

Electrophysiological Signatures of Visual Lexical Processing: Open- and Closed-Class Words

Colin M. Brown and Peter Hagoort

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Mariken ter Keurs

Max Planck Institute for Psycholinguistics and Institute of Neurology, University Hospital, Nijmegen, The Netherlands

Abstract

n This paper presents evidence of the disputed existence of an electrophysiological marker for the lexical-categorical distinction between open- and closed-class words. Event-related brain potentials were recorded from the scalp while subjects read a story. Separate waveforms were computed for open- and closed-class words. Two aspects of the waveforms could be reliably related to vocabulary class. The first was an early negativity in the 230- to 350-msec epoch, with a bilateral anterior predominance. This negativity was elicited by open- and closed-class words alike, was not affected by word frequency or word length, and had an earlier peak latency for closed-class words.

The second was a frontal slow negative shift in the 350- to 500-msec epoch, largest over the left side of the scalp. This late negativity was only elicited by closed-class words. Although the early negativity cannot serve as a qualitative marker of the open- and closed-class distinction, it does reflect the earliest electrophysiological manifestation of the availability of categorical information from the mental lexicon. These results suggest that the brain honors the distinction between open- and closed-class words, in relation to the different roles that they play in on-line sentence processing. ■

INTRODUCTION

In this paper we investigate electrophysiological manifestations of visual lexical processing, as reflected by scalp-recorded event-related brain potentials (ERPs). Our focus is on processes of language comprehension, in particular in relation to a proposed categorization within the mental lexicon that has a long tradition within both linguistics and psycholinguistics: the separation into two vocabulary classes, variously called content and function words, or open- and closed-class words. We will use the latter terms.

The category of open-class words contains nouns, verbs, and adjectives. These words are the main bearers of meaning in the language, providing the building blocks for the overall sense that is contained in a spoken or written sentence. The category of closed-class words contains, among others, articles, conjunctions, and prepositions. These words are relatively devoid of meaning—certainly in comparison with the open-class words—and primarily serve a syntactic role in language understanding, providing crucial components for the phrase-level grammatical structure that makes combinations of words interpretable.

Broadly speaking, the distinction between open- and closed-class words can be seen as a basic reflection of the separation between semantics and syntax. The open-class words do of course have important syntactic functions, but above all they contain indispensable semantic information. No coherent and overall meaning can be derived without open-class words. This contrasts with closed-class words. Although they are not without meaning (witness, for instance, the difference in semantic scope between the indefinite “a” and definite “the”), the first and foremost function of closed-class words is to provide information on the syntactic relations that hold among open-class words. An investigation, then, of the open- and closed-class distinction is not only relevant for our knowledge of the mental lexicon but also for our understanding of the interplay between meaning and structure during language comprehension.

Although not all lexical differences are captured by the open- and closed-class categorization, the distinction is widely used in linguistics and psycholinguistics and has received support from several lines of experimental work. Early work based on corpus analysis of speech errors (e.g., Garrett, 1982) and on reaction-time measurements of visual word processing (e.g., Bradley & Garrett,

1983; Bradley, Garrett, & Zurif, 1980) led to the claim that the two word classes are represented in separate subvocabularies and that during comprehension the closed-class words can be accessed via a dedicated access mechanism (while also sharing a general access mechanism with open-class words). In particular the finding that lexical decision times for open-class words were modulated by the frequency of occurrence of these words, whereas latencies for the closed-class words were unaffected by lexical frequency, was taken as strong evidence for a computational distinction (cf. Bradley et al., 1980). However, subsequent studies did not replicate this result (Gordon & Caramazza, 1982, 1985; Segui, Mehler, Frauenfelder, & Morton, 1982), possibly in part due to inherent sensitivity problems of reaction-time measurements for (relatively small) differences between very high-frequency words (cf. Gordon & Caramazza, 1985). A similar pattern of contrasting effects has been reported for the performance of aphasic patients. Most aphasiological researchers have reported impaired processing of especially closed-class words in agrammatic patients (e.g., Bradley et al., 1980; Friederici, 1985; Swinney, Zurif, & Cutler, 1980), but this has not been consistently found (Linebarger, Schwartz, & Saffran, 1983). Finally, work on lateralized tachistoscopic presentation of open- and closed-class words has yielded evidence in support of the vocabulary distinction. Chiarello and Nudling (1987) and Mohr, Pulvermüller, and Zaidel (1994) demonstrated that closed-class words are processed fastest with right visual-field presentation (i.e., primary input to the left hemisphere), with less or no differential visual-field effect for open-class words.

The research we report in the current paper adds in several ways to the work on the open- and closed-class distinction. First, we measured ERPs, that is, real-time electrical brain activity, related to the processing of open- and closed-class words. As we shall show, ERPs are a rich tool with which to investigate on-line lexical processing.

Second, given that brain potentials are a reflection of ongoing electrophysiological activity, they can, in principle, yield information about possibly distinct neuronal populations that subserve the processing and/or representation of open- and closed-class items. Several researchers have suggested that the functional separation between the primarily syntactic nature of closed-class words and the mainly semantic nature of open-class words is mirrored by a neurobiological distinction in which distinct areas of the brain subserve the representation and processing of the two vocabulary classes (e.g., Bradley et al., 1980; Nobre & McCarthy, 1994; Pulvermüller, Lutzenberger, & Birbaumer, 1995; Swinney et al., 1980). ERPs can provide directly relevant information concerning this claim.

Finally, because ERPs can be measured without having subjects perform an overt task (a feature that the method shares with eye-movement registration), it is possible to

register brain activity during uninterrupted language processing over extended periods of time. This enabled us to look at the processing of open- and closed-class words in discourse (in the present case a relatively long story). This contrasts with the word-isolation or at best single-sentence presentations of the majority of studies on the open- and closed-class distinction.

The amount of published ERP data on the open- and closed-class distinction is quite limited, and not all findings are consistent. The main contestant for an electrophysiological marker of the open- and closed-class distinction is the so-called N280 component. This negative-polarity marker (hence the label N) has been proposed by Neville, Mills, and Lawson (1992), who had subjects read individual, unrelated sentences and make acceptability judgments about each sentence. Neville et al. reported an N280 that was only elicited by closed-class words. The N280 component that they describe reached its maximal amplitude at on average 280 msec following stimulus onset and emerged most strongly over left-anterior electrode sites. A second aspect of the ERP waveform that differentiated between closed- and open-class words in the Neville et al. data was a late negative shift occurring in the 400- to 700-msec window for closed-class words. This shift was largest over left-anterior sites and was absent for open-class words.

The open-class words in the Neville et al. (1992) study were associated with a later negativity with a posterior distribution, which can be classified as the N400 component (Kutas & Hillyard, 1980). The N400 is an established ERP marker of semantic processing during language comprehension (cf. Kutas & Van Petten, 1994). This later negativity was also observed with the closed-class words, but the size of this component was greatly reduced, and its scalp topography did not match the standard distribution of the N400. On the basis of separate analyses on the effects of lexical frequency and length on the evoked potentials, Neville et al. argued that the N280 is specific to closed-class words, distinct from the N400 component, and cannot be accounted for by lexical-statistical factors. They claimed, therefore, that nonidentical neural systems mediate the different syntactic and semantic processes associated with the closed- and open-class vocabulary.

Earlier ERP work by Kutas and Hillyard (1983) and Van Petten and Kutas (1991) did not provide evidence for an N280 component for closed-class words (see also Kutas, Van Petten, & Besson, 1988). Kutas and Hillyard used prose passages consisting of several sentences, some of which contained either semantic or grammatical anomalies. Van Petten and Kutas did not use anomalous materials but presented subjects with isolated normal sentences, syntactic prose sentences (i.e., sentences that concur with the grammatical rules of the language but that do not have an overall meaning), and randomly ordered sequences of words. Although the waveforms presented by Kutas and Hillyard (1983) do seem to

contain a more frontal negativity at approximately 250 msec for the closed-class words, the difference between the two word classes is reported to reach significance with broad temporal windows of 200 to 700 msec and 400 to 700 msec, which is well outside the restricted time window claimed to be specific for the N280 (235 to 400 msec in the Neville et al., 1992, report). To the extent that the earlier work showed any difference between the two classes, the effects were either at the level of the N400 (and could be reliably related to frequency differences) or emerged as late frontal negativities (approximately 400 to 700 msec postonset) for the closed-class words, similar to the broad negativity later reported by Neville et al. (1992). This effect was classified as a variant of the contingent negative variation (CNV; cf. Hillyard, 1973; McCallum, 1988). Van Petten and Kutas (1991) speculate that this response reflects anticipatory processes related to the fact that closed-class words in sentences often provide a functional signal that a head of phrase is imminent.

Later work showed a slightly more varied ERP profile for open- and closed-class words. Nobre and McCarthy (1994) found a significantly more negative N400 amplitude for open-class words than for closed-class words, when presented in random word lists, and interpreted the attenuated N400 to closed-class words as evidence for a different neural organization and processing of these words. Pulvermüller et al. (1995), using a lexical decision paradigm with single word/nonword presentation, reported morphologically similar waveforms elicited by closed- and open-class words, with several quantitative amplitude differences as a function of word class. Although they did find an early negative effect for closed-class words when they compared electrode sites over the left and the right side of the scalp, similar early negative shifts were also observed for open-class words, albeit with less asymmetry. Neville, Coffey, Holcomb, and Tallal (1993) presented language-impaired children and matched controls with sentences that had congruous or incongruous endings. Notwithstanding the authors' statement about the similarity of their data with the original Neville et al. (1992) study, no N280 effect or in fact no early effect was observed as a function of word class for either of the subject groups. The first indication of a separation between open- and closed-class words was at the level of the N400 (with a maximum at around 450 msec), where the control children presented an N400 to closed-class words that was more maximal over anterior and temporal electrode sites and more evenly distributed for the open-class words.

A later study in which a clear N280 component was obtained was reported by King and Kutas (1998, see also Kutas, 1997). They presented unrelated written sentences to subjects and followed half of the sentences with a true/false comprehension probe. A left-frontal negativity, maximal over anterior-temporal and frontal sites, but also present at more central sites, was elicited

by *both* closed- and open-class words. The mean peak latency of this negativity for the closed-class words was on average at 280 msec (i.e., an N280), whereas the mean peak latency for the open-class words lay around 315 msec (i.e., an N315). King and Kutas hypothesized that the N280 and the N315 are a manifestation of a common stage during visual word processing and, hence, that functionally speaking the two negativities are in fact modulations of one and the same ERP component. In support of this hypothesis, King and Kutas performed a regression analysis on the peak latency of the negativity as a function of word type, with word scarcity as the single predictor.¹ King and Kutas showed that the mean peak latencies of the left-frontal negativity at three anterior-temporal electrode sites can be predicted from the length and frequency of the words presented to the subjects. Although there were clear separations between the open- and closed-class categories—due to the inherent frequency differences between the two classes—the regression analysis revealed that within each category the peak latency of the negativity was also largely affected by the frequency properties of the words. For example, the peak latency for definite articles (the most frequent word type) was at 270 msec, whereas the latency for the less frequent prepositions was around 300 msec. However, the peak latency differences between different members of the closed and open classes were not further analyzed, so the actual statistical separation within word class is unclear.

