Research Report

Monitoring in language perception: The effect of misspellings of words in highly constrained sentences

Constance Th.W.M. Vissers*, Dorothee J. Chwilla, Herman H.J. Kolk

Nijmegen Institute for Cognition and Information, University of Nijmegen, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Accepted 4 May 2006
Available online 14 July 2006

Keywords:
Language
Sentence processing
Orthography
Phonology
High/low-cloze context
Event-related potential
Functional significance of the P600

ABSTRACT

We present evidence for a monitoring process in language perception at the word level, reflected by a P600. This P600 is triggered when a conflict evolves because the brain encounters an unexpected linguistic item when another item is highly expected. To resolve this conflict between representations, the brain monitors the input to check for possible processing errors. A P600 was hypothesized to occur after orthographic anomalies, like pseudohomophones, in particular when the word from which the pseudohomophone is derived is highly expected. This hypothesis was tested by recording ERPs while participants read high-cloze sentences (’In that library the pupils borrow books …’) and low-cloze sentences (’The pillows are stuffed with books …’). In a pretest, the high-cloze sentences were produced by more than 90% of the subjects, while the low-cloze sentences were never produced. In half of the sentences, the critical word books was replaced by a pseudohomophone (e.g., bouks), which in the high-cloze sentences orthographically and phonologically resembles the highly expected word. Consistent with the monitoring hypothesis, only pseudohomophones in high-cloze sentences elicited a widely distributed P600 effect while pseudohomophones in low-cloze sentences did not. A standard N400 effect of cloze probability occurred both for words and pseudohomophones.

The present ERP results support the view that there is a process of monitoring that takes place in language perception which is reflected by the P600. It occurs whenever a conflict between a strong tendency to accept and one to reject a word brings the cognitive system in state of indecision.

1. Introduction

Monitoring refers to a process that evaluates the appropriateness or correctness of ongoing motor activity or response output. It is a function of cognitive control aimed at output optimization: to bring erroneous behavior in line with desired goals (e.g., Botvinick et al., 2001; Postma, 2000).

In the language domain, monitoring can manifest itself in the phenomenon of self-repair in speech. In ‘overt’ self-repairs, speech is interrupted and a new attempt is made at producing the correct form (e.g., ’I thought she …I thought he was looking at me). Levetel (1983) argues that, in addition to overt repairs, there are also ‘covert’ repairs in which errors are intercepted at the level of planning by an inner monitoring mechanism. This inner monitoring mechanism hence operates on a prearticulatory representation of the utterance (prearticulatory editing). Covert repairs manifest themselves as various speech disfluencies such as prolongations or pauses (’I thought… I thought he was looking at me). An important argument for the existence of prearticulatory editing is that...
sometimes repair occurs after just one phoneme has been produced. Considering the early moment of these repairs, it seems improbable that errors are always detected by the speaker while listening to her own overt speech. The dominant theory of error monitoring in speech is that the prearticulatory and postarticulatory editing in speech is accomplished by the language comprehension system (Levelt, 1983). In other words, we use the same mechanism for comprehending speech and for monitoring our own speech. Hence, Levelt’s ”perceptual loop theory” localizes monitoring in the perception apparatus. Recently, Hartsuiker and Kolk (2001) have provided computational evidence for this theory.

In the action domain, studies of event-related brain potentials (ERPs) have revealed a brain response following errors: the error-related negativity (ERN), typically occurring around 100 ms after an error (for a review see Yeung et al., 2004). ERN activity is not only observed after errors in choice reaction time tasks but also when participants are told that an error occurred, whether this was true or not. If an overt error is not necessary for an ERN to occur, what is it that elicits the ERN? Recent theories suggest that it is the conflict between two representations that triggers the ERN. Recent evidence provides support for the hypothesis that this monitoring process is generated in the anterior cingulate cortex (ACC).

Monitoring has been studied primarily in production tasks. However, besides errors in production, we also make errors in perception (e.g., misreading a word) or comprehension (e.g., misunderstanding a speaker) and it seems likely that these errors are also monitored for. A monitoring process at the sentence perception level has been described by Kolk et al. (2003). They assume that, in sentence perception, simple processing heuristics are used in addition to syntactic algorithms. These heuristics are rules of thumb: highly economical strategies that are generally but not always effective in extracting meaning. The syntactic algorithm on the other hand involves an algorithmic analysis of the syntactic structure of the sentence; this analysis is time-consuming but always comes up with the correct sentence interpretation.

Ferreira (2003) has recently caught up on the discussion of the use of simple processing heuristics in language comprehension. She tested whether the participants’ performance on deciding on the thematic roles in sentences that varied in plausibility (plausible: the dog bit the man vs. implausible: the man bit the dog) and in reversibility (reversible: the dog bit the man vs. nonreversible: the mouse ate the cheese) could be modeled by the use of two simple heuristics. One heuristic is the NVN strategy; that is, the processor assumes that the subject is a proto-agent and the object is a proto-patient. The second heuristic is the plausibility heuristic which states that the processor assumes the semantic analysis which is most consistent with world knowledge. It thus combines lexical items of a sentence in the most plausible way (see also, Ferreira et al., 2002). Although the NVN strategy was the best predictor, the combined use of both strategies could mimic the participants’ performance even better. The study thus provides evidence for the use of both the NVN and the plausibility strategies in normal speakers.

As Ferreira points out, it is presently unknown how the product of the heuristics is coordinated with the output of the syntactic algorithms. Perhaps heuristics are employed when algorithms are hard to apply, given the complexity of the sentence. Alternatively, algorithms may be used when comprehenders have little confidence in the outcome of a heuristic. Although these possibilities exist, it seems simpler and more straightforward to assume that the two routes operate in parallel and largely independent from one another. Kolk et al. (2003) and Van Herten et al. (in press) argue for the latter possibility and point to the dual route model of reading aloud, in which there is similar parallel processing along two independent routes (e.g., Coltheart et al., 1993). But if there is such parallel processing in the case of sentence interpretation, it is possible that the two routes lead to conflicting outcomes. Kolk et al. (2003) propose that such a conflict triggers a monitoring process and that it is this monitoring process which underlies their ERP findings. In particular, they used semantic reversal anomalies, which were formed by exchanging the subject and object of semantically acceptable sentences such as (1).

(1) De vos die op de stropers jog sloop door het bos (original).

The fox that at the poachers hunted [singular] stalked through the woods (literal translation).
The fox that hunted [singular] the poachers stalked through the woods (paraphrase).

