, other people may have a conversation while we read something, or sales promotions via loud speakers in the supermarket may be irrelevant to us, at least sometimes. Presumably, language processing is a highly automatic process because we speak and listen (and write and read) on a daily basis. However, it is probably also a costly process in terms of cognitive resources, and therefore it may be economic to avoid processing costs in case speech is irrelevant. In the current study, we investigated the processing of relevant and irrelevant speech errors in verbal descriptions of pictures by means of event-related potentials (ERPs).

Different types of errors are associated with different components in ERP studies. In their seminal study, Kutas and Hillyard (1980) demonstrated that semantically anomalous words in sentences elicit a more negative deflection in the ERP signal than semantically appropriate words. This ERP component peaks around 400 ms after word onset and is called the N400 effect. The N400 effect has been associated with lexical and post-lexical processing, e.g. lexical access

* Corresponding author. Leiden Institute for Brain and Cognition, Leiden University, Faculty of Social Sciences, Department of Psychology, Cognitive Psychology Unit, P.O. Box 9555, 2300 RB Leiden, The Netherlands. Fax: +31 71 5273783.
E-mail address: nschiller@fsw.leidenuniv.nl (N.O. Schiller).

and semantic integration of words into context (Kutas, and Feder-meier, 2000; Kutas and Van Petten, 1994; Van Petten and Luka, 2006 for reviews; see also Holcomb, 1993). In contrast, syntactic anomalies were found to elicit a positive deflection peaking around 600 ms following the anomalous word onset, the P600 (Hagoort, Brown, and Groothusen, 1993; see also Friederici, Pfeifer, and Hahne, 1993). Note that recent studies also found P600 effects in response to semantically implausible sentences (Kolk, Chwilla, Van Herten, and Oor, 2003; Kuperberg, 2007).

For the processing of a phonological mismatch in the onset of an expected and an actually heard word, a so-called phonological mismatch negativity (PMN) has been reported. The PMN is a negative-going ERP component that has been linked to the initial (pre-lexical) phonological processing stage of auditory speech perception (Connolly, Byrne, and Dywan, 1995; Connolly and Phillips, 1994). Temporally, the PMN precedes the (auditory) N400 and has been shown to be largely independent of the N400, since it occurred regardless of the semantic appropriateness of the spoken words (Connolly and Phillips, 1994; D'Arcy et al., 2004 for an overview). It is usually identified in the ERP signal as the most negative peak between 150 and 350 ms after stimulus onset.

Connolly et al. (1995), for instance, presented black-on-white line drawings to participants and simultaneously spoken words in English via headphones. Spoken words were either semantically congruent (e.g. *tree*) with the target picture (e.g. *TREE*) or semantically incongruent and beginning with an unexpected phoneme (e.g. *cup*). Participants' task was to decide whether the spoken word matched the visual stimulus. It was found that PMN amplitudes were significantly larger to incongruent than to congruent picture–word pairs. The onset of this PMN, which had a centro-parietal topographical distribution, occurred before the auditory N400, supporting the idea that auditory word recognition begins before the completion of a spoken word. Identifying the initial phoneme of the auditory stimulus was sufficient to detect the mismatch between picture name and spoken word (Connolly et al., 1995).

The aim of the current study is to investigate the time course of the processing of verbal descriptions of pictures including speech errors, in particular phonological in comparison to semantic errors. As mentioned above, language users are sometimes confronted with irrelevant speech. This raises the question whether or not speech errors are detected in irrelevant speech and if so whether they are processed in a similar way as in relevant speech. So far, the processing of irrelevant speech has not received much attention. It has been used to investigate the architecture of the phonological loop although with simple stimuli (syllables; Martin-Loeches, Schweinberger, and Sommer, 1997), and morphosyntactic features that are syntactically irrelevant have been investigated while the speech was attended (Koester, Gunter, Wagner, and Friederici, 2004; Koester, Gunter, and Wagner, 2007). In the first experiment of the current study, we investigated the processing of different types of speech errors when participants performed a non-linguistic probe detection task. In addition to *phonological errors* (e.g. *d*rop the *d*ish), which contain phonological overlap with the correct target word (e.g. drop the f*ish*), we also included so-called *semantic errors* which are semantically anomalous and have no phonological overlap with the target (e.g. drop the *sky*).

Experiment 1: task-irrelevant monitoring of speech errors

In our study, we displayed pictures to create a constricting context while simultaneously presenting naturally spoken sentences describing the pictures. Some of the sentences contained the above-mentioned speech errors. We were mainly interested in the electro-physiological responses to these errors while participants performed an unrelated probe detection task. If irrelevant speech is not semantically processed (e.g. to avoid cognitive processing costs), no

difference in N400 amplitude is expected between different conditions. For example, Chwilla et al. (1995) obtained an N400 priming effect in a semantic priming task (word-to-word priming) using a lexical decision task but not when they used a visual task (case judgment) in which the content of the language stimuli was irrelevant. These authors suggested that lexical–semantic information, as measured by the N400 amplitude, is not processed if the word is not part of an episodic trace of the stimulus event, i.e. if the lexical–semantic information is task-irrelevant. If, however, participants process irrelevant speech up to a semantic level, an N400 effect is expected at least for semantic errors (e.g. Kutas and Van Petten, 1994; Van Petten and Luka, 2006). Phonological errors (e.g. *dish* instead of *fish* due to the preceding word *drop*) are also semantically anomalous, and should therefore result in an N400 effect. However, they may elicit a different electrophysiological signature since they perseverate certain phonological features (e.g. the segment /d/) of the influencing syllable (e.g. *d*rop the *d*ish).

Method

Participants

Twenty students (17 women) of Maastricht University in the Netherlands participated in the current study (mean age 21.9 years; range 18–29). All participants were right-handed, native speakers of Dutch with normal or corrected-to-normal vision and audition. Participants received a financial reward for their participation.

Materials

Fifty line drawings were combined with three Dutch sentences each resulting in 150 experimental stimuli. The sentence correctly described the picture (correct condition), contained a semantically anomalous word (semantic error), or contained a phonological error. The critical words in the phonological error condition were created following actual speech error patterns (Nooteboom, 1969), so-called perseveration errors. All speech errors occurred in sentence-final position. For example, participants saw a line drawing of a twig with a red leaf (see Fig. 1). In the correct condition, they heard the sentence

Correct condition: De tak heeft een rood **blad**.
(The twig has a red leaf.)

Semantic condition: De tak heeft een rood **gebit**.
(The twig has a red denture.)

Perseveration condition: De tak heeft een rood **rad**.
(The twig has a red wheel.)

Fig. 1. Example of the picture stimuli used and the associated sentences.

de tak heeft een rood blad ('the twig has a red leaf'). In the semantic error condition, they heard *de tak heeft een rood gebit* ('the twig has a red denture'), and in the phonological error condition they were presented with *de tak heeft een rood rad* ('the twig has a red wheel'); that is, the critical words in the phonological condition always rhymed with the correct description of the picture.

The length of the sentences ranged from 5 to 9 words with an average of 6.3 words, and all sentences had similar syntactic structures (a subject noun phrase group followed by a verb and an object noun phrase group, e.g. [*de tak*]subject NP [*heeft*]V [*een rood blad*]object NP). The mean frequency of occurrence of the critical words was matched as closely as possible across conditions (correct: 54.2; semantic error: 33.1; phonological error: 41.5 occurrences per million according to the CELEX database; Baayen, Piepenbrock, and Gulikers, 1995), as was their length in number of syllables (correct: 1.2; semantic error: 1.2; phonological error: 1.3).

In addition to the experimental sentences, we used 25 filler sentences all correctly describing their corresponding pictures which were different from the experimental pictures. Filler stimuli were comparable to the experimental stimuli as they had the same syntactic structure and were of similar length (number of words). Fillers were presented four times to match the number of presentations of the experimental pictures resulting in a total of 100 fillers. They were presented once with a grey star integrated in the picture, once with a burst of 4 ms white noise integrated in the sentences, and twice without grey star or white noise. All sentences were recorded with a sampling rate of 44 kHz from a female Dutch native speaker with normal intonation and speaking rate to avoid the risk of artifacts due to artificial speech manipulations.