King and Kutas (1998) conclude that the early negativity is elicited by open- and closed-class words alike, with its latency varying as a function of the lexical characteristics of the eliciting item. Hence, in contrast to the claim of Neville et al. (1992), King and Kutas argue that the early negativity (that they rename as the lexical processing negativity, LPN) “does not reflect lexical class differences” (King & Kutas, 1998, p. 4).

Osterhout, Bersick, and McKinnon (1997) added a further piece of evidence on brain potentials and vocabulary class. They discuss two sets of data, one on ERPs elicited during the reading of a passage of prose and one on ERPs elicited during the reading of a random list of words. Unlike King and Kutas (1998) and Neville et al. (1992), Osterhout and his colleagues did not report a separate early negativity that could be distinguished from the N400. Instead, they argue for the existence of only one negative shift, of which the peak latency varies between 280 and 400 msec. In separate regression analyses, using length or frequency as the predictor for the peak latency of the negativity, Osterhout et al. found that in the prose experiment some 46% of the variance could be accounted for by lexical-statistical factors, and some 74% in the word-list experiment. On the basis of these results, Osterhout et al. concluded that the brain potential data that they obtained do not provide any evidence for a qualitative difference related to the open- and closed-class distinction.

In summary, the current situation with respect to an electrophysiological marker of vocabulary class is unclear, that is, at least as far as an early ERP effect is concerned. Although the main interest to date has been on early negativities in the waveform, it should be kept in mind that a late negative slow shift in the 400- to 700-msec epoch has been consistently found for closed-class words. The uncertainty mainly concerns the status of the so-called N280 as a specific index of vocabulary class and the sensitivity of this component to lexical-statistical factors.

One of the purposes of the present study is to add to the body of evidence on the (putative) existence of the N280. In particular, we further investigate the sensitivity of evoked potentials to the length and frequency of occurrence of their eliciting words. To this end, we use a much larger frequency-coded database than used by Neville et al. (1992), King and Kutas (1998), and Osterhout et al. (1997), who used the Brown corpus (cf. Kucera & Francis, 1967; Francis & Kucera, 1982). The frequency counts in the Brown corpus are based on just over one million tokens, from text samples that were published in 1961. The frequency analyses that we will report in this paper are based on the Dutch Celex corpus (cf. Baayen, Piepenbrock, & van Rijn, 1993), which contains over 42 million tokens, derived from corpora based on samples drawn during the 1970 to 1988 period. This more extended and more current database provides a good basis for a fine-grained analysis of lexical-statistical effects.

A further purpose of the present study is to extend the available results on ERP effects related to vocabulary class by using an extensive set of stimuli. In this way we can analyze ERP effects between the open- and closed-classes, and we can also focus on the ERP componentry elicited by individual members of each of the two categories (cf. Osterhout et al., 1997). This will provide additional information on the possibly qualitative ERP differences between and within the open- and closed-class vocabularies.

Finally, in addition to a detailed investigation of the N280 component, we will also focus on the late negative shift between 400 to 700 msec, and the extent to which this shift separates the open- and closed-class words.

RESULTS

We present four separate sections on the data. In the first section we compare the ERPs as a function of the basic vocabulary distinction between open- and closed-class words, focusing on the overall differences between the two categories. In the following two sections we look at the two word classes separately and compare the ERPs elicited as a function of each word type (i.e., Articles, Prepositions, and Conjunctions in the closed-class category and Nouns, Verbs, and Adjectives in the open-class category). In the fourth section we investigate the con-

tribution to the ERP profiles of two lexical characteristics, namely, frequency of occurrence in the language and length in number of letters.

The Closed- and Open-Class Categories Compared

As can be seen in Figures 1a and 1b, the waveforms for the closed- and open-class categories both show an N1-P2 complex, which is a standard response to physical stimulation (for purposes of graphical display, electrode Fz' is not shown in this and all following figures). The amplitude of the P2 component is reduced in the closed-class words (overall analysis in the 150- to 200-msec epoch: $F(1, 12) = 18.49$, $MSE = 7.28$, $p = 0.001$). This is probably due to the stronger development of the immediately following negativity for the closed-class words.

The waveform for the closed-class category shows two main features. The first is an early negative component that develops between about 230 and 340 msec, with a maximum at about 290 msec. The second is a late negativity that is visible from 350 to 550 msec as a slowly increasing negative shift, reaching its maximum at about 460 msec.² Both components have a clear anterior distribution. The early negativity appears to be somewhat larger over the left hemisphere and is largely restricted to the anterior and anterior-temporal sites. The late negativity is likewise most prominent over anterior sites.

The waveform for the open-class category also shows two components, again an early and a late negativity. The early negativity develops in a later and somewhat more restricted time window than in the closed-class waveform, from about 260 to 350 msec, with a peak at about 310 msec, and is overall smaller in amplitude. The later negativity immediately follows the earlier one, extending from 350 msec to about 550 msec (see note 2), with a maximum around 400 msec.

Mean Amplitude Analyses on the Early Negativities

An analysis of the mean amplitude between 230 and 340 msec over 23 electrode sites revealed a significant main effect of Vocabulary Class, with the closed-class words eliciting a larger negative shift in comparison with the open-class words ($F(1, 12) = 7.51$, $MSE = 7.28$, $p < 0.02$). This main effect was also obtained in an analysis with the additional factor of Hemisphere, collapsing over the 10 left and right electrode sites ($F(1, 12) = 7.36$, $MSE = 5.76$, $p < 0.02$), in which in addition a trend toward a main effect of Hemisphere was found ($F(1, 12) = 3.41$, $p = 0.089$ after z -score conversion), reflecting the fact that the effect is marginally more negative over the left side of the scalp. No interaction was observed between Vocabulary Class and Hemisphere ($F < 1$). (See the Methods section for a description of electrode sites.)

Further analyses of the topography of the effect of vocabulary class demonstrated that the negativity was slightly larger over anterior left than right electrode sites

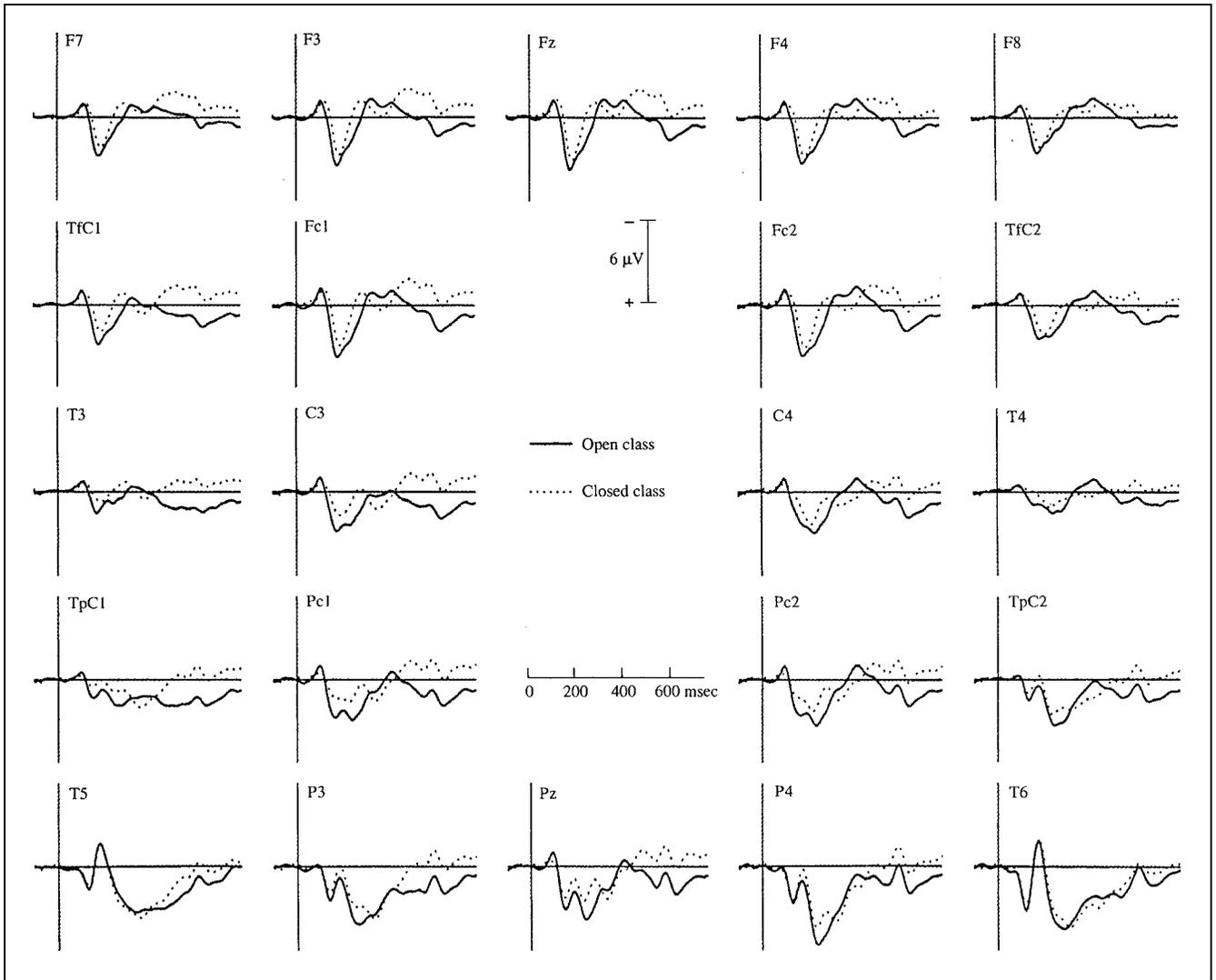
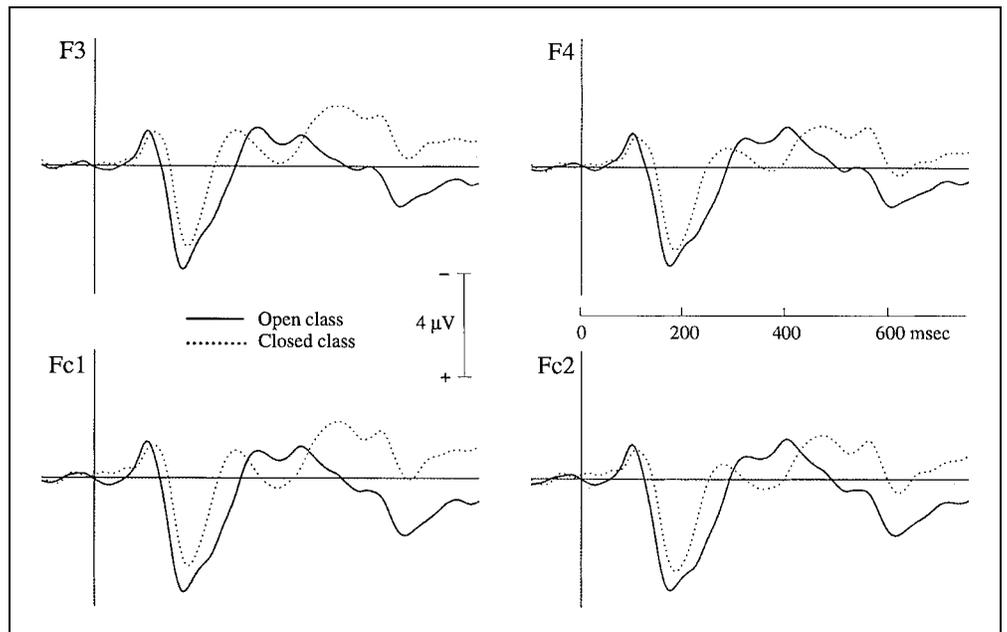


Figure 1a. Grand average waveforms ($N = 13$ subjects) elicited by the open- and closed-class categories at 22 electrode sites. The epoch in this and all other ERP figures is 850 msec long, starting 100 msec before a word was presented.