It is clear that in these sentences that plausibility heuristics and syntactic algorithms produce different thematic interpretations. Whereas the plausibility heuristic (a lexical strategy) leads to the interpretation that poachers are hunting foxes, the parsing routines lead to the interpretation that the foxes are hunting the poachers. Although the latter interpretation is not entirely impossible, it represents a highly unlikely event based on world knowledge.

Kolk et al. (2003) observed a P600 effect to these semantic reversal anomalies, and not an N400 effect—as would have been expected, given that semantic anomalies typically elicit an N400 effect. This is consistent with recent findings from other researchers who, despite differences in sentence material and language (English and Dutch), observed a P600 effect in the absence of an N400 effect to semantically implausible sentences relative to their plausible counterparts (Hoeks et al., 2004; Kim and Osterhout, 2005; Kuperberg et al., 2003, 2006; Van Herten et al., 2005). These results seem to challenge the view that P600 effects are primarily elicited by syntactic anomalies. How does the occurrence of the P600 relate to the presence of a conflict between algorithmic and heuristic processing routes?

The P600 effect is assumed to reflect an immediate consequence of the situation that the parse and the plausibility heuristic suggest different interpretations. Kolk et al. (2003) argued that the language comprehension system attempts to resolve the resulting conflict by checking the input for possible processing errors. In particular, the mismatch between the semantically plausible, highly expected (based on world knowledge) thematic
interpretation and the implausible thematic interpretation makes it necessary for the brain to reattend the unexpected linguistic unit to check upon its veridicality. After all, an inconsistency can have two sources. It can be real, in the sense that an unexpected event has indeed occurred (e.g., man bites dog). On the other hand, it can also stem from a processing error. To prevent integration of erroneous information into the current discourse, the reader will generally monitor the correctness of his or her analysis in case of a conflict. This explains the occurrence of P600 effects to semantic reversal anomalies.

Kolk et al. (2003) proposed that the absence of an N400 effect was due to the fact that the ‘lexical’ interpretation for both the plausible and implausible sentences leads to an interpretation that is plausible. Since the plausibility heuristic does not have difficulty integrating the words of the semantic reversals into a coherent message (the lexical items in both conditions are the same; the fox and the hunters), readers initially do not notice the anomaly. Hence, no N400 effect was elicited. It thus seems that a conflict at the sentence level triggers the monitoring process in perception, just as in production. This conflict can probably best be described as one between different tendencies: the tendency to reject and the tendency to accept the sentence.

Thus, we propose that if language perception leads to the activation of two incompatible interpretations, a conflict would arise, signaling the possibility of a processing error. Such a conflict could trigger a monitoring response and hence a P600, to check for the possibility of a processing error. Garden path sentences are one example of a situation known to elicit P600 effects (Osterhout and Holcomb, 1992) which can be characterized as representing some kind of conflict between response tendencies. In garden path sentences, two different analyses of the same linguistic string lead to the activation of two incompatible responses. After reading the sentence ‘The women persuaded to answer the door’, initially one interpretation is chosen, but has to be replaced by a different interpretation later on. At first, readers assume that the sentence is about women persuading someone, but after reading the sentence part following the verb, they realize that the sentence is about a woman being persuaded. The brain resolves this indecisive state by monitoring, to check for the possibility of a processing error.

The goal of the present experiment was to test whether a monitoring process triggered by a conflict could also be present at the word level. In a study with German-speaking participants, Münte et al. (1998) observed a P600 effect after orthographic anomalies (Die Hexe benutzte ihren Behsen, um zum Wald zu fliegen. Literal translation: ‘The witch used her broome to fly to the forest’). Could this P600 effect stem from a conflict, similar to what we proposed for the semantic anomalies? One might indeed argue that there is a strong tendency to accept the word Behsen (broome). First of all, it is semantically highly expected. Secondly, the phonological form of the word confirms this expectation and makes it maximally strong. It is as if we ask a participant in a Cloze test to fill in a word that refers to something witches tend to use to fly to the forest, a word that sounds like /brum/. It would seem that this word is 100% predictable. On the other hand, there will be a strong tendency to reject the word since the orthographic form does not fit the phonological form: the word is misspelled. So we may indeed have a strong conflict between a tendency to accept and a tendency to reject the misspelled word in this context. To be sure, the source of the conflict is very different from what we saw in the case of the semantic anomalies. Here, the tendency to accept stemmed from the fact that a lexical strategy indicated a highly plausible interpretation and the regular parse a highly implausible interpretation. In both cases, a conflict seems to exist nevertheless. An interpretation cannot be simultaneously plausible and implausible. Similarly, a particular word cannot simultaneously be there and not be there. A conflict like this would bring the cognitive system into a state of indecision. Thereby, it would trigger a monitoring response, involving reprocessing the critical linguistic string to detect and restore a possible processing error.

In our test of the monitoring hypothesis at the word level, we followed Münte et al. (1998) by presenting pseudohomophones in high-cloze contexts. What we added was a low-cloze condition. So we created both a high and a low-cloze context for a particular lexical item. There was no difference in high and low-cloze context sentences except for the critical item (“The pillows are stuffed with feathers which make them feel soft.” vs. “The pillows are stuffed with books which makes them feel hard”). The critical lexical item was either spelled correctly or was a pseudohomophone derived from the expected word and phonetically similar to the expected word. The pseudohomophone was created by changing the vowel of the second syllable (e.g., pseudohomophone derived from the word ‘boeken’ is ‘boeken’); so, the changed vowels were unstressed and always in the second part of the word (Table 1).

The predictions were as follows. First, for the high-cloze context, we expected the system to anticipate that a particular word will occur, and then start up a monitoring process when a different stimulus that is phonologically identical and orthographically similar to the highly expected item is actually presented. We predicted that this monitoring process will elicit a P600. Second, for the low-cloze sentences, the lexical items from which the pseudohomophones are derived are not highly expected and thus should not elicit a conflict between the expected and actually presented lexical item. Consequently, no monitoring process and hence no P600 effect were expected to occur. Third, we predict a standard N400 effect of cloze probability when comparing correctly spelled words in the high and the low-cloze condition. This is a classical finding but

\footnote{The dominant view in psycholinguistics is that language processing proceeds in a strict bottom-up fashion. However, there are at least some recent ERP studies that provide clear evidence for top down influences; that is, language users online generate expectancies for upcoming words. Of special interest is a study by DeLong et al. (2005), who show that expectancies manipulate ERP responses in a graded fashion. In particular, they show that participants do not only generate expectancies in high constrained sentences but even in less constrained sentences (i.e. for cloze probability less than 0.5).}
orthographically mediated process. Specifically, the amplitude of the N270 would be modulated by mismatches between the orthographic input and the expected orthographic form. They found evidence in some participants for this effect. Because our high-cloze and low-cloze pseudohomophones are orthographically inconsistent with the expected word, we propose that an N270 to pseudohomophones might occur in both the high- and low-cloze conditions.