The 150 experimental sentences used were submitted to a semantic rating test to check whether the semantic violations in the semantic and phonological error conditions differed from each other. A group of 30 native Dutch participants rated the match between the meaning of the sentence and the picture on a 7-point scale (1=very bad; 7=very good). The mean for correct sentences was 6.36, for sentences containing a semantic error 1.58 and for sentences containing a phonological error 2.07. A paired t-test showed that the correct condition differed from both the semantic ($t(29)=37.72$, SD=.69, $p<.01$) and the phonological error condition ($t(29)=26.53$, SD=.89, $p<.01$). The semantic and phonological error conditions differed from each other ($t(29)=5.61$, SD=.47, $p<.01$) although the difference was subtle.

Procedure

Participants were tested individually while seated in a dimly lit, soundproof room in front of a computer screen. Before participants started the experiment proper, they received 15 practice trials that corresponded to the experimental trials but were not used in the experiment proper. Participants were asked to press a button as fast and accurately as possible when they discovered a grey star or heard a short burst of noise. Using this task, semantic and phonological errors were task-irrelevant, but participants were asked whether they noted any speech errors in a debriefing following the experiment.

During each trial, a fixation cross of variable duration (500–800 ms) preceded the stimulus. Pictures and sentences were presented simultaneously, and pictures remained on the screen throughout the sentence presentation. The next trial started 1000 ms after the sentence offset.

Apparatus and recordings

The electroencephalogram (EEG) was recorded from 29 tin electrodes, mounted in an electro-cap according to the extended 10/20 system. The left mastoid was used as on-line reference. Off-line, the EEG signal was re-referenced to the mean of both mastoids. Eye movements were recorded for artifact rejection by electrodes placed at the sub- and supra-orbital ridge of the left eye (vertical movements)

and at the right and left outer canthus (horizontal movements). Signals were digitized at 250 Hz and band-pass filtered from 0.05 to 30 Hz. Impedance for all electrodes was kept below 5 kΩ.

Data analyses

Two time windows and four regions-of-interest (ROI) were defined for the ERP analyses. The first time window ranged from 50 to 150 ms and the second from 300 to 600 ms following the onset of the critical word. The ROIs were defined as follows: anterior: F7, F3, AFZ, F4, F8, FC3, FCZ, FC4; posterior: CP3, CPZ, CP4, P7, P3, PZ, P4, P8; left: F7, F3, FC3, C3, CP3, P7, P3; and right: F4, F8, FC4, C4, CP4, P4, P8. The spatial factors *Anterior–Posterior* (AP) and *Left–Right* (LR) were analyzed in separate analyses in order to include as many electrodes as possible and to keep a symmetrical arrangement. On average, after artifact rejection, 8.1% of the trials were excluded from further analysis (correct condition: 7.7%; semantic error condition: 7.6%; phonological error condition: 9.1%). Average waveforms were computed for all conditions and ROIs separately. For the correct condition, only the first presentation of each stimulus was entered into the averaging procedure. Analyses of variance were performed on the mean area amplitude relative to a pre-stimulus baseline from −200 to 0 ms.

Results

The grand averages for the three conditions and nine electrode sites are shown in Fig. 2. This figure shows a negative deflection for semantic and phonological errors compared to the correct condition. There was no clear difference in amplitude or latency between the semantic and the phonological error condition. Both negativities reached their peak amplitude around 350 ms after the onset of the critical word and the effect was largest over centro-parietal electrodes (see Fig. 3).

Time window 50–150 ms

A 3×2 repeated measures ANOVA was performed with the factors Error Type (semantic error, phonological error, and correct) and Anterior–Posterior (AP) location. No significant effect of Error Type was found ($F(2, 38)=1.64$, MSE=4.80, ns). The interaction between Error Type and AP was not significant, either ($F(2, 38)=1.51$, MSE=2.20, ns). The 3×2 ANOVA with the factors Error Type and Left–Right (LR) location did not yield a significant main effect of Error Type or an interaction with LR (all Fs<1).

Time window 300–600 ms

Again, we performed a 3×2 repeated measures ANOVA with the factors Error Type and AP. The main effect of Error Type and the interaction between Error Type and AP was significant ($F(2, 38)=20.16$, MSE=3.99, $p<.01$; $F(2, 38)=6.64$, MSE=0.80, $p<.01$, respectively).

To follow up the interaction, separate ANOVAs were performed for anterior and posterior ROIs. For the anterior ROI, the main effect of Error Type was significant ($F(2, 38)=10.82$, MSE=2.13, $p<.01$). Subsequent ANOVAs revealed a significant difference between phonological errors and the correct condition ($F(1, 19)=15.16$, MSE=4.82, $p<.01$) and between semantic errors and the correct condition ($F(1, 19)=13.12$, MSE=4.94, $p<.01$). However, the difference between semantic and phonological errors was not significant ($F(1, 19)<1$).

The ANOVA for the posterior ROI showed a significant main effect of Error Type ($F(2, 38)=23.54$, MSE=2.67, $p<.01$). Subsequent ANOVAs yielded a significant difference between semantic errors and the correct condition ($F(1, 19)=37.75$, MSE=5.54, $p<.01$) and between phonological errors and the correct condition ($F(1, 19)=32.52$, MSE=5.07, $p<.01$). The difference between the semantic and phonological errors was not significant ($F(1, 19)<1$). The 3×2 ANOVA with the factors Error Type and LR did not yield an interaction of Error Type with LR (Fs<1).
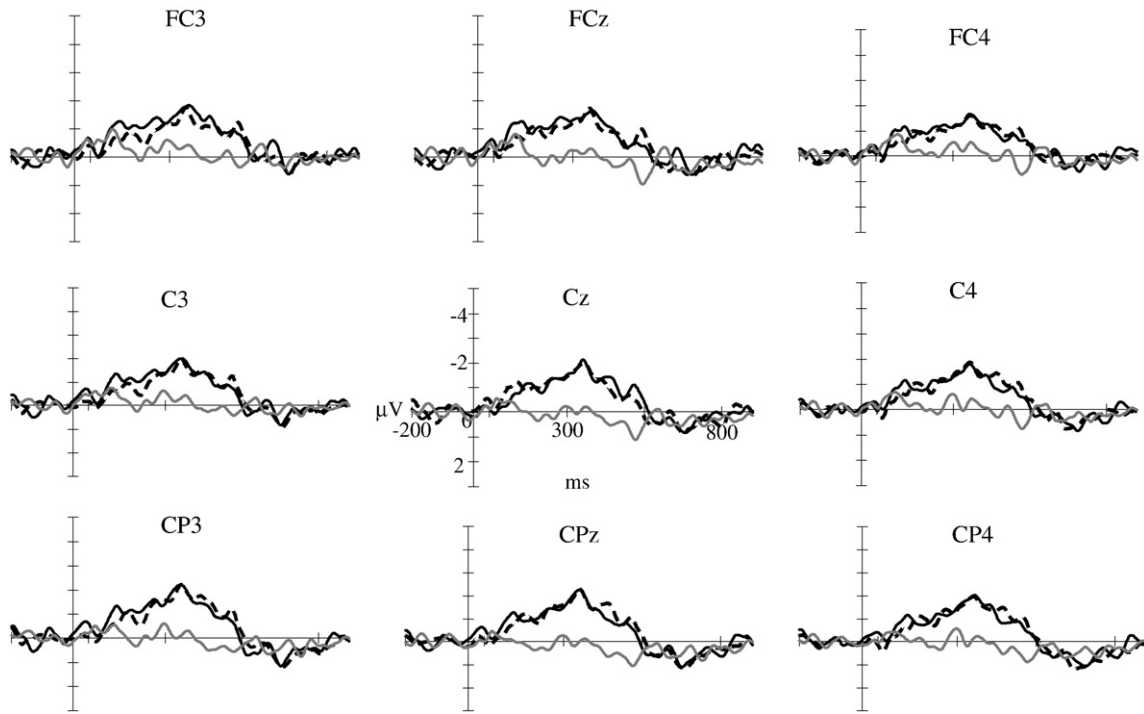
Fig. 2. Grand average ERPs for the three experimental conditions in Experiment 1 (grey line: correct condition; black dashed line: lexical error; black line: perseveration errors) for a subset of nine electrodes. Negative is plotted up in this and all subsequent ERP plots.