Figure 1b. Grand average waveforms ($N = 13$ subjects) elicited by the open- and closed-class categories at four anterior electrode sites.



(main effect of Anterior Left/Right: $F(1, 12) = 4.33$, $MSE = 2.65$, $p = 0.059$), but that this was equally so for open- and closed-class words (i.e., no interaction was obtained between the Anterior Left/Right factor and Vocabulary Class: $F < 1$).

A separate analysis of anterior left (AL) electrodes and anterior right (AR) electrodes, showed that the waveform for the closed-class words was significantly more negative than the waveform for the open-class words (main effect of Vocabulary Class at AL: $F(1, 12) = 4.82$, $MSE = 2.46$, $p = 0.048$; at AR: $F(1, 12) = 5.71$, $MSE = 2.35$, $p = 0.034$). A marginal difference was found for the posterior left and right sites (PL: $F(1, 12) = 4.38$, $MSE = 0.95$, $p = 0.058$; PR: $F(1, 12) = 4.15$, $MSE = 1.67$, $p = 0.064$). This latter effect can be ascribed to the presence—albeit attenuated—of the early negativity for the closed-class words at temporal and posterior-temporal sites (i.e., TpC1, TpC2, Pc1, Pc2), whereas at these sites the waveform for the open-class words shows hardly any modulation.

In summary, in the 230 to 340 msec time window, where the early negativity for the closed-class items is maximal, a clear effect of Vocabulary Class is obtained. The effect is most prominent over anterior electrode sites, about equally so over left and right sides of the scalp, with closed-class words being consistently more negative than open-class words.

In contrast to the effects in the 230- to 340-msec window, no effects of Vocabulary Class emerged in analyses in the 260 to 350 msec window, despite the fact that in this latter window the morphology for the open-class items is most distinct. The null effect holds both for an analysis over all electrodes and in analyses restricted to the same subsets of electrodes as reported above (all F values < 1.55). The most probable cause of this lack of an effect lies in the overlap of the descending flank of the early negativity for closed-class items, with the ascending flank of the early negativity for the open-class items. Nevertheless, the early waveform for the open-class items is clearly separate from the waveform for the closed-class items, primarily in terms of the latency of its peak amplitude. This conspicuous contrast between the two early negativities is underscored by the outcome of statistical analyses on the peak latencies, to which we now turn.

Peak Latency Analyses on the Early Negativities

In the light of the predominantly anterior distribution of the early components, an analysis of peak latency values within the 250- to 350-msec window was performed over only the anterior left and right electrode sites. The results of this analysis suggest a differential distribution of the peak latency of the early negativity as a function of Vocabulary Class. Not only were the main effects of Vocabulary Class ($F(1, 12) = 35.05$, $MSE = 1089$, $p =$

0.0001) and of Anterior Left/Right significant ($F(1, 12) = 15.30$, $MSE = 398$, $p = 0.002$), but the interaction of these two factors also approached significance ($F(1, 12) = 4.07$, $MSE = 330$, $p = 0.066$). The open-class words peaked at 312 msec, whereas the closed-class words peaked at 285 msec. The marginal interaction arises from the fact that whereas the peak latency for the open-class negativity over left and right sites only differs by 5 msec (LH 309 msec, RH 314 msec), this difference is 15 msec for the closed-class negativity, with the left sites showing the earliest maximum (LH 277 msec, RH 292 msec).

Mean Amplitude Analyses on the Late Negativities

An analysis over 23 electrodes in the 350- to 550-msec epoch revealed that the closed-class items elicited a significantly more negative waveform compared to the open-class items ($F(1, 12) = 6.98$, $MSE = 9.86$, $p = 0.021$). An analysis with Hemisphere as additional factor yielded a significant interaction with Vocabulary Class ($F(1, 12) = 18.75$, $p = 0.001$ after z -score conversion), reflecting that the difference between the open- and the closed-class items was larger over the left than the right hemisphere. A more detailed investigation of the distribution of the effect showed that the closed-class waveform was significantly more negative than the open-class waveform over anterior-left sites but not over anterior-right sites (AL: $F(1, 12) = 15.74$, $MSE = 2.05$, $p < 0.002$; AR: $F < 1$). The same pattern was observed in the posterior analyses (PL: $F(1, 12) = 14.78$, $MSE = 2.13$, $p < 0.003$; PR: $F < 1$).

A final analysis in which the two vocabulary classes were compared between anterior left and posterior left sites revealed a significant effect of Vocabulary Class ($F(1, 12) = 18.58$, $MSE = 3.43$, $p = 0.001$), reflecting the overall more negative waveform for the closed-class items and a significant effect of Anterior-Posterior ($F(1, 12) = 31.87$, $MSE = 6.95$, $p = 0.0001$), reflecting the overall more negative waveforms over anterior sites, with no interaction between these two factors ($F < 1$). The absence of the interaction occurred because the size of the difference between the closed- and open-class items is the same at anterior and posterior electrode sites.

In summary, closed-class items elicited a late negativity with a clear left-hemisphere preponderance. Over the left side of the scalp, the effect is evenly distributed over anterior and posterior electrode sites.

The open-class words also elicited a late negativity in the 350- to 550-msec epoch. However, due to the larger size and wider scalp distribution of the late negativity for the closed-class words, the amplitude analyses comparing the open- and closed-class words do not really capture the effect that is present for the open-class words.

Latency Analyses on the Late Negativities

We performed analyses of the latency of the negative shifts for open- and closed-class words in the 350- to 500-msec epoch. This window contains a clear peak negativity for the open-class words. There is a less pronounced peak for the closed-class words, where the negativity at some sites is steadily increasing without an obvious peak. Nevertheless, a negative maximum for the closed-class words can be discerned for most electrode sites in the 350- to 500-msec time window, which enables a latency comparison between the two vocabulary classes.

A statistically significant effect emerged in the analysis over all electrodes ($F(1, 12) = 150.20$, $MSE = 3193$, $p < 0.0001$), with the maximum for the open-class words occurring earlier than the maximum for the closed-class words (422 vs. 479 msec). However, the average latencies that emerge from this overall analysis might not be representative, given the topographic differences between the late negativities for the two word classes. The largest differences are seen over anterior and temporal electrode sites and to a lesser extent over temporal-parietal sites. We therefore performed an additional analysis with the factor Anterior Left/Right. This analysis again yielded a main effect of Vocabulary Class ($F(1, 12) = 94.58$, $MSE = 2054$, $p < 0.0001$), as well as a main effect of Anterior Left/Right ($F(1, 12) = 7.96$, $MSE = 620$, $p = 0.015$) but with no interaction between these two factors ($F < 1$). The mean latency for the open-class items at anterior-left and anterior-right sites was 413 and 407 msec, respectively. The maxima for the closed-class items were at 478 and 464 msec.

Given the topographical differences that were reported above for the analyses on the mean amplitudes, and in the light of both the morphological and the latency differences between the open- and closed-class words, it is reasonable to assume that the two vocabulary classes elicit different late negative components. The negativity for the open-class words is most prominent over the right side of the scalp and reaches its overall peak latency at about 420 msec. In contrast, the negativity for the closed-class words is most prominent over the left side of the scalp, with an overall peak latency at about 480 msec over anterior electrode sites.

Comparisons within the Open-Class Category

The separate waveforms for the Nouns, Verbs, and Adjectives show both commonalities and divergences (see Figures 2a and 2b). All three waveforms contain the early and the late negativity but to a different degree. The two negativities are most clearly seen for the Adjectives and the Verbs, with the Adjectives showing a slightly larger early negativity than the Verbs. The Nouns elicited a less

prominent early negativity, but the late negativity is strongly present in the waveform.

In terms of topography the three waveforms are generally consistent and point toward a qualitative distinction between the early and the late negativity. The early negativity is essentially confined to the anterior and anterior-temporal electrode sites, with a slight persistence over temporal and central sites and without any systematic lateralization. The late negativity is more widespread and tends to be larger over the right hemisphere.

Mean Amplitude Analyses on the Early Negativity

An analysis of the mean amplitude within the 260- to 350-msec latency range, over all 23 electrodes, yielded a main effect of Word Type ($F(1, 12) = 4.64$, $MSE = 14.23$, $0.05 < p < 0.10$) with Nouns being less negative than Adjectives and Verbs. The equal distribution of the negativity over the left and right side of the scalp was revealed by an analysis with Hemisphere as additional factor, which showed neither a main effect of this factor ($F(1, 12) = 1.63$, $MSE = 22.47$, $p = 0.22$) nor an interaction with Word Type ($F < 1$). The anterior predominance of the early negativity was demonstrated by an analysis with the additional factor Anterior/Posterior. This analysis resulted in a significant main effect of Word Type ($F(1, 12) = 4.64$, $MSE = 10.10$, $p < 0.02$) and Anterior/Posterior ($F(1, 12) = 28.25$, $MSE = 58.18$, $p = 0.0002$) with, however, no interaction between the two factors.

Peak Latency Analyses on the Early Negativity

Inspection of the waveforms for the three word types indicates that there is little difference in the peak latency of the early negativity. This is substantiated by a statistical analysis on the latency of the maximum negative peak in the 250- to 350-msec epoch, restricted to the anterior left and right electrodes. Although a significant effect was found for the factor Anterior Left/Right ($F(1, 12) = 5.95$, $MSE = 431$, $p = 0.031$), due to an overall 5-msec earlier onset over left electrode sites (309 versus 314), this analysis did not yield an interaction of Anterior Left/Right and Word Type ($F < 1$). Overall, Nouns, Verbs, and Adjectives elicited an early negativity over anterior electrode sites with a peak at about 312 msec, with no significant latency differences between the three word types.