2. Results

2.1. Reaction time pilot study

MANOVAs were performed for the RT and the error data with repeated measures on Cloze (high vs. low) and Lexicality (word vs. pseudohomophone).

For RT, a main effect of cloze ($F(1,32) = 51.56, P < 0.001$) revealed that mean RT for high-cloze sentences was shorter than for low-cloze sentences (see Table 2). In addition, a cloze by lexicality interaction ($F(1,32) = 14.56, P = 0.001$) reflected that only for the low-cloze sentences mean RT to words was longer than that to pseudohomophones ($F(1,32) = 17.43, P < 0.001$). The error analyses revealed a main effect of cloze ($F(1,32) = 6.11, P < 0.02$), indicating that participants made less errors on high-cloze sentences (6%) than on low-cloze sentences (9%). The cloze by lexicality interaction was marginally significant ($F(1,32) = 3.132, P = 0.086$). The trend reflected that, for the low-cloze sentences, more errors were made on the words than on the pseudohomophones ($F(1,32) = 4.11, P = 0.051$). No such difference was found for the high-cloze sentences ($F < 1$).

2.2. Event-related potentials

Grand averages for the high-cloze sentences and for the low-cloze sentences time-locked to the onset of the critical letter strings are presented in Figs. 1 and 2, respectively. All conditions elicited an early ERP response that is characteristic for visual stimuli that is, an N1 followed by a P2 which at occipital sites was preceded by a P1. These early components were followed by a broad negative-going wave peaking at about 400 ms, the N400, which was followed by a slow positive shift, the P600, starting at about 500 ms and extending up to 1000 ms. Inspection of Figs. 1 and 2 suggests that (1) words and pseudowords elicited an N400, and (2) pseudohomophones elicited a large P600 (mean amplitudes

<table>
<thead>
<tr>
<th>Table 1 – Examples of the high-cloze and low-cloze versions and the correct word and pseudohomophone versions of the critical sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Correct</strong></td>
</tr>
<tr>
<td>High cloze (Word-by-word translation) (Paraphrase)</td>
</tr>
<tr>
<td>In that library the pupils borrow books to take home</td>
</tr>
<tr>
<td>Low cloze (Word-by-word translation) (Paraphrase)</td>
</tr>
<tr>
<td>The pillows are stuffed with books which make them feel hard</td>
</tr>
</tbody>
</table>

| Table 2 – Mean reaction time (RT) and error percentages (Error) with standard deviations (SD), for the high-cloze sentences, low-cloze sentences and correct words, and pseudohomophones |
|----------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                      | **RT** | **SD** | **Error** | **SD** | **RT** | **SD** | **Error** | **SD** |
| **High-cloze context** |        |        |        |        |        |        |        |        |
| Correct Word         | 804.10  | 193.21 | 0.06   | 0.06   | 892.96 | 196.70 | 0.11   | 0.11   |
| Pseudohomophone      | 811.77  | 186.36 | 0.06   | 0.05   | 849.75 | 176.87 | 0.07   | 0.07   |
| Mean                 | 807.93  | 0.06   | 871.35 | 0.09   |        |        |        |        |
| **Low-cloze context** |        |        |        |        |        |        |        |        |
| Correct Word         |        |        |        |        |        |        |        |        |
| Pseudohomophone      |        |        |        |        |        |        |        |        |
| Mean                 |        |        |        |        |        |        |        |        |

The means are marginal means averaged over either cloze or lexicality.
were more positive for pseudohomophones than for words) in high-cloze sentences but not in low-cloze sentences.

2.2.1. Statistical analyses
The mean percentage of trials that was rejected based on artifacts for the high-cloze condition and the control condition was 26.12% and 26.03%, and for the low-cloze and the control condition was 26.91% and 26.97%, respectively.

The description of the ERP results will be restricted to main effects and interactions that are relevant for the cognitive-functional interpretation of the condition effects.

2.2.2. N400 window (400–500 ms)
A main effect of cloze reflected that mean N400 amplitude across conditions was larger (i.e., more negative-going) to the critical letter strings in low-cloze sentences than in high-cloze sentences (midline sites: \(F(1,32) = 48.63, P < 0.001\); lateral sites: \(F(1,32) = 30.43, P < 0.001\)). For the midline sites, a trend for a lexicality effect \(F(1,32) = 3.30, P = 0.079\) and a lexicality by site interaction \(F(4,29) = 3.12, P < 0.05\) reflected that at posterior sites mean N400 amplitude was more negative for words than for pseudohomophones. For the lateral sites, a five-way interaction between cloze, lexicality, ROI, hemisphere, and site was found \(F(4,29) = 3.07, P < 0.05\). This interaction revealed differences in N400 pattern between words and pseudowords as a function of cloze. Follow-up analyses for the high-cloze condition revealed larger N400 amplitudes for pseudohomophones than words at two sites of the left-hemisphere (LT and T5: Ps < 0.05). The analyses for the low-cloze condition revealed larger N400 amplitudes for words than pseudohomophones at posterior sites of the right-hemisphere and the left occipital site (P4, P4P, OR, T6, RTP, RT, and OL: all Ps < 0.05), whereas N400 amplitude for a subset of left-hemisphere sites (F7, LAT, and LT: all Ps < 0.02) was larger for pseudowords than for words (see Fig. 2).

2.2.3. P600 window (650–850 ms)
Effects of cloze for the midline sites \(F(1,32) = 33.18, P < 0.001\) and for the lateral sites \(F(1,32) = 35.63, P < 0.001\), reflected that overall mean amplitudes were more positive for the high-cloze sentences than for the low-cloze sentences. In addition, a cloze by lexicality interaction for the midline sites \(F(1,32) = 25.10,\)
Effects of lexicality confirmed that for the high-cloze sentences mean P600 amplitude was larger for pseudohomophones than for words [midline sites ($F(1,32) = 26.17, P < 0.001$); lateral sites ($F(1,32) = 29.67, P < 0.001$)]. The midline analysis yielded a lexicality by site interaction ($F(4,29) = 14.48, P < 0.001$). Follow-up single site analyses revealed larger P600 amplitudes for pseudohomophones than words at Fz, Cz, Pz, and Oz ($Ps < 0.05$). The lateral analyses revealed interactions between lexicality and hemisphere ($F(1,32) = 13.43, P < 0.01$), lexicality and ROI ($F(1,32) = 59.67, P < 0.001$), and between lexicality, ROI, and site ($F(4,29) = 7.81, P < 0.001$). Follow-up analyses revealed larger P600 amplitudes to pseudohomophones than words at bilateral temporal and posterior sites (Lt, Rt, Ltp, Rtp, T5, T6, P3, P4, T5, T6, P3p, P4p, Ol, and Or; $Ps < 0.05$) and at two anterior sites (F4 and Rat; $P < 0.01$).