Discussion

A similar negativity was observed for semantic and phonological errors compared to the correct condition between 300 and 600 ms. Both negativities showed a centro-parietal maximum peaking around 350 ms following the anomalous word onset. Due to the negative polarity of the effects, and their temporal (300–600 ms) and spatial characteristics (centro-parietal maximum), we interpret these negativities as N400 effects.

Fig. 3. Scalp distribution maps of the difference waves (speech error − correct) for the time windows 50–150 ms (upper panel) and 300–600 ms (lower panel) in Experiment 1.

These results are important for several reasons: first, the differences between the processing of correct and anomalous words demonstrate that auditory speech was processed even though participants were engaged in an unrelated, non-linguistic task with relatively low cognitive processing demands. Whether or not irrelevant speech is similarly processed, i.e. elicits similar N400 effects if the processing demands of the primary task are high, e.g. an n-back task, remains a topic for further inquiry (for a related study see Sabri, Binder, Desai, Medler, Leitl, and Liebenthal, 2008). Second, the N400 effects suggest not only superficial, shallow processing of the speech signal, but instead deeper processing involving semantic and presumably conceptual information. Therefore, it is suggested that the present N400 effects reflect the difficulty of integrating the critical words into the situational context even though the speech was task-irrelevant (Kutas, and Federmeier, 2000). The integration difficulty may also be associated with a decreased semantic expectancy of the critical words in both error conditions (Kutas, and Hillyard, 1984).

Finally, no difference was observed for the processing of semantic and phonological errors. This may have one of the following reasons: (1) these two types of errors may be processed in the same way by our language processing system. (2) It may be the case that a specific ERP signature for phonological errors is only obtained if the speech signal is task-relevant. That is, while (1) suggests that the two types of errors are cognitively processed similarly, (2) conceives of phonological errors as a different error category than semantic errors, which in the current experiment did not lead to different processing signatures due to the specific task instructions. (3) We do not want to exclude the possibility that the approach employed in the first experiment may not have been sensitive enough to reveal the rather subtle differences between semantic and phonological errors.

After the experiment, all participants reported noticing (different kinds of) speech errors in some sentences supporting our interpretation that the task-irrelevant speech was processed. Sometimes a complete word and sometimes only some sounds of a word were reported to be anomalous. This distinction may relate to the small difference in the semantic rating of semantic and phonological errors
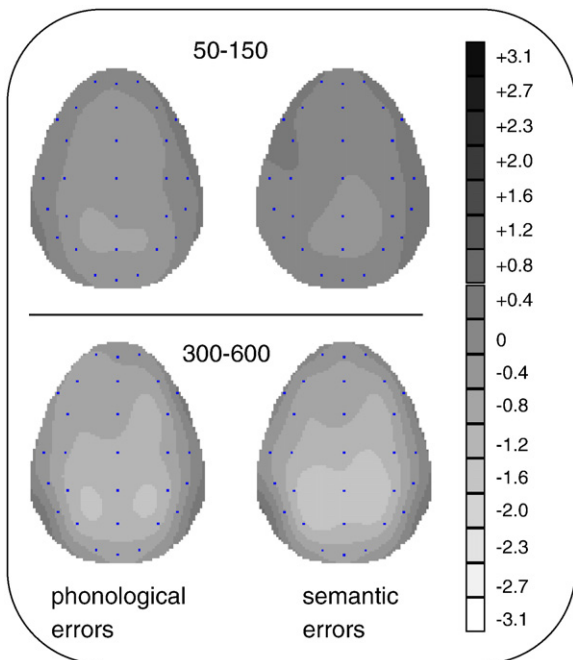
regarding their mismatch with the picture. However, the difference found in the off-line rating task cannot explain our present results because the ERP waveforms did not differ between semantic and phonological errors.

The N400 has been shown to begin around 200 ms after word presentation (Van Petten, Coulson, Rubin, Plante, and Parks, 1999). The onset of the N400 in Experiment 1 is roughly in accordance with the typical N400 onset. The slightly earlier onset (around 150 ms after critical word onset) in the present study might be due to the fact that the experimental situation was of reduced complexity and thus processing might have been relatively easy. For example, the sentences used in Van Petten et al. (1999) comprised on average 12.3 words (range 6–29) while sentences in the present experiment contained on average only 6.3 words (range 5–9).

To investigate whether phonological errors are processed differently from semantic errors only if speech is task-relevant, we carried out a second experiment asking another group of participants to detect anomalous words in the same picture descriptions. This also allows us to show whether or not phonological errors can be distinguished cognitively from semantic errors using ERPs.

Experiment 2: task-relevant monitoring of speech errors

The same experimental design and stimuli were used as in the first experiment. However, participants were instructed to press a button as fast and accurately as possible whenever they detected an anomalous word in the picture descriptions. By hypothesis, this situation makes the speech signal relevant to participants. Possibly, this induces a different processing strategy because, in contrast to irrelevant speech, avoiding processing costs is not helpful for the task at hand.

One way to facilitate the task of detecting anomalous words and communication in general, is by trying to predict upcoming words (see Pickering, and Garrod, 2007 for a review). Evidence for the prediction of words in speech processing comes from different observations. For example, we can often utter target words when our interlocutor pauses or has temporary word finding difficulties. Furthermore, semantic priming and the syntactic garden-path phenomenon can be seen as prediction or anticipation of upcoming words. Similarly, story completion and the cloze test (Taylor, 1953) are taken to reflect word prediction (see Van Berkum, Brown, Zwitserlood, Kooijman, and Hagoort, 2005 for a detailed discussion; but see also Jackendoff, 2002).

It has been suggested that language users not only try to narrow the range of expected words, but can, given sufficient context information, predict specific word forms before the words have been presented (DeLong, Urbach, and Kutas, 2005; Van Berkum et al., 2005; Wicha, Bates, Moreno, and Kutas, 2003). In these studies, the question of whether specific word forms were anticipated on-line was tested during the processing of so-called test words that preceded the anticipated word. Test words and anticipated words were linguistically related (e.g. phonologically or by grammatical gender agreement). In some experimental conditions, the test words violated these relations and the observed ERP effects during test word presentation implied that a specific word form had been predicted. That is, these studies suggest that language users not only try to narrow the range of expected words but they are, in principle, able to predict *specific* word forms.

It is important to differentiate the (semantic) pre-activation of a lexical field and the prediction of a specific lexical entry. The prediction effects in the above-mentioned studies suggest that predicted words included morphosyntactic (e.g. syntactic gender; Van Berkum et al., 2005; Wicha et al., 2003) and word form information (e.g. phonology; DeLong et al., 2005). However, so far it is not clear whether or not listeners generally predict upcoming words as suggested by Pickering and Garrod (2007).

Here, we would like to suggest that semantic and phonological errors are processed differently only in relevant speech because in relevant, but not in irrelevant speech prediction of upcoming words is attempted. This argument is in line with the idea that cognitively costly processes, such as – presumably – prediction, are avoided if unnecessary, i.e. if speech is irrelevant. Given the relatively simple sentences and pictures used here, the object nouns are easy to predict (e.g. see Fig. 1), especially since the syntactic structure of the sentences was very similar across the experiment (all sentences had a subject–verb–object structure where the object was formed by a determiner–adjective–noun phrase). Note that reduced N400 amplitude, as found in Experiment 1, does not imply a specific word form prediction. It indexes only a decreased semantic processing effort, e.g. an easier semantic integration.

Consider how prediction could make word processing more efficient. If a specific word is predicted based on the available context information (i.e. the combination of the auditory description and the picture), the perceived word can be checked for congruence with the predicted word before the former is completely perceived, i.e. before the recognition point of this word. As soon as a phoneme including the very first deviates from the predicted word form, processing can be adapted, e.g. a new, more adequate lexical search might be initiated. Similarly, as long as the incoming auditory word form matches the predicted word form, some processing resources might be spared or used for other (linguistic) processes because no exhaustive lexico-semantic search is necessary. In this case, a specific word candidate has already been predicted, i.e. activated.