Mean Amplitude Analyses on the Late Negativity

All three word types elicited a strong negative-going shift. The size of the component is essentially the same for Adjectives and Verbs and somewhat smaller for Nouns. In an analysis over all electrodes in the 350- to 550-msec epoch, a significant main effect of Word Type emerged ($F(1, 12) = 5.55$, $MSE = 10.52$, $p < 0.05$), with

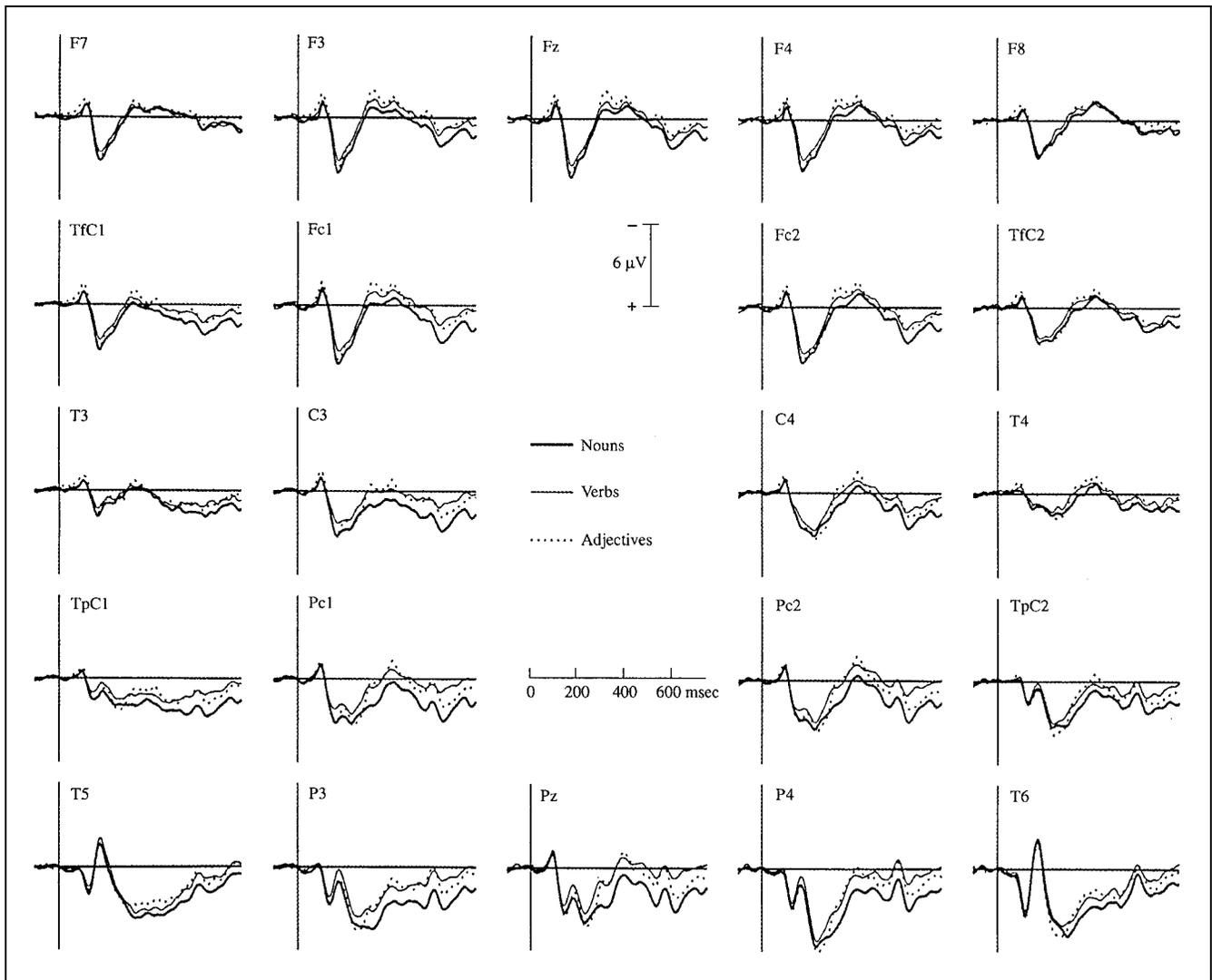
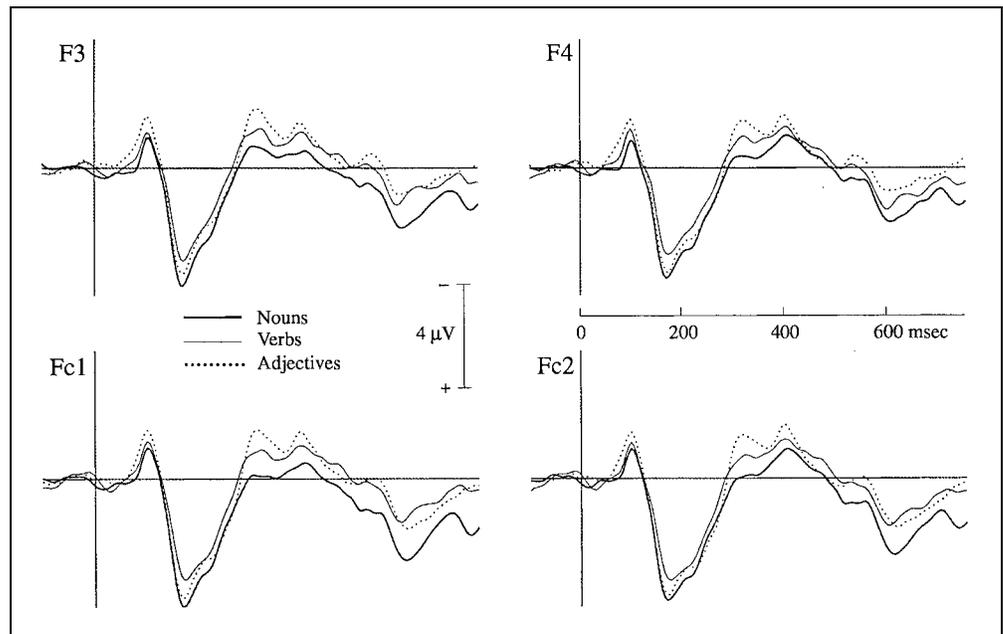


Figure 2a. Grand average waveforms ($N = 13$ subjects) elicited by nouns, verbs, and adjectives at 22 electrode sites.

Figure 2b. Grand average waveforms ($N = 13$ subjects) elicited by nouns, verbs, and adjectives at four anterior electrode sites.



Nouns overall less negative-going than Adjectives and Verbs. An analysis with the additional factor Hemisphere showed that the late component is more negative-going over the right side of the scalp (main effect of Hemisphere: $F(1, 12) = 5.45$, $MSE = 14.78$, $p = 0.011$), equally so for all word types (interaction of Word Type with Hemisphere: $F < 1$).

Peak Latency Analyses on the Late Negativity

Visual inspection of the waveforms indicates that the peak of the late negativity is essentially the same for the three word types. This was confirmed by an analysis of the latency of the peak negativity in the 350- to 450-msec epoch. The peak latency was at 395 msec for the Nouns, at 390 msec for the Adjectives, and at 399 msec for the Verbs. The small differences in latency did not yield any statistically significant result over all electrodes or over any of the subsets that have been reported on above.

In the light of the peak latencies, and given the quite widespread scalp distribution with a predominance over the right hemisphere, it is clear that the late negativity for the open-class words can be characterized as the N400 component (cf. Kutas & Hillyard, 1980).

Comparisons within the Closed-Class Category

The waveforms for the Articles, Conjunctions, and Prepositions contain a clear early and late negativity, with few visible differences between the three word types (see Figures 3a and 3b). The only exception here concerns the Prepositions, which appear to have a somewhat later peak latency and a slightly more enhanced overall negative amplitude.

Mean Amplitude Analyses on the Early Negativity

Neither an analysis over all electrodes in the 230- to 340-msec range nor an analysis including the factor Hemisphere resulted in a significant effect of Word Type (both F values for Word Type < 1 ; main effect of Hemisphere: $F(1, 12) = 3.18$, $MSE = 21.90$, $p = 0.099$). An analysis with the additional factor Anterior/Posterior yielded a strong main effect of this factor ($F(1, 12) = 50.28$, $MSE = 32.44$, $p < 0.0001$), reflecting the presence of the early negativity over anterior electrode sites and its absence over posterior sites.

The Anterior/Posterior analysis yielded neither a main effect of Word Type ($F < 1$), nor an interaction ($F(1, 12) = 1.03$). In addition, no Word Type effects were observed for either Anterior Left/Right (Word Type: $F < 1$; Anterior: $F(1, 12) = 2.66$, $MSE = 5.90$, $p = 0.13$; Interaction: $F < 1$) or Posterior Left/Right comparisons (Word Type: $F < 1$; Posterior: $F(1, 12) = 2.04$, $MSE = 18.64$, $p = 0.18$; Interaction: $F < 1$). In the light of the apparent increased negativity for Prepositions, two separate analy-

ses were performed over only the Anterior Left and the Anterior Right electrode sites. Neither of these analyses resulted in a significant effect for Word Type (AL: $F(1, 12) = 1.13$; AR: $F < 1$).

In sum, all three word types elicit essentially the same early negativity in terms of mean amplitude and topography. The negativity has a strong anterior distribution and is equally present over the left and right sides of the scalp.

Peak Latency Analyses on the Early Negativity

On the basis of the topography of the early negativity, the peak latency statistics were restricted to the anterior electrode sites. In an analysis with the factor Anterior Left/Right on the peak negative latencies in the 250- to 350-msec epoch, a main effect of Word Type ($F(1, 12) = 8.30$, $MSE = 1636$, $p < 0.05$) and of Anterior Left/Right ($F(1, 12) = 11.26$, $MSE = 1750$, $p = 0.006$) was obtained, with no interaction between these two factors ($F(1, 12) = 1.85$, $MSE = 722$, $p < 0.18$). The main effect of Anterior Left/Right reflects that the peak latencies occur earlier over left anterior sites. The main effect of Word Type indicates that there are differences among the three word types in their peak latencies. Articles had a mean peak latency of 277 msec, Prepositions, of 298 msec, and Conjunctions, of 278 msec. The peak latency for Articles is 15 msec earlier over the left side of the scalp (AL: 270 msec, AR: 285 msec). Prepositions reach their maximum negativity 23 msec earlier over left electrode sites (AL: 286 msec, AR: 309 msec), whereas Conjunctions peak 9 msec earlier (AL: 274 msec, AR: 283 msec).