2.2.5. Low-cloze sentences

Main effects of lexicality (midline sites: ($F(1,32) = 4.93, P < 0.05$); lateral sites: ($F(1,32) = 9.28, P < 0.01$)) disclosed that overall mean amplitudes were more negative-going for pseudohomophones than for words. Thus, at first sight, a reversed P600 pattern seemed to be elicited by the pseudohomophones in the low-cloze sentences. Closer inspection of the waveforms, however, seems more compatible with a different interpretation, namely, that the unexpected pattern was caused by the unrelated word condition and not by the pseudohomophones. Specifically, the low-cloze word condition at central/posterior sites shows a biphasic pattern—that is, a large N400 followed by a P600 (see e.g., Fig. 2: Pz, P4, and RTP). Therefore, in the following, the ERP differences will be described as an increase in P600 amplitude to the low-cloze words compared with the pseudohomophones.

For the midline sites, a lexicality by site interaction ($F(4,29) = 3.082, P < 0.05$) reflected larger P600 amplitudes for words than pseudohomophones at Fz, Pz, and Oz (all $Ps < 0.05$). The lateral analyses yielded interactions between lexicality and hemisphere ($F(1,32) = 10.54, P < 0.01$), between lexicality and ROI ($F(1,32) = 10.91, P < 0.01$), and between lexicality, ROI, and site ($F(4,29) = 18.79, P < 0.001$). Separate analyses for the single sites indicated that mean P600 amplitude for words was larger than for pseudohomophones at the following temporal and temporoparietal sites (Lt, Rt, Ltp, Rtp, T5, and T6), bilateral posterior sites (Ol, Or,
P3, P4, P3p, and P4p; Ps < 0.01), and two left anterior sites (F3 and LAT; P < 0.05).

To sum up, the follow-up analyses of the cloze by lexicality interactions confirmed that different P600 patterns were obtained for the two levels of cloze probability. As predicted pseudohomophones embedded in high-cloze sentences elicited a large P600 compared to the correctly spelled highly expected word (counterpart). The P600 effect was widely distributed across the scalp and showed a central/posterior maximum. An unexpected result was that the critical words in low-cloze sentences elicited a larger P600 than pseudohomophones.

2.2.6. N400 effect of cloze probability for words and pseudohomophones

The N400 analyses presented above focused on the comparison of the critical items—that is, words vs. pseudohomophones. The question remained whether a standard N400 effect of cloze probability occurred in the present study (see Figs. 3 and 4). Therefore, supplementary analyses were conducted for the words only. These analyses revealed clear effects of cloze for the midline sites \(F(1,32) = 28.23, P < 0.001\) and for the lateral sites \(F(1,32) = 26.57, P < 0.001\). A cloze by site interaction for the midline sites \(F(4,29) = 9.00, P < 0.001\) indicated that a standard N400 effect was present at central/posterior sites (Cz, Pz, and Oz: Ps < 0.05). The lateral analysis yielded a four-way interaction of cloze by ROI by hemisphere by site, \(F(4,29) = 3.54, P < 0.05\). The interaction reflected: first, that an N400 effect was present at bilateral temporal (Lt, Rt, LTP, RTP, T5, and T6: P < 0.05), and bilateral posterior sites (P3p, P4p, F3, P4, Ol, and OR: P < 0.05), and, second, that only for the right hemisphere N400 effects extended to anterior/temporal sites (F4a, F4, F8, and Rat: P < 0.05).

To check whether an N400 effect of cloze was also obtained for pseudohomophones, the same set of analyses was carried out for pseudowords only. These analyses, also, revealed clear effects of cloze for the midline sites \(F(1,32) = 33.94, P < 0.001\) and for the lateral sites \(F(1,32) = 14.67, P < 0.001\). For the midline sites, a cloze by site interaction \(F(4,29) = 9.50, P < 0.001\) reflected that – although significant N400 effects occurred at all sites – the N400 effect was most pronounced at Cz and Pz. For the lateral sites, an interaction between cloze, ROI, and site was found \(F(4,29) = 11.80, P < 0.001\). Follow-up analyses indicated that N400 effects were present for a subset of frontal and temporal sites (F3, F4a, F4, Lt, and Rt: all Ps < 0.05) and for bilateral posterior sites (LTP, RTP, P3p, P4p, F3, P4, and OR: all Ps < 0.05).

In sum, the analyses in which the words and pseudowords were analyzed separately, showed that clear N400 effects of
cloze probability were present for words and pseudohomophones. The main difference as a function of lexicality was that the N400 effect to pseudohomophones was more widely distributed than that for words, including frontal midline and left frontal sites.

2.2.7. N270 window

Newman and Connolly (2004) observed an N270 after orthographically incongruent words and pseudohomophones. To check if an N270 was present—like Newman and Connolly—we performed analyses on the most negative peak in the 200 to 350 ms window. These analyses revealed effects of cloze for the midline sites \(F(1,32) = 7.26, P < 0.05\) and for the lateral sites \(F(1,32) = 5.67, P < 0.05\). The amplitude of the N270 was larger for low-cloze sentences than for high-cloze sentences. The midline analysis revealed an interaction of cloze by site \(F(4,29) = 5.01, P < 0.01\). However, separate MANOVAs for the individual sites did not yield reliable effects. More importantly, the lateral analysis yielded interactions of cloze by lexicality by site \(F(4,29) = 2.95, P < 0.05\) and of cloze by lexicality by ROI by hemisphere \(F(1,32) = 4.44, P < 0.05\). Hence, separate analyses were performed for the two levels of cloze.

For the high-cloze sentences, no effects or interactions with lexicality were found \(F < 2.5\). For the low-cloze sentences, a lexicality by site interaction was obtained \(F(4,29) = 3.10, P < 0.05\). Follow-up analyses indicated that only for the low-cloze sentences N270 amplitude was larger for pseudohomophones than for words at the following left frontal sites \(F(7, F7a, \text{and } F3a; \text{all } Ps < 0.05)\).