More importantly, if prediction is attempted in relevant but not irrelevant speech, different results may be predicted for Experiment 2. If a specific word is predicted, e.g. *fish* in our example above, the initial phoneme should be found to mismatch when *dish* is heard. Such a mismatch is known to elicit a PMN (Connolly et al., 1995; Connolly and Phillips, 1994), and the PMN has been suggested to reflect access to pre-lexical speech segments necessary during phonological analysis and possibly verbal working memory during speech comprehension (D'Arcy et al., 2004).

While the suggested prediction mechanism should lead to a PMN for both errors types, a differential prediction is made for the expected N400 effects. Priming studies suggest that form overlap between prime and target words facilitates lexical access (Deacon, Dynowska, Ritter, and Grose-Fifer, 2004; Swinney, 1979; Zwitserlood, 1996). Such a facilitation effect may be reflected in a reduced N400 (Besson, Kutas, and Van Petten, 1992; Deacon, Hewitt, Yang, and Nagata, 2000; Kiefer, 2002; Praamstra, Meyer, and Levelt, 1994; Rugg, 1990) and reduced reaction times (RTs; Radeau, Morais, and Segui, 1995).

In the present study, the *prediction* of specific upcoming words may facilitate lexical processing. When predicting upcoming words, the lexical entry of the predicted word should be selected and the according word form activated within the production system. Any activated word form (i.e. phonological representation) can be transmitted to the comprehension system via the internal monitoring loop and activate the corresponding form representations in the lexicon, including the corresponding phonemes. If, in close temporal proximity, another word is fed into the comprehension system from the external auditory channel, lexical processing of this latter word may be facilitated due to the previously activated phonemes. That is, facilitated lexical processing may be reflected in a reduced N400, if the predicted and the actually perceived words have some form overlap (i.e. for phonological errors). In contrast, a full N400 effect is expected for lexical errors which are also semantically incongruent but do not have phonological overlap with the predicted words.

In contrast, speech processing could be highly automatic and independent of context, i.e. task relevance. In that case, the same effects are predicted for the processing of irrelevant and relevant speech, i.e. the same N400 effects would be expected for both error

conditions as in Experiment 1. It is also possible that irrelevant speech is processed using the same mechanisms as for relevant speech but the processing is more shallow (Chwilla et al., 1995; Craik and Lockhart, 1972). In that case, both error types are expected to yield similar N400 effects, but the effects should be larger than in Experiment 1.

Method

Participants

Twenty students (19 women) from Maastricht University in the Netherlands took part in Experiment 2 (mean age 20.7 years; range 18–24). One participant was left-handed. Participants were native speakers of Dutch, had normal or corrected-to-normal vision and audition, and received a small financial reward for their participation.

Materials

Pictures and sentences were the same as in the first experiment. However, filler pictures and sentences did not contain grey stars or short noise bursts.

Procedure

The procedure was identical to Experiment 1 with the exception that participants were instructed to press a button as fast and accurately as possible when they discovered an error in the auditory description of the picture, for instance, when the verbal description was not congruent with the picture. Before the experiment proper started, they received 15 practice trials to ensure they understood the task.

Apparatus and recordings

The same apparatus and recordings were used as in Experiment 1.

Data analyses

The data were analyzed in the same way as in Experiment 1. After artifact rejection, on average 4.9% of the trials were excluded from further analysis (correct condition: 4.8%; semantic error condition: 5%; phonological error condition: 5%).

Results

Behavioral data

Only reaction times that were within the range of the participants' mean ± 2 standard deviations (SDs) were analyzed (Ratcliff, 1993). Participants detected phonological errors (566 ms) faster than the semantic errors (586 ms; $t(19) = 2.19$, SD = 41.19, $p < .05$). Since participants made on average only 1.5% errors, no statistical analysis was performed on the error rates.

Electrophysiological data

The grand average waveforms for the three conditions are depicted in Fig. 4 for nine electrode sites. These ERPs show a negative deflection that has a similar magnitude for semantic and phonological errors in the early time window. In the latter time window, both error types show an increased negativity compared to the correct condition although the negativity appears to be larger for semantic than for phonological errors. The topographic maps of the effects for both time windows are shown in Fig. 5.

Time window 50–150 ms

As in the first experiment, we performed a 3 × 2 repeated measures ANOVA with the factors Error Type and AP. The interaction between Error Type and AP was significant ($F(2, 38) = 4.45$, MSE = 1.74, $p < .01$).

To follow up the interaction, two separate ANOVAs were run for the anterior and the posterior ROI. In the anterior ROI, the main effect of Error Type was significant ($F(2, 38) = 4.32$, MSE = 3.36, $p < .05$). Subsequent ANOVAs showed a significant difference between semantic errors and the correct condition ($F(1, 19) = 6.08$, MSE = 6.37, $p < .05$), and between phonological errors and the correct condition ($F(1, 19) = 6.61$, MSE = 7.21, $p < .01$). The difference between semantic and phonological errors did not reach significance ($F(1, 19) < 1$).

The ANOVA for the posterior ROI also yielded a significant main effect of Error Type ($F(2, 38) = 13.05$, MSE = 4.55, $p < .01$). The differences between the semantic errors and the correct condition as well as between phonological errors and the correct condition were significant ($F(1, 19) = 15.90$, MSE = 8.69, $p < .01$; $F(1, 19) = 30.24$, MSE = 6.95, $p < .01$, respectively). The difference between the phonological and semantic
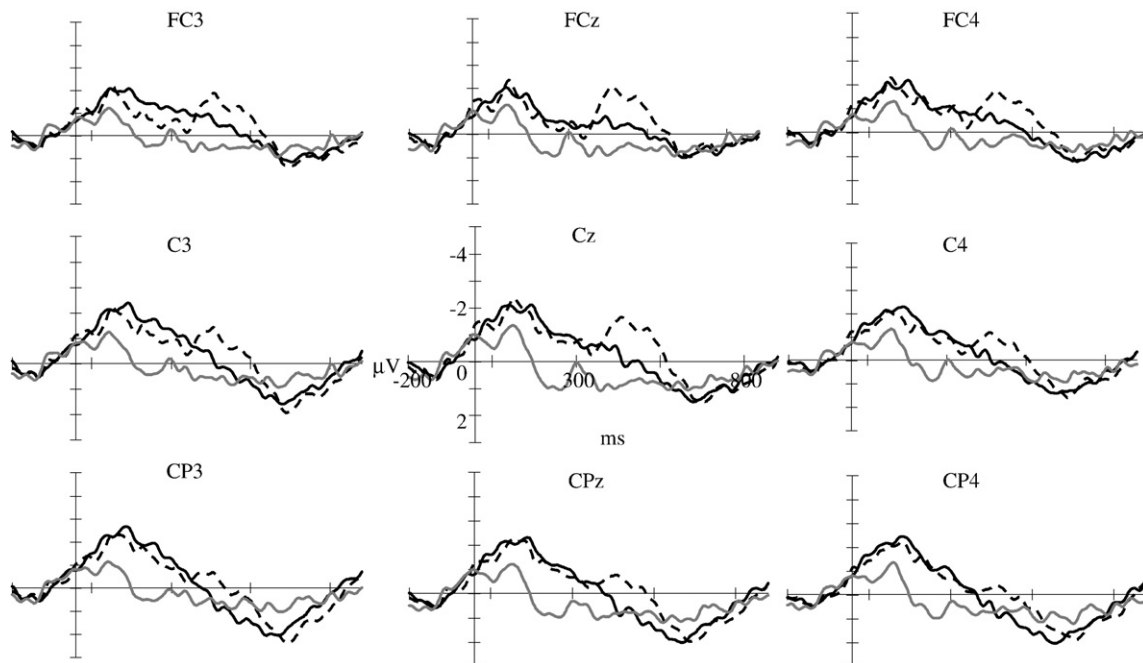