The differences in the peak latencies of the three word types were further investigated in separate analyses of the Anterior Left and the Anterior Right electrode sites. Both analyses revealed a main effect of Word Type (AL: $F(1, 12) = 5.61$, $MSE = 676$, $p < 0.05$; AR: $F(1, 12) = 6.62$, $MSE = 1682$, $p < 0.05$). In separate simple effects analyses, the 15-msec difference between left and right electrode sites for the Articles was significant ($F(1, 12) = 5.74$, $MSE = 1113$, $p = 0.034$), as was the 23-msec difference for Prepositions ($F(1, 12) = 9.27$, $MSE = 1505$, $p = 0.01$). The 9-msec difference for Conjunctions was in the same direction but failed to reach significance ($F(1, 12) = 3.52$, $MSE = 577$, $p = 0.085$).

Pair-wise comparisons between word types, made separately for Anterior Left and Anterior Right electrodes, demonstrated that Articles and Conjunctions did not differ in their peak latencies but did peak significantly earlier compared to Prepositions. Over Anterior Left sites, the 16-msec difference between Articles and Prepositions is significant ($F(1, 12) = 8.38$, $MSE = 838$, $p = 0.01$), and the 12-msec difference between Conjunctions and Prepositions is marginally significant ($F(1, 12) = 3.98$, $MSE = 974$, $p = 0.069$). The 4-msec difference between Articles and Conjunctions does not approach signifi-

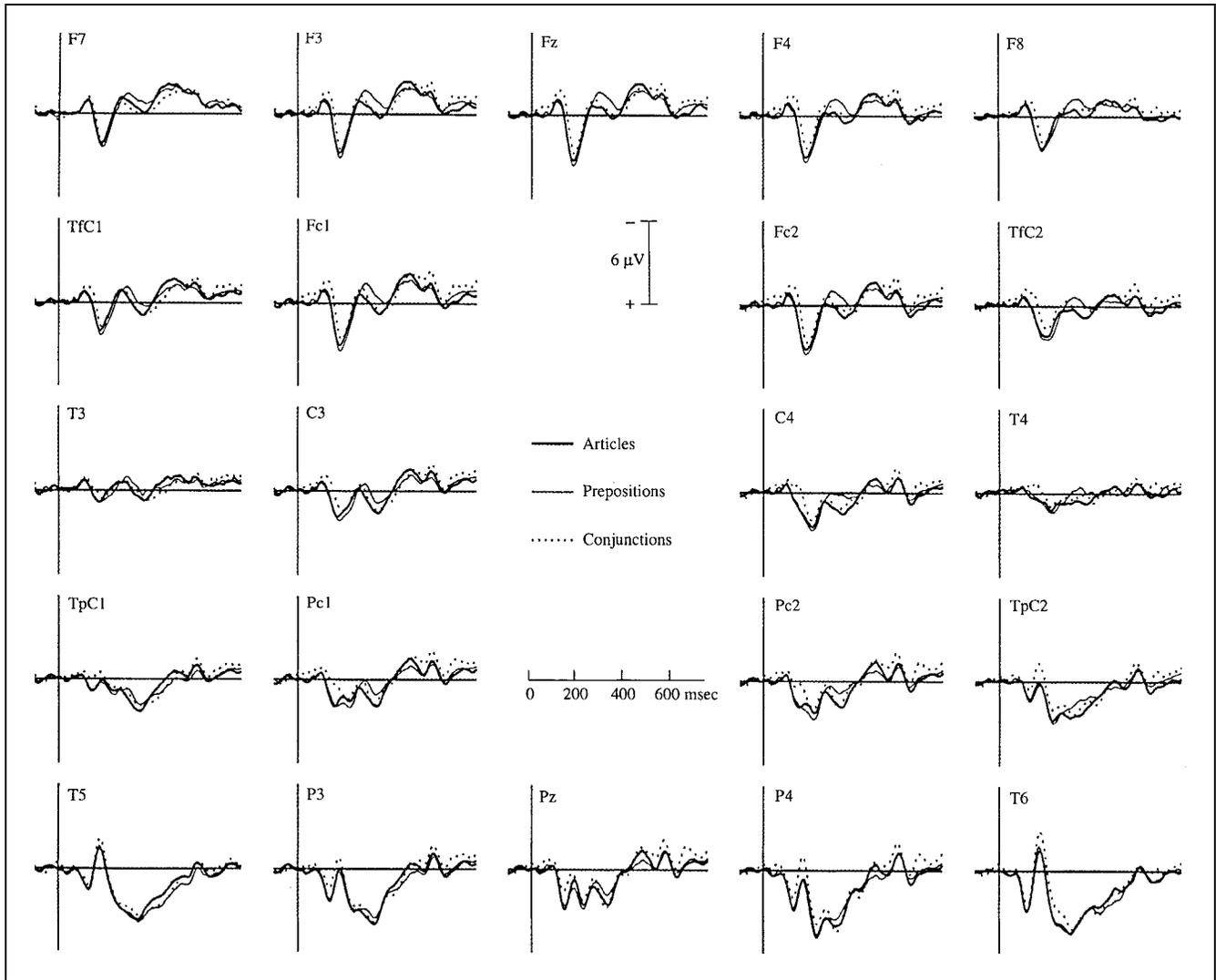
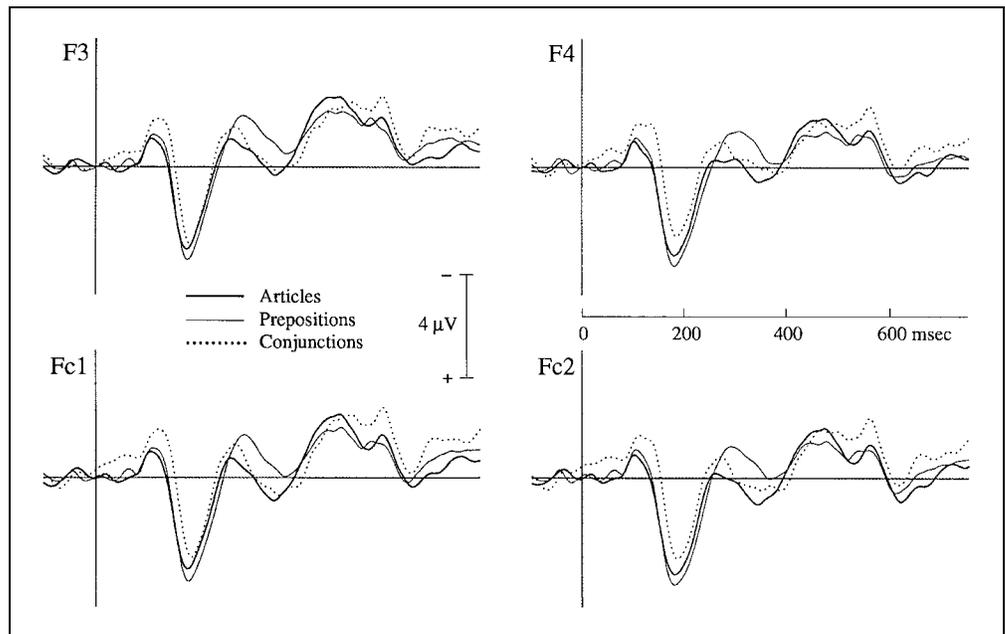


Figure 3a. Grand average waveforms (N = 13 subjects) elicited by articles, prepositions, and conjunctions at 22 electrode sites.

Figure 3b. Grand average waveforms (N = 13 subjects) elicited by articles, prepositions, and conjunctions at four anterior electrode sites.



cance ($F(1, 12) = 2.16, MSE = 215, p = 0.17$). The same pattern obtains over Anterior Right sites. The 24-msec difference between Articles and Prepositions is significant ($F(1, 12) = 6.12, MSE = 2434, p = 0.03$), as is the 26 msec difference between Conjunctions and Prepositions ($F(1, 12) = 13.86, MSE = 1321, p = 0.003$). The 3-msec difference between Articles and Conjunctions is not significant ($F < 1$).

In sum, in contrast to the mean amplitude analyses, the results of the analyses of peak latencies reveal differences between the three word types. Whereas the early negativity elicited by Articles and by Conjunctions does not differ in peak latency, for both word types the peak occurs significantly earlier than for Prepositions (by 19 msec on average). This difference holds equally over left and right anterior electrode sites, but all three word types have a consistently earlier peak over the left side of the scalp (16 msec on average).

Mean Amplitude Analyses on the Late Negativity

None of the analyses on the mean amplitude in the 350- to 550-msec epoch resulted in a significant main effect of Word Type (all p values were larger than 0.15 for the various configurational analyses). In other words, all three word types elicited a late negativity of similar extent and amplitude. An analysis with the factor Anterior/Posterior yielded a significant main effect of this factor ($F(1, 12) = 52.45, MSE = 10.26, p < 0.0001$), reflecting the fact that the late negativity is most prominent over anterior electrode sites. This holds equally for the three word types, as is demonstrated by the nonsignificant interaction of Word Type with Anterior/Posterior ($F(1, 12) = 1.05$). A separate analysis over the anterior electrodes, with the factor Anterior Left/Right, showed that the late component is significantly more negative over the left side of the scalp ($F(1, 12) = 4.59, MSE = 4.68, p = 0.053$), again equally so for the three word types (interaction of Word Type with Anterior Left/Right: $F(1, 12) = 2.86, p > 0.10$).

Latency Analyses on the Late Negativity

The three word types elicited a similar late negativity in terms of latencies: None of the analyses of the maxima in the 400- to 500-msec epoch resulted in a significant main effect of Word Type (all p values were larger than 0.15 for the various configurational analyses). In an analysis restricted to anterior electrode sites, a tenuous trend emerged for an earlier maximum over right electrode sites (AL: 463 msec, AR: 459 msec; $F(1, 12) = 3.74, MSE = 319, p = 0.077$). Collapsing over the anterior left and right electrodes yielded a negative maximum at 462 msec for Articles, at 456 msec for Prepositions, and at 464 msec for Conjunctions.

Effects of Lexical Frequency and Length

To what extent might the vocabulary-class effects be the result of lexical-statistical factors? As we discussed in the Introduction, there is a clear correlation between in particular the frequency but also the length of a word and its membership of the open or closed class. Before we can make any statements about category-specific effects, we need to examine and partial out possible lexical-statistical effects. The way in which we approached this issue in the current data set is by performing extensive analyses collapsing over all word types (i.e., ignoring vocabulary class), in which we compared waveforms that were averaged as a function of different frequency ranges, keeping word length constant, or as a function of different lengths, keeping frequency constant. The rationale behind this approach is that if the different early and late negative effects that we have reported for the two vocabulary classes are indeed related to the vocabulary class distinction, these differences should not emerge when the data are averaged as a function of lexical-statistical factors. For instance, in a comparison between different frequency ranges (e.g., high, middle, low) collapsing over all six word types, the early negativity that distinguishes between the closed- and open-class words should not be differentially modulated by the factor Frequency. Similarly, the early negative difference should not be modulated by the factor Length (e.g., comparing between all words with length 4, 5, and 6) within a particular frequency range and over several ranges.