3. Discussion

In the Introduction, we proposed that a conflict triggers a monitoring process in perception, similar to what has been shown in the action domain. This conflict was said to arise between different tendencies: the tendency to reject and the tendency to accept the sentence. Previously, Kolk et al. (2003) had claimed that a conflict between two kinds of processing (that is, heuristic and algorithmic processing) triggers a monitoring response and that it is this process that underlies the P600 effect observed to semantically implausible sentences relative to their plausible counterparts.
The prediction of the present study was that the pseudohomophone/high-cloze sentences would lead to a conflict at the word level, between the tendency to accept the pseudohomophone, and the tendency to reject it. The tendency to accept the pseudohomophone was supposed to be very strong not only because it corresponded to a word which was semantically highly expected but also because the phonological form of the word confirmed this expectation and made it maximally strong. On the other hand, the tendency to reject the pseudohomophone would also be very strong, because it is orthographically ill-formed. The resulting conflict was expected to bring the brain into a state of indecision and elicit a monitoring response that should give rise to a P600. In the low-cloze condition, the lexical items from which the pseudohomophones were derived were not expected and thus should not create a mismatch between an expected and an actually presented lexical item. Hence, no monitoring process and no P600 were expected to occur.

The ERP data confirmed the present prediction in that only pseudohomophones embedded in a high-cloze context gave rise to a P600 effect. Because the words from which the pseudohomophones were derived were highly expected, initially the pseudohomophones were easily integrated into the higher order meaning representation of the context. After all, the phonological representation of the pseudohomophone is congruent with the sentential constraints. But when the subject detected the misspelling, which signals a possible processing error, a monitoring response was triggered. This monitoring process gave rise to the P600 effect. Münte et al. (1998) also observed a P600 effect after orthographic anomalies in stories (see Introduction). In the present experiment, we systematically varied the degree of expectation of the critical word by comparing a high-cloze and a low-cloze context. Thereby, in the low-cloze sentences, the words and pseudohomophones were not expected. Consequently, the pseudohomophones were not qualified as possible processing errors and did not confuse the reader; hence, no monitoring process or P600 was triggered. Apparently, a P600 effect is only elicited in cases of a strong conflict, when an unexpected linguistic event is observed while another event is predicted with more or less 100% certainty.

There was one unexpected finding: namely that a P600 effect followed the N400 effect of cloze probability for the low-cloze words. How can we explain this biphasic pattern? The individual content words that induce the strong expectation for a particular lexical item were the same in the high and low-cloze condition (e.g., The pillows are stuffed with feathers vs. The pillows are stuffed with books). Given the high mean cloze value (above 90%), the expectation for a particular noun was very strong. The strong sentence constraint as such could have resulted in a bias to accept the noun, even in the absence of a pseudohomophone that was phonologically identical and orthographically similar to the highly expected word (as was the case for the pseudohomophones in the high-cloze condition). The strong sentence constraint could have resulted in a bias to accept the noun in the low-cloze condition, that is to assume that the highly expected noun was actually presented, to ‘fill it in’ so to speak. Having actually perceived the low-cloze item, a conflict would arise, leading to reprocessing. This could have resulted in some monitoring activity that gave rise to a larger P600 for the words compared to the pseudohomophones. This explanation also seems in accordance with the RT data. Only for the low-cloze sentences, mean RT to words was longer than that to pseudohomophones.\(^2\) Hence, both the EEG as well as the reaction time data suggest that some monitoring activity occurred to an unexpected event in a strong biasing sentential context.

In line with the literature, a standard N400 effect of cloze probability was found for the high-cloze words vs. low-cloze words, when the critical items were in mid-sentence position. Interestingly, an N400 effect of cloze probability was also observed for pseudohomophones when the critical items were in mid-sentence position. This latter finding fits with the proposal by Newman and Connolly (2004) that the phonological representation of the letter string influences the integration of word meaning with sentential context. Because the phonological representation is congruent with the semantic context in the high-cloze sentences but incongruent with the semantic context in the low-cloze sentences, an N400 effect was elicited. The N400 effect for pseudohomophones observed in the present study speaks in favor of the existence of a phonologically mediated pathway that facilitates semantic integration independently of orthography.

According to Newman and Connolly (2004) an N270 should be elicited whenever a mismatch occurs between a given word and a representation of its orthographic form. In the present study, however, the N270 effect was only observed to pseudohomophones in the low-cloze condition. It is unclear why no N270 was observed to pseudohomophones in the high-cloze sentences. One possibility is that we are dealing with a phenomenon related to what we observed in our sentence processing studies (Kolk et al., 2003). Here, it was found that participants were subject to a temporary semantic illusion, as they did not notice the difference in plausibility between ‘the poacher that hunted the fox’ and ‘the fox that hunted the poacher’. In a similar way, our participants may be subject to an orthographic illusion, at least a temporary one, in the high-cloze condition. Just as subjects appear to think that the anomalous sentence is correct in the reversal anomalies, they might think that the misspelled word is orthographically correct in the high-cloze condition. This temporary short-sighted way may be caused by the fact that phonologically and semantically the sentence is intact. It may also be related with the difficulty with proofreading, during which orthographic errors must be detected in otherwise meaningful texts. At some later point, however, the mismatch is detected, leading to a monitoring response, reflected in a P600 effect. In the low-cloze condition, the bias to accept the misspelled word is less strong, and therefore an N270 is elicited. It is clear that future studies are needed to further determine the antecedent conditions for the N270.

The P600 effect has typically been described as an index of syntactic processing (e.g., Hagoort et al., 1993; Osterhout and