Fig. 4. Grand average ERPs for the three experimental conditions in Experiment 2 (grey line: correct condition; black dashed line: lexical errors; black line: perseveration errors).
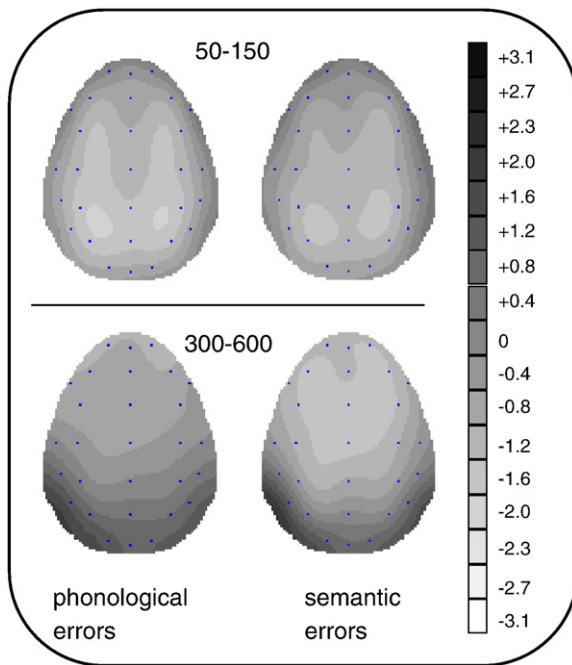
Fig. 5. Scalp distribution maps of the difference waves (speech error − correct) for the time windows 50–150 ms (upper panel) and 300–600 ms (lower panel) in Experiment 2.

errors was not significant (F(1, 19)<1). The 3×2 ANOVA with the factors Error Type and LR orientation did not yield an interaction of Error Type with LR either (F<1). Both Error Types elicited a similar and broadly distributed PMN.

Time window 300–600 ms

A 3×2 repeated measures ANOVA, performed with the factors Error Type and AP, revealed a significant interaction between Error Type and AP (F(2, 38)=22.28, MSE=1.68, p<.01). Separate ANOVAs were run for the anterior and the posterior ROI.

In the anterior ROI, there was a significant main effect of Error Type (F(2, 38)=23.87, MSE=3.22, p<.01). Subsequent ANOVAs showed a significant difference between semantic errors and the correct condition (F(1, 19)=62.86, MSE=4.72, p<.01), and between phonological errors and the correct condition (F(1, 19)=15.79, MSE=8.33, p<.01). The difference between semantic and phonological errors was also significant (F(1, 19)=5.27, MSE=6.28, p<.05).

For the posterior ROI, there was no effect of Error Type (F(1, 19)=2.17, MSE=4.11, ns). In addition, the 3×2 ANOVA with the factors Error Type and LR orientation did not yield an interaction of Error Type with LR (F<1).

The results of Experiment 2 demonstrated a qualitatively different pattern compared to Experiment 1. As this difference is of high theoretical interest, we sought additional statistical validation by analyzing the ERPs of both experiments using the same time windows with Experiment as a between-subjects factor. An interaction involving the factors Experiment and Error Type is expected in each time window if the ERP effects differ between the two experiments.

Combined analysis, time window 50–150 ms

An ANOVA was performed with the within-subjects factors Error Type (3), AP (2), and the between-subjects factor Experiment (2). This ANOVA revealed a significant interaction between Error Type and Experiment (F(2, 76)=5.54, MSE=5.48, p<.01). The three-way interaction was not significant (F<1).

Combined analysis, time window 300–600 ms

The same 3×2×2 ANOVA in the later time window revealed a significant three-way interaction between Error Type, AP, and Experi-

ment (F(2, 76)=27.95, MSE=1.24, p<.001). The interaction between Error Type and Experiment was not significant (F(2, 76)=2.62, MSE=4.82, ns).

Discussion

Detection accuracy was high, suggesting that participants processed the stimuli according to the instructions. Two negativities were observed in the ERP signal. The first negativity was similar for both error types. The second negativity differed in amplitude, i.e. it was larger for semantic than for phonological errors.

Both early negativities started to differ from the correct condition very early, i.e. around 50 ms after critical word onset. These negativities peaked around 100 ms at frontal electrodes and around 150 ms at posterior electrodes. The early negativities were maximal over central electrodes. The amplitude differences were significant in the time window from 50–150 ms. Given our experimental manipulation and the spatio-temporal characteristics of these components, we interpret them both as PMNs.

Previous work suggests that the PMN reflects the access to speech segments of the critical word, i.e. the pre-lexical processing of phonological features (D'Arcy et al., 2004). Increased amplitude of the PMN reflects the phonological mismatch of initial segments compared to an expected candidate word (Connolly et al., 1995; Connolly and Phillips, 1994). In the present experiment, the expected word must have been predicted because the correct sentence endings were not contained in the anomalous picture descriptions. In addition, if the correct sentence endings were not available, no phonological mismatch could have occurred. However, our data demonstrated comparable PMNs for both error types. This suggests the availability of the correct word proposed to be produced tacitly because critical words in both error conditions deviated in their initial phonemes from the correct sentence ending. Thus, we conclude that the correct completion must have been predicted using the linguistic together with the visual context. This prediction is suggested to be done using the language production system (Pickering, and Garrod, 2007) and the internal monitoring loop (see Introduction).

Both PMNs were broadly distributed in the present study. In contrast, earlier studies reported frontal scalp distributions for the PMN. Note, however, that the use of picture–word pairs can lead to a central–parietal scalp distribution of the PMN (Connolly et al. 1995; see above). The more extended distribution into posterior areas could also be due to a number of differences with earlier reports (e.g. language or task differences).

A noteworthy feature of the observed PMNs is their early onset around 50 ms (see D'Arcy et al., 2004). However, in contrast to previous studies that used only sentence materials, we presented additionally pictures that were visible from sentence onset. Therefore, a prediction of upcoming words was relatively easy and the incoming spoken words may have been processed faster. Thus, it is feasible that such faster processing is reflected in the early onset of the PMNs. In addition, effects of lexical access (Marslen-Wilson and Tyler, 1975; Marslen-Wilson and Welsh, 1978) and syntactic processing (Friederici et al., 1993; Neville, Nicol, Barss, Forster, and Garrett, 1991) in language comprehension have been observed to begin within 100–200 ms. Considering that phonological effects should occur prior to lexical access and syntactic processing, the early effects of phonological mismatch reported here become conceivable.

Connolly et al. (1995) did not find such an early PMN onset, although they also presented pictures and spoken words simultaneously. Crucially, our sentence stimuli were longer and therefore provided more processing time for participants to make specific word predictions. This reasoning leads to a testable prediction. If the picture presentation precedes the acoustically presented word to permit tacit picture naming before word onset, a PMN with an earlier onset should be observed. That is, the high predictability and the use of additional

Table 1
Mean amplitude of central electrode sites for the three experimental conditions

Condition	0–50 ms	50–100 ms	100–150 ms	150–200 ms	200–250 ms	250–300 ms	300–350 ms
Correct	−0.27	−0.29	−0.54	0.47	1.09	0.99	0.78
Lexical error	−0.92	−1.22	−1.75	−1.33	−0.87	−0.37	−0.04
Perseveration error	−0.94	−1.58	−1.83	−1.74	−1.08	−0.77	−0.27

non-linguistic context information may have resulted in the early PMN onset in our study.

Subsequent to the PMNs, both error types elicited a second, broadly distributed negativity that peaked shortly after 400 ms, and differed in amplitude from correct sentence endings between 300 and 600 ms. Both negativities displayed a maximum over fronto-central electrodes (see General Discussion) and are interpreted as N400 components. The N400 effect was larger for semantic than for phonological errors. As mentioned before, the N400 is sensitive to lexical and post-lexical processes (Kutas, and Federmeier, 2000; Kutas, and Van Petten, 1994; Van Petten, and Luka, 2006). Importantly, the present N400 effects cannot be explained by the motor responses that accompanied both error types. The N400 effects differed in magnitude although the (button press) responses were the same for both error types, and in Experiment 1, N400 effects were also obtained in the absence of a motor response. Furthermore, the obtained ERP effects are not comparable to known ERP components of motor responses regarding their polarity, magnitude and component shape (Kornhuber, and Deecke, 1965; Walter, Cooper, Aldridge, McCallum, and Winter, 1964).