Mean Amplitude and Peak Latency Analyses as a Function of Lexical Frequency

We performed two separate sets of amplitude analyses with the factor Frequency. In the first set of analyses we collapsed over all words with lengths 4, 5, and 6 and fractionated the factor Frequency into four levels, corresponding to the logarithmic bins 1.2 to 1.4, 2.2 to 2.4, 2.6 to 2.8, and 3.2 to 3.4. The maximum possible number of words in each of these bins was 31, 54, 56, and 36, respectively. These numbers ensure an adequate signal-to-noise ratio. In an additional analysis on these word lengths, we restricted the factor Frequency to two levels, namely, 1.2 to 1.4 and 3.2 to 3.4. In this way we maximized the possibility of finding an effect due to frequency. In the second set of analyses we collapsed over lengths 6, 7, and 8 and compared frequency bins 1.2 to 1.4 and 2.2 to 2.4 (with a maximum possible number of words of 32 and 36, respectively). Within each set of analyses we performed the same topographical analyses that were used in the previous sections, and each set was executed on the time window for both the early and the late negativity.

With respect to the effects of the factor Frequency on mean amplitude, we can be brief. At the level of the P2,

no effects were found.³ Within the 230- to 340-msec epoch (i.e., for the early negativity) not a single effect of this factor was obtained for any of the comparisons listed above, not even as a trend (the lowest p value was 0.21, with the majority of F values below 1). The same pattern of null effects was found for the 350- to 550-msec window.

The only effect of frequency that we did find was in the 230- to 340-msec window in an analysis of all words of only length 4, contrasting frequency bins 0.2 to 2.0, 2.2 to 2.8, and 3.0 to 3.8. Here the highest frequency words were associated with the smallest mean amplitude. This effect was also present in the 350- to 550-msec window, mainly as a trend. The small effect that we have in hand here is most likely due to a confounding with vocabulary class: The 0.2 to 2.0 and 2.2 to 2.8 frequency bins are composed almost entirely of open-class words, whereas the 3.0 to 3.8 bin has an open- to closed-class ratio of 1:2.

In two additional analyses comparing frequency bins 0.2 to 2.0 and 2.2 to 2.8, one of all words of length 5 and one of all words of length 6, no effects of frequency were obtained.

The same set of analyses that were performed on the mean amplitudes were also performed on the peak latency of the early and late negativities. None of the analyses of peak latency as a function of frequency reached significance.

Mean Amplitude Analyses as a Function of Word Length

In the analyses with the factor Length, we compared different lengths within particular frequency bins. Words with lengths 4 through 8 were compared within frequency bin 0.2 to 2.0, lengths 4 through 6 were compared within bin 2.2 to 2.8, and lengths 3 and 4 were compared within bin 3.0 to 3.8. As with the frequency analyses, the selection of particular lengths within particular bins was dictated by having a sufficient number of words over which to compute an averaged potential. For the five different word lengths in frequency bin 0.2 to 2.0, the maximum number of trials was 46, 25, 45, 32, and 31, respectively. For the three lengths in bin 2.2 to 2.8, the numbers were 33, 47, and 31. Finally, in bin 3.0 to 3.8 there were 31 words of length 3, and 49 of length 4.

We can again be brief in reporting the results of the full set of ANOVAs with the factor Length: No differential effects of length were found in any of the frequency bins for either the 230- to 340-msec epoch or the 350- to 550-msec epoch (the lowest p value was 0.20, with the majority of F values below 1).

Peak Latency Analyses as a Function of Word Length

The full set of configurational analyses used to test the effects of word length on the mean amplitude was also

used to assess the possible effects of length on the latency of the maximal peak of the early and late negativities. As with the previous analyses, no effects of length were found in the latency range of the early negativity. Basically the same holds for the late negativity, with one exception. In an analysis over all electrodes comparing words of three and four letters, in frequency bin 3.0 to 3.8, the shorter words were associated with a longer peak latency (417 versus 401 msec, $F(1, 12) = 5.30$, $MSE = 5413$, $p = 0.04$). This effect was also significant over anterior right electrodes (416 versus 386 msec, $F(1, 12) = 13.71$, $MSE = 1663$, $p = 0.03$) but not over anterior left sites (424 versus 408 msec, $F(1, 12) = 2.99$, $MSE = 2315$, $p = 0.11$).

DISCUSSION

To recap, the data reveal two latency windows within which negative-polarity brain potential modulations occur that can be related to lexical processing: an early epoch, around 230 to 350 msec, and a late epoch, from 350 to 550 msec. The early window is preceded by a reduced P2 component for the closed-class compared to the open-class words, which is most likely due to the stronger development of the early negativity for the closed-class words.

The early negative shift is elicited by both vocabulary classes. It has a strong anterior distribution and is present over both hemispheres, with a tendency toward greater amplitudes over the left side of the scalp. The morphology of the early negative shift for the open and closed class is very similar, with no clear topographical differences emerging as a function of vocabulary class.

There are, then, no indications in the early time window that we are dealing with an electrophysiologically different effect for the two lexical categories. The only clear difference that was obtained was in the peak latency of the negativity, which was significantly earlier for closed-class words (289 msec compared to 312 msec for open-class words, in an analysis restricted to anterior electrode sites) and which cannot be attributed to frequency and/or length effects. On the basis of the similar morphology and topography of the negative shift for the two vocabulary classes in the early time window, we suggest that here we are dealing with one electrophysiological effect as a manifestation of the same underlying process for both open- and closed-class words. This is not the case for the late time window. Here the morphology and topography of the negative shift elicited by open-class words is very different from that elicited by the closed-class words. We propose, therefore, that qualitatively distinct processes are operative in the later epoch. Moreover, based on the topographical differences between the early negativity and the late negativity for open-class words, we claim that these are two different effects. This contrasts with for instance the position of

Osterhout et al. (1997), who claim that in effect only one negativity is present.

When we compare our data with previously reported results, it is clear that the late negativities are a consistent finding. Both the N400-700 for closed-class words and the N400 for open-class words have been reported by most, if not all, studies. The situation for the early negativity is different: Its putative existence, sensitivity, and interpretation are contentious.

In a total of 20 ERP datasets in which open- and closed-class words were explicitly compared (which is the extent of the ERP data available to us), 14 of these (or 15 if our interpretation of the Van Petten and Kutas 1991 data is correct) have shown an early negativity for the closed-class category. Ten of these 14 also showed an early negativity for open-class words. We list the 20 datasets in Table 1 (some of the datasets are not, as yet, published; these waveforms can be obtained from the authors).

Our results on the early negativity both agree and contrast with the Neville et al. (1992), King and Kutas (1998), and Osterhout et al. (1997) studies. In agreement with the Neville et al. data, we observed an early negativity for closed-class words. However, in contrast to Neville et al., our data also show an early negativity for open-class words. In this respect, our data accord with the King and Kutas results; like us they report an early negativity for both closed- and open-class words. Moreover, the peak latencies that we obtained for the two classes (collapsing over word types and over anterior electrode sites) are very similar to the peak latencies that King and Kutas report: We obtained a peak latency of 289 and 312 msec for closed- and open-class words, respectively, which compares well with the 280 and 315 msec reported by King and Kutas in their overall analysis. However, unlike King and Kutas, we found no effect for frequency. Here, too, our work contrasts with that of Osterhout et al., who like King and Kutas reported that

Table 1. ERP Datasets on the Open- and Closed-Class Vocabularies

<i>Reference^a</i>	<i>Material</i>	<i>Dur.</i>	<i>SOA (msec)</i>	<i>Task</i>	<i>Neg. CC</i>	<i>Neg. OC</i>	<i>Remarks</i>
a	Prose	200	640-760	None	Yes	No	Not analyzed
b	Words	20	4500	Lex. dec.	Yes	Yes	Not analyzed
c	1. Sentences	150	500	None	Yes	No	Not analyzed
	2. Sentences	200	900	None	No	No	Not analyzed
	3. Sentences	132	700	None	No	No	Not analyzed
d	Sentences	200	600	Probe rec.	Yes?	No	Not analyzed
e	Sentences	200	700	(Non)sense	Yes	No	
f	Sentences	300	700	(Non)sense	No	No	
g	Words	500	2100-2200	Word det.	Yes	No	Also on pseudowords
h	Words	100	3500-4500	Lex. dec.	Yes	Yes	
i	Sentences	200	500	True/false	Yes	Yes	
j	1. Prose	300	700	None	No	No	Early effects on some word types
	2. Words	300	700	None	No	No	
k	1. Prose	400	800	None	Yes	Yes	1. Broca's aphasics
	2. Words	400	800	None	Yes	Yes	2. Broca's aphasics
	3. Prose	400	800	None	Yes	Yes	3. RH patients
	4. Words	400	800	None	Yes	Yes	4. RH patients
	5. Prose	400	800	None	Yes	Yes	5. Controls
	6. Words	400	800	None	Yes	Yes	6. Controls
l	Prose	400	800	None	Yes	Yes	

a Reference to literature. a: Kutas & Hillyard (1983); b: Garnsey, 1985; c: Kutas et al. (1988); d: Van Petten & Kutas (1991); e: Neville et al. (1992); f: Neville et al. (1993); g: Nobre & McCarthy, 1994; h: Pulvermüller et al. (1995); i: King & Kutas (1998); j: Osterhout et al. (1997); k: Ter Keurs, Brown, Hagoort & Stegeman (1999); l: Brown et al. (present data).

Key: *Dur.*: Stimulus presentation duration in msec; *SOA*: Stimulus onset asynchrony; *Neg. CC*: Early negativity for closed-class words; *Neg. OC*: Early negativity for open-class words. *Not analyzed*: An early negativity is present in the waveforms, but was not statistically analyzed by the authors. *RH*: Patients with lesions in the right hemisphere, without aphasia. *Lex. dec.*: lexical decision; *Probe rec.*: probe recognition; *(Non)sense*: meaningfulness judgment; *Word det.*: word detection; *True/false*: veridicality decision; *None*: No explicit task during or immediately following a stimulus.

lexical-statistical factors play a dominant role in the waveforms elicited as a function of word type. But the Osterhout et al. results stand out in one respect: These researchers did not report any early negativity akin to the N280 for closed-class words, not at left anterior frontal electrode sites or any other site. In this respect, their description of the data agrees most with that of Van Petten and Kutas (1991).

The split in the evidence for and against the existence of an early negativity obviously leaves some room for debate. One reason for the diverging results on the presence of an early negativity might be found in the extent to which multiple repetitions of the same word entered into the averaged waveform. Including multiple repetitions of the same word in the averaged ERP can have a detrimental effect on the possibility of observing early negativities. The well-established ERP repetition effect (e.g., Besson & Kutas, 1993; Rugg, 1995; Rugg & Doyle, 1992) emerges as a sustained positive-going deflection in the waveform for repeated items, most saliently in the latency range from 200 to 400 to 500 msec. The impact of the ERP repetition effect on the early negativity would, then, be to reduce the overall amplitude of this negativity, and potentially to entirely obscure it. This might be one of the reasons why several of the studies listed in Table 1 did not observe an early negativity to their open-class words.⁴ However, this argument does not hold for the closed-class words, because despite multiple tokens of these words entering into our averaged waveforms, we—and others—do observe a clear early negativity for these words. To date, the ERP repetition research has used open-class words as the stimuli with which to investigate memory effects. It is, therefore, not known whether the same kind of positive shift is obtained for the repetition of closed-class words. The fact that a clear early negativity is present in our waveforms for the closed-class category, as well as for the separate closed-class word types, might indicate that the repetition effect for closed-class words is either considerably weaker than for open-class words or in fact is absent. This issue needs to be resolved in further research.