\(^2\) This opposite to the standard lexicality effect which implies that RTs are shorter to words than to nonwords.
This interpretation has been challenged by studies demonstrating P600 effects after semantic anomalies. Current accounts for a P600 elicited by semantic anomalies propose that individual word meanings 'cue', 'suggest', or 'prime' a plausible role assignment for both plausible and implausible sentences, even in syntactically unambiguous sentences. Furthermore, they assume that the P600 effect is as an immediate consequence of a difference in interpretation: a plausible interpretation on the basis of the individual word meanings and an implausible one by the parser. The accounts differ, however, in their description of this immediate consequence. One possible consequence of the mismatch between lexical and syntactic analysis was investigated by Van Herten et al. (2005) (see Kim and Osterhout, 2005 for a related idea). They hypothesized that P600 effects to semantically anomalous sentences could arise if the semantic interpretation on the basis of the lexical analysis leads to a strong bias to expect a particular grammatical morphological. The discrepancy between the expected and the observed morphology would then underline the P600. Van Herten et al. (2005) showed that the P600 effect to reversal anomalies was not due to a syntactic mismatch, but was a response to the semantic anomaly (the meaning of the expected verb) as such. A second possibility was suggested by Kuperberg et al. (2003). Since the semantic relationship between the individual words suggests one set of role assignments and the regular parse another, a mismatch occurs. In response to this mismatch, the processing system is said to 'repair the anomaly by reassigning thematic roles' (Kuperberg et al., 2003, p. 128). This repair process is of a syntactic nature because it involves a process of restructuring. The notion of reprocessing embodied in this hypothesis certainly seems part of the picture. Syntactic reprocessing is also assumed to occur in garden path sentences, where it serves to uncover the alternative parse and interpretation (e.g., Friederici, 1995). Our view on the P600 is that it reflects reprocessing indeed but that its function is more general than syntactic repair. Its function is to monitor for processing errors. As a consequence, it could involve reprocessing at a number of linguistic levels, just as speech repair may involve phonological, syntactic, lexical, and conceptual levels, leading to phonological, syntactic, lexical, and appropriateness repairs, respectively. In support of this approach, the present study has demonstrated that, in language perception, in addition to a monitoring process at the conceptual/semantic level, there is also a process of monitoring for errors on the orthographic level.

As shown by previous studies, the P600 effect to syntactic and certain semantic violations shows a central/posterior scalp distribution (Coulson et al., 1998; Kolk et al., 2003). The P600 to pseudohomophones had a slightly different scalp distribution. We observed that the present P600 was centroparietally distributed across the scalp extending to two right anterior sites. One might argue that this could indicate that the P600 to pseudohomophones is qualitatively different to the earlier reported P600. However, it must be realized that there is evidence for quite some variation in scalp distribution of the P600 after syntactic and semantic anomalies. A more frontal/broad distribution of the P600 effect has for example also been reported for locally ambiguous sentences (Friederici et al., 1996; Hagoort et al., 1999; Osterhout and Holcomb, 1995; Van Berkum et al., 1999) while Kaan and Swaab (2003) observed a more posterior distribution. As Fig. 1 shows, the present P600 had an early onset. With respect to P600 latencies, Friederici (1995) has proposed that differences in latency may reflect the complexity of processing necessary for the revision of the initially preferred reading. Longer latencies seem to be correlated with more complex restructuring. An early positivity, referred to as P345, has been observed after disambiguating auxiliaries (Friederici, 1995). Intermediate positivities have been shown to occur after syntactic and certain semantic anomalies (time range: 650–850 ms). Finally, even later positivities (time range: 700–1300 ms) have been reported in a recent discourse study (Nieuwland and Van Berkum, 2005). Hence, differences in latency may indeed be systematically related to the complexity of the language input that has to be checked for possible processing errors.

From this, we conclude that the reported variability in scalp distribution and latency of the P600 does not imply qualitatively different processes. According to the monitoring hypothesis, this variability is understandable. This could be understood as a function of type of material that needs to be processed and its complexity.

It thus seems that, when we are looking at P600 effects of syntactic violation or syntactic ambiguity, we are not observing manifestations of syntactic processing as such, but of higher order – executive – processes of conflict resolution. This conclusion bears a striking resemblance to the one reached by Thompson–Schill and her colleagues, on the basis of fMRI research (for a review see Novick et al., 2005). One essential finding is that garden path sentences, which as we saw above reliably elicit P600 effects, also lead to activity in the left inferior frontal gyrus (LIFG). This is not – in view of the authors – because this area is specialized in syntactic processing, but because this is an area for conflict resolution. One important argument is that this area becomes routinely activated in incongruent trials during the Stroop task. Another argument is that damage to just this area leads to only a very minor and transient language impairment, and not to something likeagrammatism, as one might expect. On the other hand, patients with LIFG damage have severe difficulty in completing sentences when there are many competing possibilities. It has to be noted that the authors assume that both the LIFG and the Anterior Cingulate are involved in monitoring and conflict resolution, but that they do so under slightly different circumstances. We may conclude that both ERP and fMRI studies point to the presence of higher order processes of cognitive control and monitoring during language comprehension.