The N400 effect as found for both error types is proposed to reflect the difficulty of integrating the anomalous word into the context established by the preceding sentence in combination with the picture. The reduced N400 effect for phonological compared to the N400 effect for semantic errors is suggested to reflect facilitated lexical processing of perseveration errors. Such facilitation may result from the phonological overlap between the phonological error and the adequate word if, and only if, the adequate word form has been predicted because it was not contained in the stimulus. Such a prediction may be achieved by tacitly naming the adequate sentence-final word. Subsequently, activated phonological segments may cross-over to the comprehension system via the internal monitoring loop (Özdemir, Roelofs, and Levelt, 2007). This benefit in lexical processing might reflect facilitated lexical access (Besson et al., 1992; Rugg, 1990) or be an instance of rhyme priming (Praamstra et al., 1994; Radeau, Besson, Fonteneau, and Castro, 1998).

Lexical access may be facilitated because activated phonological segments due to cross-over from the (internal) production system via the internal monitoring loop may reduce the selection threshold for phonologically anomalous words. Although the behavioral responses for both error types occurred relatively late within the time window of the N400 effects (586 ms and 566 ms), they may nevertheless be influenced by the cognitive processes reflected in the early phase of the N400 component. Furthermore, the behavioral responses may in part have been based on the initial phonological mismatch between the perceived and the predicted word.

It may be argued that lexical access must precede lexico-semantic integration and therefore these processes should not affect the ERP in the same time window. However, the onset of ERP effects provides an estimate of the upper time limit of when a cognitive process begins (Rugg, and Coles, 1995). That is, lexical access and integration may begin at different times, but still affect the ERP during the same N400 time window.

Alternatively, the facilitated lexical processing may be an instance of rhyme priming. In word–word priming studies, a phonological overlap in the rhyme led to a reduced N400 magnitude (Praamstra et al., 1994; Radeau et al., 1995, 1998; but see Van Petten et al., 1999). These studies are not directly comparable to our experiments

because we used sentences as picture descriptions and the rhyming word was not contained in the stimulus. Nevertheless, it is interesting to note that the reduced N400 effects were obtained without a phonological judgment task. Recently, such a phonological component of the task has been discussed as a prerequisite for an N400 rhyme priming effect (Perrin, and García-Larrea, 2003) although such effects have been reported using a lexical decision task (Praamstra et al., 1994; Radeau et al., 1998). Based on the present data, we cannot decide whether the facilitated lexical processing reflects a benefit in lexical access or rhyme priming. Future research is necessary to disentangle these alternative processing mechanisms. Importantly, both mechanisms are functional only if the adequate word has been activated including its word form representation because it was not part of the verbal stimulus.

Phonological errors were detected faster than semantic errors. This outcome is in line with a result obtained by Oomen and Postma (2002) who also found that participants were faster in detecting phonological compared to semantic speech errors. The different RTs for phonological and semantic errors are in accordance with the respective, differential N400 effects and converging evidence for facilitated lexical processing of phonological errors if speech is relevant.

In contrast to the RTs, no latency differences were obtained in the ERP measures. The amplitude values were largest for our experimental conditions between 100 and 150 ms after word onset (see Table 1). RTs and ERPs may reflect different processes because RTs are assumed to reflect the summation of sensory, cognitive, and motor execution processes, whereas ERPs provide a more continuous measure of electrical brain activity (e.g. Holcomb, 1993). We suggest that the RT difference between semantic and phonological errors originates at later processing stages than phonological processing (lexical or post-lexical).

As mentioned before, the semantic rating test showed that semantic-error sentences matched the pictures less than phonological-error sentences. Importantly, the difference between either of the two error types and the adequate sentences was almost ten times as large as the difference between both error types. Thus, the difference in the semantic rating between semantic and phonological errors cannot fully explain the differential N400 effects. That is, the difference in the semantic rating seems to play only a minor role during on-line processing.

General discussion

This paper investigated the monitoring of semantic and phonological errors that were task-irrelevant (Experiment 1) or relevant (Experiment 2). During irrelevant speech, anomalous words elicited similar N400 effects for semantic and phonological errors. In contrast, during task-relevant speech, similar early PMNs were elicited for both error types that were followed by N400 effects of different amplitude (enlarged for semantic errors as compared to phonological). The qualitatively different ERP patterns were supported by the interaction involving Error Type and Experiment in the combined statistical analysis of both experiments (see above). The data pattern suggested that task-irrelevant speech is processed semantically. In contrast, for task-relevant speech, prediction of upcoming words appears to be attempted, and the predicted word form may facilitate the comprehension of phonological errors via the internal monitoring loop.

The PMNs observed in Experiment 2 suggest that participants have predicted sentence-final words on the basis of both the picture and the sentence. Presumably, the PMN reflects the phonological mismatch between the onset of the perceived, anomalous word and the predicted, adequate word form. Neither the sentence nor the picture alone allows the listener to predict the exact sentence-final word. While the picture alone provides information about the content of a potential description, several different sentences including different syntactic structures are possible. The PMNs and the implied prediction mechanism are in accordance with previous work suggesting a prediction mechanism on the basis of verbal material alone (DeLong et al., 2005; Van Berkum et al., 2005; Wicha et al., 2003).[1]

We assume that prediction is avoided when speech was task-irrelevant due to the associated processing costs. Accordingly, no PMN was observed in Experiment 1, and it is unlikely that this null result reflects a sensitivity issue because reliable ERP effects were obtained in both experiments. The absence of a PMN in Experiment 1 and the presence of a PMN in Experiment 2 suggest that prediction is a rather strategic process employed only when speech is relevant.

Regarding the N400 effects, a different pattern was observed. Both experiments yielded N400 effects that are interpreted to reflect post-lexical integration processes, in particular the difficulty of integrating sentence-final words into the preceding context (Friederici, 2002; Hagoort, Hald, Bastiaansen, and Petersson, 2004; Kutas, and Van Petten, 1994). The different magnitude of the N400 between semantic and phonological errors in Experiment 2 suggests an additional difference in lexical processing (Besson et al., 1992; Praamstra et al., 1994; Radeau et al., 1995, 1998; Rugg, 1990; Van Petten, and Luka, 2006). Presumably, lexical processing is facilitated for phonological but not for semantic errors because only phonological errors have some form overlap with the predicted (adequate) sentence ending. The adequate word for a given anomalous picture description is suggested to be produced tacitly and made available to the comprehension system via the internal monitoring loop. Here, the form overlap between the predicted adequate word and the perceived phonological error can facilitate lexical processing. The prediction cannot facilitate lexical processing for semantic errors because they have no form overlap with the predicted word. In accordance with the proposal that prediction is avoided during task-irrelevant speech perception, no such modulation of the N400 effect was observed in Experiment 1. The fact that participants reported on different kinds of errors in Experiment 1 suggests that this distinction was not made on-line but rather during later processing stages, possibly reflecting meta-linguistic decisions.

One might argue that the reduced N400 for phonological errors is simply due to the repeated onset of the influencing syllable (e.g. *r*ood *r*ad) and in that sense truly a perseveration (Boomer and Laver, 1968; Fromkin, 1971). For semantic errors, the onset did not share the onset with the preceding word (e.g. *r*ood *gebit*) and were also semantically anomalous regarding the context. Thus, a slightly reduced N400 might be expected for phonological errors compared to semantic errors (Radeau et al., 1998). Such a mechanism, however, does not seem to be correct because this account assumes that the differential N400 effect is based on sensory information processing. If true, the same effect would have been expected for the first experiment but was not observed. Also, there should be no PMN for phonological errors because the onset (of the influencing syllable) is repeated and does not mismatch during phonological processing. Therefore, we consider the alternative account for phonological errors inadequate for the present results.