In addition to, and perhaps more important than, repetition issues, there are considerable differences in stimulus selection and presentation, experimental procedure, and statistical analysis between the various reports. Nevertheless, it is exactly because of these differences that we would claim that there is a sufficient body of evidence in the literature to suggest that an early negativity in the 250- to 350-latency band is elicited as a function of lexical processing and that this negativity can be distinguished from the by now well-established N400 component. Despite methodological differences, a sizable number of datasets contain early negativities for either open- or closed-class words, or both. The data that we have presented are an example in point: Both in the waveforms by category, collapsed over word type, and

in the separate waveforms by word type, a quite prominent early negativity is present in all cases.

We now turn to a further consideration of the effects of lexical factors on the elicitation and morphology of the early negativity.

The Effect of Lexical Factors

Only a minority of the 20 studies listed in Table 1 explicitly compared the effects of frequency and length on the brain potentials elicited by open- and closed-class words, and as already noted above, the pattern of results is not consistent. A number of reasons can be put forward about why such inconsistent results have been obtained.

First, the notorious natural confound of in particular frequency with vocabulary class makes it difficult to independently test for lexical-statistical effects. Specifically, it is in practice almost impossible to find sufficient tokens in both vocabulary classes to statistically test for the main effects and interaction of vocabulary class with a sufficiently distributed factor of frequency and/or length. This is certainly the case with brain-potential data, where to achieve an acceptable signal-to-noise ratio a minimum of 20 to 40 separate items in a condition is usually required. ERP researchers have adopted different approaches to deal with this problem. Neville et al. (1992), for instance, analyzed frequency and length effects only for open-class words. They argued that because no early negative effects were found in these analyses, the early negativity observed for the closed-class words was not dependent on lexical-statistical factors. However, to assess the effect of frequency Neville et al. split their stimuli into two frequency classes, above and below 90 occurrences per million, as indicated by the Kucera and Francis (1967) database. This is a very broad split that is not representative of the sensitivity to frequency that has been reported in the reaction-time literature (see, for example, Coltheart 1987) or representative of the frequency distribution within and between the open and closed classes.

The second approach in the open- and closed-class ERP literature to assess the impact of lexical-statistical factors has been the regression analysis procedure chosen by King and Kutas (1998) and Osterhout et al. (1997). Here, the factor of vocabulary class is not present as such in the analysis, but a frequency and/or length variable is defined over the entire set of words. This variable is used for the regression with peak negative latency.

It should be noted that the regression analyses performed by King and Kutas (1998) and by Osterhout et al. (1997) differ in two potentially important ways. First, King and Kutas distinguished between 10 word types: nouns, verbs, and adjectives in the open-class and infinitival “to”, definite articles, indefinite articles, noun phrase prepositions, verb phrase prepositions, conjunctions, and forms of the verb “to be” in the closed class. Osterhout

et al. do not provide a complete listing of the word types that entered into the averages for the open- and closed-class categories, but in the separate analyses by word type they fractionate the open-class category into auxiliaries, nouns, and verbs and the closed-class category into articles, prepositions, and pronouns. There are, then, considerable differences in the kinds of words that entered into the two analyses.

Second, King and Kutas (1998) performed their analyses on the data from individual electrode sites, whereas Osterhout et al. (1997) performed their analyses on peak latencies averaged over 10 lateral electrode sites, distributed from frontal to occipital sites. Although it is unclear exactly what consequences follow from these differences, the separate sets of words entering into the analyses in the two studies, as well as the disparate number and placement of electrode sites used, might underlie the substantial difference between the amount of explained variance in the two studies. The King and Kutas experiment included four frequency bins that contained both open- and closed-class words. The regression analysis on one left-frontal electrode site (F7) accounted for 92% of the variance in peak latency of the early negativity (two other anterior-temporal electrodes showed a similar pattern). The two separate analyses reported by Osterhout et al. for their prose experiment (which compares best with the King and Kutas sentence material) each account for some 46% of the variance.

In the frequency and length analyses that we performed on our data, we also cannot escape from the confound with word type. However, in contrast to the Van Petten and Kutas (1991), the Neville et al. (1992), the Osterhout et al. (1997), and the King and Kutas (1998) studies that all used American English and based their frequency analyses on the Brown corpus (cf. Francis & Kucera, 1982; Kucera & Francis, 1967), we used Dutch stimuli and based our frequency analyses on the Celex corpus (cf. Baayen et al., 1993). The Celex corpus is a considerably more extensive, up to date, and representative lexical database than the Brown corpus. The database contains some 42 million tokens, mainly derived from texts that were printed between 1970 and 1988. By contrast, the Brown corpus contains only some 1 million tokens, based on texts that appeared in 1961.⁵ The large size of the Celex corpus presents a much more fine-grained distributional pattern than can be found in a 1 million token corpus, and it enabled us to perform a series of stepwise frequency analyses in which we held word length constant, capturing a wider frequency range than the studies mentioned above, particularly with respect to words belonging to the open-class category. Despite this extensive frequency analysis, we did not observe any significant modulation of the early negativity in terms of its amplitude or in terms of its latency. Most importantly, irrespective of the frequency and length bin that we focused on, the early negativity was invariably present in the waveform. In other words, in

our data neither the existence nor the morphology of the early negativity was affected by the factors of frequency or length. We propose, therefore, that the elicitation of the early negativity is not a primary function of lexical-statistical factors but is one of the standard electrical brain responses during lexical processing. At the same time, however, we agree with Van Petten and Kutas (1991), Osterhout et al. (1997), and King and Kutas (1998) that the early negativity cannot serve as a qualitative marker of vocabulary class. The effect is present for both categories and cannot be distinguished in terms of its shape or its distribution over the scalp.

In summary, the inconsistencies in the ERP literature could have arisen as a function of the validity of the frequency counts, the disparate materials entering into the analyses, the different frequency ranges used in the analyses, the different (number of) electrode sites entering into the analyses, or, in all likelihood, some interaction of two or more of these factors. Further research is required before the actual impact of these (and perhaps other) factors can be described in any detail. Nevertheless, our position based on the currently available ERP literature is that the existence of an early negativity during lexical processing that is separate from the N400 can be accepted with a reasonable degree of confidence.

A Functional Interpretation of the Early Negative Effects

We interpret the early negativity as the temporally first electrophysiological manifestation of the availability of categorical information from the mental lexicon. We hypothesize that the difference in peak latency between the closed- and the open-class words reflects a difference in the relative timing with which lexical-categorical information becomes available during on-line word processing. Our suggestion is that the lexical information that is associated with closed-class items is more rapidly retrieved than the information related to open-class items. This difference could emerge because of a higher baseline activation level of the closed-class lemmas, which in turn could be related to the restricted and unchanging number of items in this vocabulary class, together with their ubiquitous usage in language. These intrinsic features of closed-class words enable a more efficient lexical selection process, which results in the more rapid retrieval of the information associated with closed-class words, in comparison with the selection and retrieval processes for open-class words.⁶

Note that our functional interpretation does not imply that we are postulating differences in the speed of lexical access. By lexical access we mean the on-line processes that are involved in analyzing the physical input and connecting this analysis with the mental lexicon. The product of lexical access is the activation of a subset of lexical representations that match the physical input to a sufficient degree (cf. Coltheart, Curtis, Atkins, &

Haller, 1993; Forster, 1987; Henderson, 1987). We assume that word-category information is a product of lexical access and not part of the access process itself. Hence, the modulation of the early negativity as a function of vocabulary class is not directly related to access processes. This interpretation is supported by the absence of frequency effects on the early negativity. This absence is more compatible with the involvement of lexical selection than access processes. Our claim, then, is that following access to the lexicon, the lexical selection process results in an earlier availability of the lexical-categorical information associated with closed-class items, relative to open-class items. This is, by hypothesis, what is being reflected in the differences in peak latency of the early negativity for the closed- and open-class categories: It is a reflection of the presence of this information, not its initial activation.

A Functional Interpretation of the Late Negative Effects

In the time window of 350 to 550 msec both the closed- and the open-class words elicited a negative shift. In contrast to the early negativity, the late effect is clearly qualitatively different as a function of lexical category. Although both categories show a broadly distributed negativity over anterior to posterior electrode sites, the waveform for the closed-class words shows a sustained and increasing negative shift that is most prominent over the left hemisphere, whereas the open-class words show a more sharply defined negative component that is strongest over the right hemisphere, with a clear maximum at around 400 msec.

On the basis of its morphology, its right-hemisphere preponderance, and its peak latency at 400 msec, we classify the late negativity for the open-class words as the N400 component. It is by now firmly established that in the context of written or spoken language stimulation, the N400 is a reflection of real-time semantic processing (cf. Kutas & Van Petten, 1994; Osterhout & Holcomb, 1995). The extensive literature on the N400 has demonstrated that each open-class word elicits an N400. Although the exact processing nature of the N400 is still a matter of debate, it has been argued that the modulation of N400 amplitude is a reflection of lexical-semantic integration processes, related to the insertion of word meanings into the message-level representation of the sentential and discourse context in which they occur (e.g., Brown & Hagoort, 1993, 1999; Kellenbach & Michie, 1996; Swaab, Brown, & Hagoort, 1997).

The fact that we did not observe any frequency effects at the level of the N400 amplitude requires brief discussion in the light of several detailed reports of such effects (e.g., Van Petten & Kutas, 1991). In our analyses, as was the case for the frequency analyses of Neville et al. (1992), King and Kutas (1998), and Osterhout et al. (1997), we averaged over words irrespective of their

position in a sentence. This was required to achieve an adequate signal-to-noise ratio for the statistical analyses. It has been shown that the effect of frequency on the N400, which takes the form of a reduction in amplitude with increasing frequency, is markedly reduced as a function of word position. The frequency effect is only observed for the first three or so open-class words of a sentence, after which it is no longer present. Words further "downstream" in a sentence no longer show any modulation of the N400 component as a function of frequency (cf. Van Petten & Kutas, 1990). Given that we collapsed over word positions in our analyses, it is likely that the modulation of the N400 by frequency at early positions in a sentence was obscured by the absence of this effect at later positions.