4. **Experimental procedures**

4.1. **Participants**

Thirty-three students (mean age = 22 years; age range = 18 to 33) participated in the experiment. All were native speakers of Dutch, had no reading disabilities, were right-
<table>
<thead>
<tr>
<th>#</th>
<th>Word</th>
<th>Pseudo Homophone</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>flessen</td>
<td>flessun</td>
</tr>
<tr>
<td>2</td>
<td>schilderijen</td>
<td>schilderijun</td>
</tr>
<tr>
<td>3</td>
<td>klompen</td>
<td>klompun</td>
</tr>
<tr>
<td>4</td>
<td>sigaretten</td>
<td>sigarettun</td>
</tr>
<tr>
<td>5</td>
<td>muisjes</td>
<td>muisjus</td>
</tr>
<tr>
<td>6</td>
<td>talen</td>
<td>talun</td>
</tr>
<tr>
<td>7</td>
<td>camera</td>
<td>camura</td>
</tr>
<tr>
<td>8</td>
<td>werknemers</td>
<td>werkemurs</td>
</tr>
<tr>
<td>9</td>
<td>oksels</td>
<td>oksuls</td>
</tr>
<tr>
<td>10</td>
<td>honger</td>
<td>hongur</td>
</tr>
<tr>
<td>11</td>
<td>winkel</td>
<td>winkul</td>
</tr>
<tr>
<td>12</td>
<td>paarden</td>
<td>paardun</td>
</tr>
<tr>
<td>13</td>
<td>nagels</td>
<td>naguls</td>
</tr>
<tr>
<td>14</td>
<td>armen</td>
<td>armun</td>
</tr>
<tr>
<td>15</td>
<td>bodem</td>
<td>bodum</td>
</tr>
<tr>
<td>16</td>
<td>poten</td>
<td>potun</td>
</tr>
<tr>
<td>17</td>
<td>lepels</td>
<td>lepuls</td>
</tr>
<tr>
<td>18</td>
<td>ongelnuk</td>
<td>ongulkun</td>
</tr>
<tr>
<td>19</td>
<td>bloemen</td>
<td>bloemun</td>
</tr>
<tr>
<td>20</td>
<td>tulpen</td>
<td>tulpun</td>
</tr>
<tr>
<td>21</td>
<td>eieren</td>
<td>eirun</td>
</tr>
<tr>
<td>22</td>
<td>kalender</td>
<td>kalendar</td>
</tr>
<tr>
<td>23</td>
<td>veren</td>
<td>verun</td>
</tr>
<tr>
<td>24</td>
<td>boeken</td>
<td>boekun</td>
</tr>
<tr>
<td>25</td>
<td>zeilen</td>
<td>zeilun</td>
</tr>
<tr>
<td>26</td>
<td>kapper</td>
<td>kappur</td>
</tr>
<tr>
<td>27</td>
<td>duikers</td>
<td>duikurs</td>
</tr>
<tr>
<td>28</td>
<td>water</td>
<td>watur</td>
</tr>
<tr>
<td>29</td>
<td>kolen</td>
<td>kolun</td>
</tr>
<tr>
<td>30</td>
<td>computers</td>
<td>computurs</td>
</tr>
<tr>
<td>31</td>
<td>bomen</td>
<td>bomun</td>
</tr>
<tr>
<td>32</td>
<td>dobber</td>
<td>doobur</td>
</tr>
<tr>
<td>33</td>
<td>babbels</td>
<td>babbuls</td>
</tr>
<tr>
<td>34</td>
<td>vlaggen</td>
<td>vlaggun</td>
</tr>
<tr>
<td>35</td>
<td>ogen</td>
<td>ogun</td>
</tr>
<tr>
<td>36</td>
<td>ramen</td>
<td>ramun</td>
</tr>
<tr>
<td>37</td>
<td>klieren</td>
<td>klierun</td>
</tr>
<tr>
<td>38</td>
<td>schapen</td>
<td>schapun</td>
</tr>
<tr>
<td>39</td>
<td>tenen</td>
<td>tenun</td>
</tr>
<tr>
<td>40</td>
<td>klanten</td>
<td>klantun</td>
</tr>
<tr>
<td>41</td>
<td>rekening</td>
<td>reikuning</td>
</tr>
<tr>
<td>42</td>
<td>vissen</td>
<td>vissun</td>
</tr>
<tr>
<td>43</td>
<td>lippen</td>
<td>lippun</td>
</tr>
<tr>
<td>44</td>
<td>bliskaem</td>
<td>bliskum</td>
</tr>
<tr>
<td>45</td>
<td>hengel</td>
<td>hengul</td>
</tr>
<tr>
<td>46</td>
<td>obers</td>
<td>oburs</td>
</tr>
<tr>
<td>47</td>
<td>benen</td>
<td>benun</td>
</tr>
<tr>
<td>48</td>
<td>suiker</td>
<td>suikur</td>
</tr>
<tr>
<td>49</td>
<td>kamers</td>
<td>kamurs</td>
</tr>
<tr>
<td>50</td>
<td>stenen</td>
<td>stenun</td>
</tr>
<tr>
<td>51</td>
<td>granten</td>
<td>grantun</td>
</tr>
<tr>
<td>52</td>
<td>brommer</td>
<td>brommur</td>
</tr>
<tr>
<td>53</td>
<td>hemel</td>
<td>hemul</td>
</tr>
<tr>
<td>54</td>
<td>pilaren</td>
<td>pilurun</td>
</tr>
<tr>
<td>55</td>
<td>kaarsen</td>
<td>kaarsun</td>
</tr>
<tr>
<td>56</td>
<td>vingers</td>
<td>vingurs</td>
</tr>
<tr>
<td>57</td>
<td>kater</td>
<td>katur</td>
</tr>
</tbody>
</table>

**Table 3 (continued)**

-t is changed to -d

62 | kist | kisd |
63 | talent | talend |
64 | bruiloft | bruijofd |
65 | taart | taard |
66 | koorts | koords |
67 | diefstal | diefsdal |
68 | toekomst | toekomsd |
69 | botsing | bodsing |
70 | paspoort | paspoord |
71 | dirigent | dirigend |
72 | koelkast | koelkasd |
73 | bult | buld |

-d is changed to -t

74 | dienblad | dienblat |
75 | brood | broot |
76 | draad | draat |
77 | eendjes | eenjjes |
78 | stad | stat |
79 | mond | mont |
80 | zand | zant |

-i is changed to -ei

81 | pijn | peil |
82 | batterij | batterei |
83 | bedrijf | bedreif |
84 | termijn | termein |
85 | krijtje | kreitje |
86 | bijbel | beibel |
87 | liijn | leim |
88 | diijken | deiken |

-g is changed to -ch

89 | toegift | toechift |
90 | begrip | bechrip |
91 | congres | conchres |

-ch is changed to -g

92 | bocht | bogg |
93 | vlucht | vlugt |

-ou is changed to -au

94 | zout | zaut |
95 | berouw | berauw |

-au is changed to -ou

96 | paus | pous |

-v is changed to -f

97 | vervuiling | verfuiling |
98 | druiven | druifen |
99 | brieven | briefen |

-f is changed to -v

100 | tafel | tavel |

-i is changed to -ie

101 | casino | casieno |
102 | kilo | kielo |
103 | tribune | tribune |
104 | conditie | condietie |
Table 3 (continued)

<table>
<thead>
<tr>
<th>#</th>
<th>Word Replace</th>
<th>Word Original</th>
</tr>
</thead>
<tbody>
<tr>
<td>105</td>
<td>huisen</td>
<td>huisen</td>
</tr>
<tr>
<td>106</td>
<td>bezems</td>
<td>besems</td>
</tr>
<tr>
<td>107</td>
<td>kiezers</td>
<td>kiezers</td>
</tr>
<tr>
<td>108</td>
<td>kiezen</td>
<td>kiesen</td>
</tr>
<tr>
<td>109</td>
<td>muziek</td>
<td>musiek</td>
</tr>
<tr>
<td>110</td>
<td>panty</td>
<td>penty</td>
</tr>
<tr>
<td>111</td>
<td>pensioen</td>
<td>pensjoen</td>
</tr>
<tr>
<td>112</td>
<td>koeien</td>
<td>koejen</td>
</tr>
<tr>
<td>113</td>
<td>patient</td>
<td>pasient</td>
</tr>
<tr>
<td>114</td>
<td>proces</td>
<td>proses</td>
</tr>
<tr>
<td>115</td>
<td>trein</td>
<td>trijn</td>
</tr>
<tr>
<td>116</td>
<td>web</td>
<td>wep</td>
</tr>
</tbody>
</table>

The critical words are sorted by type of anomaly of the pseudo homophone.

4.2. Materials

We first constructed 127 simple declarative sentence fragments and used these in a cloze test with 25 subjects to obtain highly expected (‘high-cloze’) critical words. Of these 127 sentences, 116 sentences were completed with the same word by 91% of the participants. These were used as the high-cloze context sentence fragments in this study.