Although the N400 is sensitive to the semantic relation (or cloze probability) between words (Kutas, and Hillyard, 1980; 1984), the pattern of N400 effects obtained in both experiments cannot be explained by the semantic relations within the spoken sentences alone. If the N400 effects were solely due to the semantic relations within the spoken sentence, the same ERP effects should be obtained in both experiments because the critical sentences were identical. However, this was not the case suggesting that the processing of the stimuli changed under different task instructions.

In the second experiment, the N400 effects had a fronto-central maximum, which is not unexpected. In studies investigating picture priming, an N300 effect with a frontal scalp distribution was found (e.g. Barrett, and Rugg, 1990; Holcomb, and McPherson, 1994). The N300 is more negative for unrelated compared to related pictures. Similarly, studies investigating the integration of pictorial stimuli (e.g. pictures, line drawings, or gestures) into a context have reported a more frontal scalp distribution for the N400 effects (e.g. Federmeier, and Kutas, 2001; Ganis, Kutas, and Sereno, 1996; Holle, and Gunter, 2007; Willems et al., 2007). Here, it is suggested that in Experiment 2 the additional attempt to predict upcoming words involves (parts of) the language production system (Pickering, and Garrod, 2007). Language comprehension and production are associated with activity in the (pre-) frontal cortex, especially when semantic information has to be judged (Bookheimer, 2002; Gernsbacher, and Kaschak, 2003), which may have resulted in a shift of the N400 effects towards anterior electrodes in Experiment 2 as opposed to Experiment 1. Further research is needed to confirm this hypothesis.

The observation of N400 effects in Experiment 1 contrasts with some previous work (Chwilla et al., 1995). These authors observed no N400 effect in a semantic priming paradigm when the lexical–semantic content was irrelevant to the task (case judgment). The diverging results can be explained by the different paradigms, tasks, modalities, and the fact that we did not use single words but sentences. These observations indicate that further research is necessary to determine the precise conditions under which an N400 can be elicited in task-irrelevant language processing.

After all, our pattern of results suggests that semantic and phonological errors are not processed in fundamentally different ways. Both types of errors elicited the same ERP pattern, i.e. the reduced N400 effect for phonological errors is presumably due to the partial form overlap with the predicted word. That is, semantic and phonological errors involve the same processing mechanisms, but phonological errors benefit from facilitated lexical processing if the correct word, that overlaps partially in its form, is predicted. Whether or not the correct word form can be predicted depends on whether or not the speech is relevant and whether or not the context information is sufficient to make a prediction.

Recently, semantic context has been proposed to be an important variable that can affect the early processing of words in sentences. For example, semantic integration of a critical word has been suggested to begin before the isolation point of the word (Van den Brink, Brown, and Hagoort, 2006; Van Petten et al., 1999). Acoustically presented words may be processed differently depending on the semantic coherence of the context, for instance, whether they are embedded in sentences or word lists (Diaz, and Swaab, 2007). When manipulating semantic and phonological relations, Diaz and Swaab (2007) obtained N400 effects in sentences for both relations, whereas in word lists only the semantic manipulation resulted in an N400 effect. The phonological manipulation (onset consistency of the last word in the list) yielded a frontal positivity and an occipital negativity. It was argued that semantic coherence of the context has a very early influence on lexical processing. Although the present experiments are not exactly parallel to these previous studies, it is instructive to see that the ERP patterns in Experiments 1 and 2 differ qualitatively while the contextual setting, i.e. the task varies between them. Presumably, the manipulation of the speech relevance is part of the context that

---

[1] Work by West and Holcomb (2002; see also Gunter, & Bach, 2004) suggests that meaningful prediction can also be achieved by pictorial stimuli without verbal material.

influences the processing of speech in general and very early during pre-lexical phonological processing stages.

In conclusion, the present study strongly suggests that natural speech is processed semantically even if it is not task-relevant. Our results also suggest that listeners attempt to predict upcoming words if speech is relevant. This apparently controlled process of prediction seems to facilitate comprehension and communication in general. Furthermore, our data suggest that the phonologically deviant condition investigated here is not processed in a fundamentally different way from the semantically deviant condition. Our findings, thus, underline the importance of the internal monitoring loop. Future research needs to determine the scope of cognitive control over the processing of irrelevant speech.

Acknowledgments

References

Baayen, R.H., Piepenbrock, R., Gulikers, L., 1995. The CELEX lexical database (CD-ROM), LDC. University of Pennsylvania, Philadelphia, PA.

Barrett, S.E., Rugg, M.D., 1990. Event-related potentials and the semantic matching of pictures. Brain Cogn. 14, 201–212.

Besson, M., Kutas, M., Van Petten, C., 1992. An event-related potential (ERP) analysis of semantic congruity and repetition effects in sentences. J. Cogn. Neurosci. 4, 132–149.

Bookheimer, S., 2002. Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. Annu. Rev. Neurosci. 25, 151–188.

Boomer, D.S., Laver, J.D., 1968. Slips of the tongue. Br. J. Disord. Commun. 3, 2–12.

Christoffels, I.K., Formisano, E., Schiller, N.O., 2007. Neural correlates of verbal feedback processing: an fMRI study employing overt speech. Hum. Brain Mapp. 28, 868–879.

Chwilla, D.J., Brown, C.M., Hagoort, P., 1995. The N400 as a function of the level of processing. Psychophysiology 32, 274–285.

Connolly, J.F., Phillips, N.A., 1994. Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. J. Cogn. Neurosci. 6, 256–266.

Connolly, J.F., Byrne, J.M., Dywan, C.A., 1995. Assessing adult receptive vocabulary with event-related potentials: an investigation of cross-modal and cross-form priming. J. Clin. Exp. Neuropsychol. 17, 548–565.

Craik, F.I.M., Lockhart, R.S., 1972. Levels of processing: a framework for memory research. J. Verbal Learn. Verbal Behav. 11, 671–684.

D'Arcy, R.C.N., Connolly, J.F., Service, E., Hawco, C.S., Houlihan, M.E., 2004. Separating phonological and semantic processing in auditory sentence processing: a high-resolution event-related brain potential study. Hum. Brain Mapp. 22, 40–51.

Deacon, D., Hewitt, S., Yang, C.-M., Nagata, M., 2000. Event-related potential indices of semantic priming using masked and unmasked words: evidence that the N400 does not reflect a post-lexical process. Cogn. Brain Res. 9, 137–146.

Deacon, D., Dynowska, A., Ritter, W., Grose-Fifer, J., 2004. Repetition and semantic priming of nonwords: implications for theories of N400 and word recognition. Psychophysiology 41, 60–74.

DeLong, K.A., Urbach, T.P., Kutas, M., 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. Nat. Neurosci. 8, 1117–1121.

Diaz, M.T., Swaab, T.Y., 2007. Electrophysiological differentiation of phonological and semantic integration in word and sentence contexts. Brain Res. 1146, 85–100.

Federmeier, K.D., Kutas, M., 2001. Meaning and modality: influences of context, semantic memory organization, and perceptual predictability on picture processing. J. Exper. Psychol. Learn. Mem. Cogn. 27, 202–224.

Friederici, A.D., 2002. Towards a neural basis of auditory sentence processing. Trends Cogn. Sci. 6, 78–84.

Friederici, A.D., Pfeifer, E., Hahne, A., 1993. Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. Cogn. Brain Res. 1, 183–192.

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

Ganis, G., Kutas, M., Sereno, M.I., 1996. The search for "common sense": an electrophysiological study of the comprehension of words and pictures in reading. J. Cogn. Neurosci. 8, 89–106.

Garnes, S., Bond, Z.S., 1980. A slip of the ear: a snip of the ear? A slip of the year? In: Fromkin, V. (Ed.), Errors in Linguistic Performance: Slips of the Tongue, Ear, Pen and Hand. Academic Press, New York, pp. 231–239.

Gernsbacher, M.A., Kaschak, M.P., 2003. Neuroimaging studies of language production and comprehension. Annu. Rev. Psychol. 54, 91–114.

Gunter, T.C., Bach, P., 2004. Communicating hands: ERPs elicited by meaningful symbolic hand postures. Neurosci. Lett. 372, 52–56.