We then created 116 low-cloze context sentences by exchanging the critical word from a high-cloze context fragment with the critical word from another high-cloze context fragment. For example, we exchanged the critical word from ‘In that library the pupils borrow books to take home.’ with the critical word from ‘The pillows are stuffed with feathers which makes them feel soft.’ resulting in the following low-cloze fragment ‘The pillows are stuffed with books which makes them feel hard.’ The critical word was always in mid-sentence position.

A further experimental manipulation was Lexicality (correct word vs. from the correct word derived pseudo-homophone). Every critical word occurred in a correct version and a pseudohomophonic version. The pseudohomophone was created by changing the vowel of the second syllable, keeping phonology the same. Table 3 contains all types of pseudohomophones we used; note that every noun contained two syllables and that the changed vowels were always in the second part of the word.

The two experimental manipulations Context and Lexicality were crossed. As a result, there were four conditions and thus, four experimental sentence types: high-cloze correct word sentences, high-cloze pseudohomophone sentences, low-cloze correct word sentences, and low-cloze pseudohomophone sentences; yielding a total set of 464 sentences. The four versions of each sentence were counterbalanced across lists. Each list contained each sentence context (in a high-cloze or a low-cloze version) and each critical word (in a correct word or a pseudohomophone version) only once. So, each list contained 29 high-cloze correct word sentences, 29 high-cloze pseudohomophone sentences, 29 low-cloze correct word sentences, and 29 low-cloze pseudohomophone sentences. To each list, 60 filler sentences were added: 30 correct sentences, 10 sentences with a pseudohomophone at the beginning of the sentence, 10 sentences with a pseudohomophone in the middle of the sentence, and 10 sentences with a pseudohomophone at the end of the sentence. To avoid sentence wrap-up processes to affect our measurement, every critical word was presented at a mid-sentence position.

4.3. Procedure

For the EEG study, participants were seated in an experimental room. Sentences were presented in serial visual presentation mode at the center of a PC monitor. Word duration was 345 ms and the stimulus-onset asynchrony (SOA) was 645 ms. Sentence final words were followed by a full stop. The intertrial interval was 2 s. Words were presented in black capitals on a white background in a 9 cm by 2 cm window at a viewing distance of approximately 1 m. Each sentence was preceded by a fixation cross (duration 510 ms) followed by a 500 ms blank screen. The experimental list was split up into five blocks; there was a brief pause between blocks and each block was preceded by two filler items. Participants were instructed to attentively read the sentences. Because eye movements distort the EEG recording, participants were trained to make eye movements, e.g., blinks, only in the period between the end of the last sentence and the beginning of the next sentence.

Prior to this EEG study, a reaction time (RT) pilot study was conducted as a pretest of the material, to test if participants were successful in detecting the pseudohomophones. A separate group of 33 participants was tested that fulfilled the same criteria as those in the ERP study. The procedure differed in two aspects from the ERP study: first, the critical letter string was presented in sentence-final position (e.g., paraphrase: “In that library the pupils borrow books.”) and, second, participants performed a lexical decision task. They had to indicate as fast as possible by pressing a button with the right or left index finger if the critical letter string was a real word (right-hand response) or not (left-hand response). A response device with three
4.4. Electrophysiological recording

The electroencephalogram (EEG) was recorded with 27 tin electrodes mounted in an elastic electrode cap (Electrocap International; see Fig. 5 for the montage).

The electrode positions included standard International 10–20 system locations over the left and right hemispheres at the frontal (F3, F4, F7, and F8), midline (Fz, Cz, Pz, and Oz), parietal (P3 and P4), and temporal (T5 and T6) sites. Eight extra electrodes were placed at the frontal (F3A, F4A, F7A, and F8A), midline (Fza and Oz), and parietal (P3P and P4P) sites. In addition, eight electrodes were placed at nonstandard electrode positions previously found to be sensitive to language manipulations (e.g., Holcomb and Neville, 1990): left and right anteriortemporal sites (LAT and RAT: 50% of the distance between T3/4 and F7/8), left and right temporal sites (LT and RT: 33% of the interaural distance lateral to Cz), left and right temporoparietal (LTP and RTP: Wernicke’s area and its right hemisphere homologue: 30% of the interaural distance lateral to a point 13% of the nasion–inion distance posterior to Cz), and left and right occipital sites (OL and OR: 50% of the distance between T5/6 and O1/2). The left mastoid served as reference. Electrode impedance was less than 3 kΩ. The electro-oculogram (EOG) was recorded bipolarly; vertical EOG was recorded by placing an electrode above and below the right eye and the horizontal EOG was recorded via a right to left canthal montage. The signals were amplified (time constant = 8 s, bandpass = 0.02–30 Hz), and digitized online at 200 Hz. Presentation of stimuli and recording of performance data were accomplished by a Macintosh computer.

4.5. Data analyses

Before analyzing EEG and EOG, records were examined for artifacts and for excessive EOG amplitude (>100 μV) from 100 ms before the onset of the critical letter string ending the relative clause to 1000 ms following its onset. Averages were aligned to a 100-ms baseline period preceding the critical letter string. Based on visual analysis and previous studies (Kolk et al., 2003; Van Herten et al., 2005), mean amplitudes in the 400–500 ms and 650–850 ms window were taken as N400 window and P600 window, respectively.

For the two time windows, repeated measures analyses of variance (ANOVAs) were conducted separately for the midline sites and for the lateral sites with cloze (high vs. low) and lexicality (word vs. pseudohomophone) as factors. The midline analyses included the additional factor site (Fza, Fz, Cz, Pz, Oz). To explore the scalp distribution of the ERP effects for the lateral analyses, we used a region of interest (ROI: anterior vs. posterior) by hemisphere by lateral site (F7a/F3a/F7/F3/LAT vs. LTP/P3/P3p/T5/OL vs. F8a/F4a/F8/F4/RAT vs. RTP/P4/P4p/T6/OR) design. The multivariate approach to repeated measurements was used to avoid problems concerning sphericity (e.g., Vasey and Tayer, 1987). Interactions with the factor site were followed up by single site analyses.

Acknowledgments

Portions of this research were presented at the 12th Annual Meeting of Cognitive Neuroscience Society in New York, 2005. We thank Gina Kuperberg and an anonymous reviewer for their very helpful comments on an earlier version of this article. We are grateful to the ERG group of the NICI for technical assistance and to Maarten vd Meulen for his support.

References