Hagoort, P., Brown, C.M., Groothusen, J., 1993. The syntactic positive shift (SPS) as a measure of syntactic processing. Lang. Cogn. Processes 8, 439–483.

Hagoort, P., Hald, L., Bastiaansen, M., Petersson, K.M., 2004. Integration of word meaning and world knowledge in language comprehension. Science 304, 438–441.

Hartsuiker, R.J., Kolk, H.H.J., 2001. Error monitoring in speech production: a computational test of the perceptual loop theory. Cogn. Psychol. 42, 113–157.

Holcomb, P.J., 1993. Semantic priming and stimulus degradation: implications for the role of the N400 in language processing. Psychophysiology 30, 47–61.

Holcomb, P.J., McPherson, W.B., 1994. Event-related brain potentials reflect semantic priming in an object decision task. Brain Cogn. 24, 259–276.

Holle, H., Gunter, T.C., 2007. The role of iconic gestures in speech disambiguation: ERP evidence. J. Cogn. Neurosci. 19, 1175–1192.

Jackendoff, R., 2002. Foundations of language: brain, meaning, grammar, evolution. Oxford Univ. Press, Oxford.

Kiefer, M., 2002. The N400 is modulated by unconsciously perceived masked words: further evidence for an automatic spreading activation account of N400 priming effects. Cogn. Brain Res. 13, 27–39.

Koester, D., Gunter, Th.C., Wagner, S., Friederici, A.D., 2004. Morphosyntax, prosody, and linking elements: the auditory processing of German nominal compounds. J. Cogn. Neurosci. 16, 1647–1668.

Koester, D., Gunter, T.C., Wagner, S., 2007. The morphosyntactic decomposition and semantic composition of German compound words investigated by ERPs. Brain Lang. 102, 64–79.

Kolk, H.H., Chwilla, D.J., Van Herten, M., Oor, P.J., 2003. Structure and limited capacity in verbal working memory: a study with event-related potentials. Brain Lang. 85, 1–36.

Kornhuber, H.H., Deecke, L., 1965. Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. Pflügers Archiv. 284, 1–17.

Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic anomaly. Science 207, 203–205.

Kutas, M., Hillyard, S.A., 1984. Brain potentials during reading reflect word expectancy and semantic association. Nature 307, 161–163.

Kutas, M., Van Petten, C.K., 1994. Psycholinguistics electrified: event-related brain potential investigations. In: Gernsbacher, M.A. (Ed.), Handbook of Psycholinguistics. Academic Press, San Diego, CA, pp. 83–143.

Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. Trends Cogn. Sci. 4, 463–470.

Kuperberg, G.R., 2007. Neural mechanisms of language comprehension: challenges to syntax. Brain Res. 1146, 23–49.

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., 1999. A theory of lexical access in speech production. Behav. Brain Sci. 22, 1–75.

Marslen-Wilson, W.D., Tyler, L.K., 1975. Processing structure of sentence perception. Nature 257, 784–786.

Marslen-Wilson, W.D., Welsh, A., 1978. Processing interactions and lexical access during word recognition in continuous speech. Cogn. Psychol. 10, 29–63.

Martin-Loeches, M., Schweinberger, S.R., Sommer, W., 1997. The phonological loop model of working memory: an ERP study of irrelevant speech and phonological similarity effects. Mem. Cogn. 25, 471–483.

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. J. Verbal Learn. Verbal Behav. 21, 578–594.

Neville, H., Nicol, J.L., Barss, A., Forster, K.I., Garrett, M.F., 1991. Syntactically based sentence processing classes: evidence from event-related brain potentials. J. Cogn. Neurosci. 3, 151–165.

Nooteboom, S.G., 1969. The tongue slips into patterns. In: Sciarone, A., van Essen, A., van Raad, A. (Eds.), Leiden studies in linguistics and phonetics. Mouton, The Hague, pp. 114–132.

Oomen, C.C.E., Postma, A., 2002. Limitations in processing resources and speech monitoring. Lang. Cogn. Processes 17, 163–184.

Özdemir, R., Roelofs, A., Levelt, W.J.M., 2007. Perceptual uniqueness point effects in monitoring internal speech. Cognition 105, 457–465.

Perrin, F., García-Larrea, L., 2003. Modulation of the N400 potential during auditory phonological/semantic interaction. Cogn. Brain Res. 17, 36–47.

Pickering, M.J., Garrod, S., 2007. Do people use language production to make predictions during comprehension? Trends Cogn. Sci. 11, 105–110.

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

Praamstra, P., Meyer, A.S., Levelt, W.J.M., 1994. Neurophysiological manifestations of phonological processing: latency variation of a negative ERP component time-locked to phonological mismatch. J. Cogn. Neurosci. 6, 204–219.

Radeau, M., Morais, J., Segui, J., 1995. Phonological priming between monosyllabic spoken words. J. Exp. Psychol. Hum. Percept. Perform. 21, 1297–1311.

Radeau, M., Besson, M., Fonteneau, E., Castro, S.L., 1998. Semantic, repetition and rime

priming between spoken words: behavioral and electrophysiological evidence. Biol. Psychol. 48, 183–204.

Ratcliff, R., 1993. Methods for dealing with reaction time outliers. Psychol. Bull. 114, 510–532.

Rugg, M.D., 1990. Event-related brain potentials dissociate repetition effects of high- and low-frequency words. Mem. Cogn. 18, 367–379.

Rugg, M.D., Coles, M.G.H., 1995. The ERP and cognitive psychology: conceptual issues. In: Coles, M.G.H., Rugg, M.D. (Eds.), Electrophysiology of Mind: Event-Related Brain Potentials and Cognition. Oxford University Press, Oxford, pp. 27–39.

Sabri, M., Binder, J.R., Desai, R., Medler, D.A., Leitl, M.D., Liebenthal, E., 2008. Attentional and linguistic interactions in speech perception. NeuroImage 39, 1444–1456.

Schiller, N.O., Meyer, A.S., 2003. Phonetics and Phonology in Language Comprehension and Production. Differences and Similarities. Mouton de Gruyter, Berlin.

Swinney, D.A., 1979. Lexical access during sentence comprehension: (re)consideration of context effects. J. Verbal Learn. Verbal Behav. 18, 645–659.

Taylor, W., 1953. 'Cloze' procedure: a new tool for measuring readability. J. Quart. 30, 415–433.

Van Berkum, J.J.A., Brown, C.M., Zwitserlood, P., Kooijman, V., Hagoort, P., 2005. Anticipating upcoming words in discourse: evidence from ERPs and reading times. J. Exper. Psychol. Learn. Mem. Cogn. 31, 443–467.

Van den Brink, D., Brown, C.M., Hagoort, P., 2006. The cascaded nature of lexical selection and integration in auditory sentence processing. J. Exper. Psychol. Learn. Mem. Cogn. 32, 364–372.

Van Petten, C., Luka, B.J., 2006. Neural localization of semantic context effects in electromagnetic and hemodynamic studies. Brain Lang. 97, 279–293.

Van Petten, C., Coulson, S., Rubin, S., Plante, S., Parks, M., 1999. Time course of word identification and semantic integration in spoken language. J. Exper. Psychol. Learn. Mem. Cogn. 25, 394–417.

Walter, W.G., Cooper, R., Aldridge, V.J., McCallum, W.C., Winter, A.L., 1964. Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. Nature 203, 380–384.

Wicha, N.Y.Y., Bates, E.A., Moreno, E.M., Kutas, M., 2003. Potato not Pope: human brain potentials to gender expectation and agreement in Spanish spoken sentences. Neurosci. Lett. 346, 165–168.

West, W.C., Holcomb, P.J., 2002. Event-related potentials during discourse-level semantic integration of complex pictures. Cogn. Brain Res. 13, 363–375.

Willems, R.M., Özyürek, A., Hagoort, P., 2007. When language meets action: the neural integration of gesture and speech. Cereb. Cortex 17, 2322–2333.

Zwitserlood, P., 1996. Form priming. Lang. Cogn. Processes 11, 589–596.