The slow ascending negative flank of the late negativity for the closed-class words, reaching its maximum just before the presentation of the following word, is very suggestive of a classical preparation response in the electrophysiological literature: the contingent negative variation, or CNV (cf. Hillyard, 1973). The CNV is typically observed in situations where subjects are expecting a relevant stimulation event to occur at a particular time. This latter situation might be underlying the negativity for the closed-class words. Van Petten and Kutas (1991) observed a slow ascending late frontal negativity for closed-class words in their sentence comprehension study, similar to the effect that we report here. The same effect was also reported by Neville et al. (1992), Osterhout et al. (1997), and King and Kutas (1998). Van Petten and Kutas labeled the effect a CNV and speculated that the closed-class words serve as a syntactic signal to the reader that a new head of a constituent is imminent. The expectation of this important event for the comprehension process is reflected in a CNV that develops contingent upon the processing of a closed-class word. This hypothesis fits well with the presumed syntactic function of closed-class words. They are relatively devoid of meaning (as is underscored by the absence of an N400 response to these words in our study), and their primary role lies in determining the syntactic structure of a sentence. Under the assumption that readers engage in some form of prediction about upcoming structures, a CNV response to closed-class words makes sense.

It should be noted, however, that it is at present unclear whether this kind of predictive process necessarily concerns specifically syntactic structures. It is also possible that the closed-class words serve to signal a much more general expectation, namely, quite simply that the next word is likely to be a meaningful word. This expectation is still predicated on the syntactic nature of the closed-class words, but on-line predictions as to what structures might follow no longer have to be postulated. Instead, the distinction is between meaning-devoid and meaning-bearing words, which maps quite naturally onto the closed- and open-class distinction without having to invoke parsing-specific prediction processes. For these

reasons, we prefer this latter interpretation of the CNV elicited by the closed-class words. However, the currently available data do not directly address the precise functional nature of the late negativity for closed-class words, and more work is called for.

CONCLUSION

Our study adds to the literature on event-related brain potential manifestations of on-line lexical processing. We have observed three separable components in the ERP waveform, each of which can be related to the comprehension process: (1) an early negativity, elicited by open- and closed-class words alike (2) the classical N400 component, elicited only by open-class words in the present study, and (3) a late slow negative frontal shift, elicited only by closed-class words.

The N400 to open-class words in our data fits well with the literature and can be related to semantic integration processes. The late negative shift also accords with several studies and is presumed to relate to some kind of on-line expectancy process. It seems warranted to suggest that in addition to these two by now well-attested ERP effects, the early negativity can be similarly described as one of the components of the brain's electrical activity that can be reliably related to on-line written language comprehension. The challenge now lies in further determining the functional nature of these electrophysiological markers of language comprehension.

METHOD

Subjects

Thirteen students (mean age 25 years, six males) from the University of Nijmegen were paid for their participation in the experiment, after having given informed consent according to the Declaration of Helsinki. All subjects had normal or corrected-to-normal vision and were right handed according to an abridged Dutch version of the Oldfield Handedness Inventory (Oldfield, 1971). Four subjects reported left-handed relatives in their immediate family. According to their responses on a questionnaire, none of the subjects had any neurological impairment, had experienced any neurological trauma, or used neuroleptics.

Materials

A simple, fairy-tale-like story was selected for visual presentation. The text contained 151 sentences. The average length of the sentences was 10 words (range 4 to 16). Active sentences made up 92% of the story. The text contained a total of 1540 words, including nouns, verbs, adjectives, adverbs, auxiliaries, conjunctions, prepositions, numerals, articles, and pronouns. From this set of words, 439 open-class words (202 nouns, 86 adjectives,

151 verbs) and 398 closed-class words (212 articles, 115 prepositions, 71 conjunctions) were included in the analyses reported below. Auxiliaries were excluded from analysis. Although this word type is usually classified under the closed-class category, many auxiliaries represent an intermediate form in terms of their lexical semantics (consider, for example, the extended meaning of the verb *have*). Sentence-initial and sentence-final words were also excluded from analysis to avoid contamination from start-up and wrap-up effects (cf. Hagoort, Brown, & Groothusen, 1993; Kutas & King, 1996).

The open-class words that were selected for analysis ranged in length from 3 to 12 letters. The selected closed-class words ranged from 2 to 5 letters. Frequency of occurrence for the open-class words was between 0.0 and 4.3 on a log frequency scale, and from 2.42 to 4.44 for the closed-class words. The frequencies were obtained from a Dutch frequency-coded corpus based on over 42 million tokens (Baayen et al., 1993). For 84% of the nouns, 89% of the verbs, and 77% of the adjectives, imagability was established on the basis of Dutch imagability norms from a corpus of 6000 words (van Loon-Vervoon, 1985). On a seven-point scale, the mean imagability for nouns was 5.65 (range 2.2 to 6.9), for verbs 5.27 (range 2.93 to 6.6), and for adjectives 5.38 (range 2.83 to 6.57).

The test story was preceded by a short practice story to familiarize the subjects with the experimental procedure.

Procedure

All subjects were tested in a quiet room, seated in a comfortable reclining chair. The text was presented visually, word by word, for 400 msec in a 10- by 2.5-cm window on a high-resolution PC screen that was covered by a black nonreflecting shield. Each word was followed by a 400-msec blank-screen interval (i.e., the stimulus-onset asynchrony was 800 msec). This relatively slow presentation rate was used to avoid analysis problems due to overlapping ERP components elicited by adjacent words.⁷ Sentence-initial words were written with an initial capital letter; all other letters were in lowercase. Sentence-final words appeared together with a period, question, or exclamation mark, as appropriate. A sentence was followed by a blank screen interval of 2200 msec. During this period subjects were told that they could blink their eyes (which was not allowed during sentence presentation, to avoid ocular artifacts in the EEG signal). Viewing distance was between 70 and 80 cm, and the stimuli subtended a vertical visual angle of approximately 3°.

Subjects were instructed to read the text carefully for comprehension, with no additional task. They were also instructed to move as little as possible and to keep their eyes fixated on the window in the center of the screen.

The test story was presented in four blocks of ap-

proximately 5 min of stimulation. To make sure that the subjects were actually reading the text, the experimenter would at various moments during the breaks between the blocks question the subjects about the story. A session lasted approximately 2 h, including electrode application and removal.

EEG Recording

Continuous EEG was recorded from 29 tin electrodes attached to an electrode cap, each referred to the left mastoid. Activity over the right mastoid was actively recorded on a separate channel to determine whether there were any effects of the experimental variables on the mastoid recordings. No such effects were obtained. We will, therefore, not report on the right mastoid recording. Eighteen electrodes were placed according to the International 10–20 system (Jasper, 1958), including sixteen electrodes over homologous positions over left and right prefrontal (Fp1-Fp2), frontal (F7-F8, F3-F4), posterior-temporal frontal (T3-T4, C3-C4), posterior-temporal parietal (T5-T6, P3-P4), and occipital (O1-O2) regions, as well as two electrodes over midline frontal (Fz) and parietal (Pz) sites. The remaining eleven electrodes were placed over nonstandard sites. A prefrontal midline electrode (Fz') was placed halfway between Fp1-Fp2 and Fz. Six electrodes were evenly spaced halfway between standard frontal and posterior-temporal frontal electrode sites over the left and right anterior to temporal frontal regions (Tf1-Tf2, Tfc1-Tfc2, Fc1-Fc2). Four electrodes were evenly spaced halfway between standard posterior-temporal frontal and posterior-temporal parietal sites over the left and right anterior to temporal parietal regions (Tpc1-Tpc2, Pc1-Pc2). Vertical and horizontal eye movements were monitored bipolarly via sub- and

supraorbital electrodes and left and right external canthal montages, respectively. The ground electrode was placed on the standard midline central site Cz. Figure 4 presents a flat-projection of the electrode configuration with site labels.

The EEG and EOG recordings were amplified by a Neurotop MME-3100 multichannel bioelectric amplifier system, using a band-pass filter of 0.016 to 35 Hz. Impedances were kept below 5 k Ω . The EEG signal and word-type trigger codes for subsequent off-line averaging and data analysis were digitized on-line with a sample frequency of 200 Hz. No on-line signal-rejection procedures were used.

EEG Analysis

Trials that were contaminated by eye movements, other forms of muscular activity, excessive electrode drift, or amplifier saturation were removed prior to averaging. In total, 16% of the trials were identified as being contaminated, of which 51 were open-class and 49 were closed-class items. In addition, data recorded from electrode sites Fp1, Fp2, Tf1, Tf2, O1, and O2 were not analyzed. Electrodes Fp1 and Fp2 were excluded because of the smeared morphology of the waveforms over these sites compared to other anterior sites, presumably related to their very frontal position (just above the eyes). The waveforms from Tf1 and Tf2 were excluded because of their essentially flat morphology. Electrodes O1 and O2 were excluded because the waveforms from these sites were heavily dominated—as is often the case—by exogenous potentials (in particular the N1-P2 complex), which in our data tended to obscure in particular any early effects in the signal. Figure 5 gives the waveforms for the eye channels, as well as Fp1 and Fp2.

Figure 4. Flat projection of the full electrode configuration.

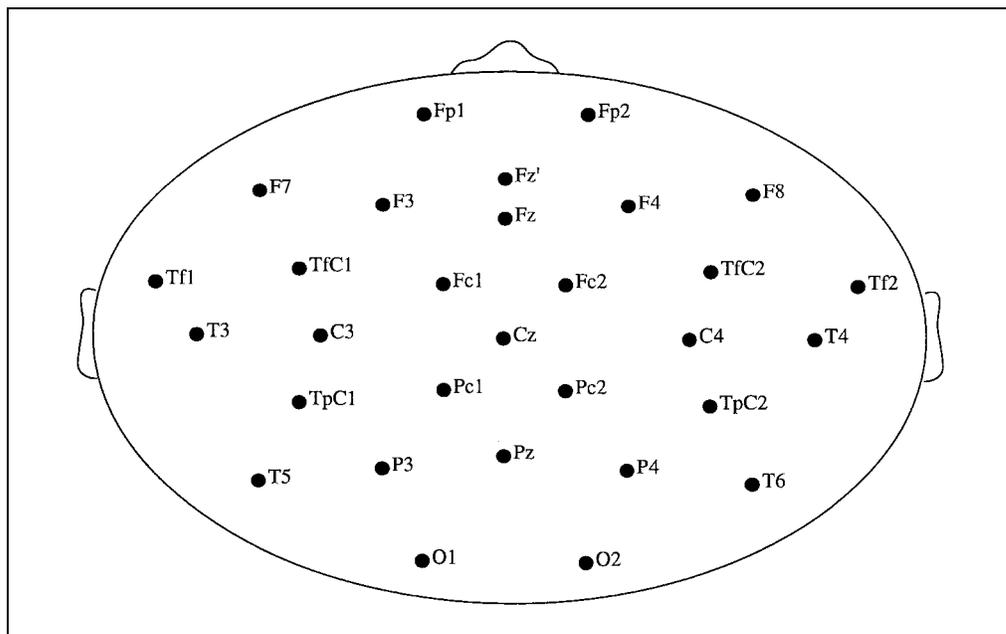
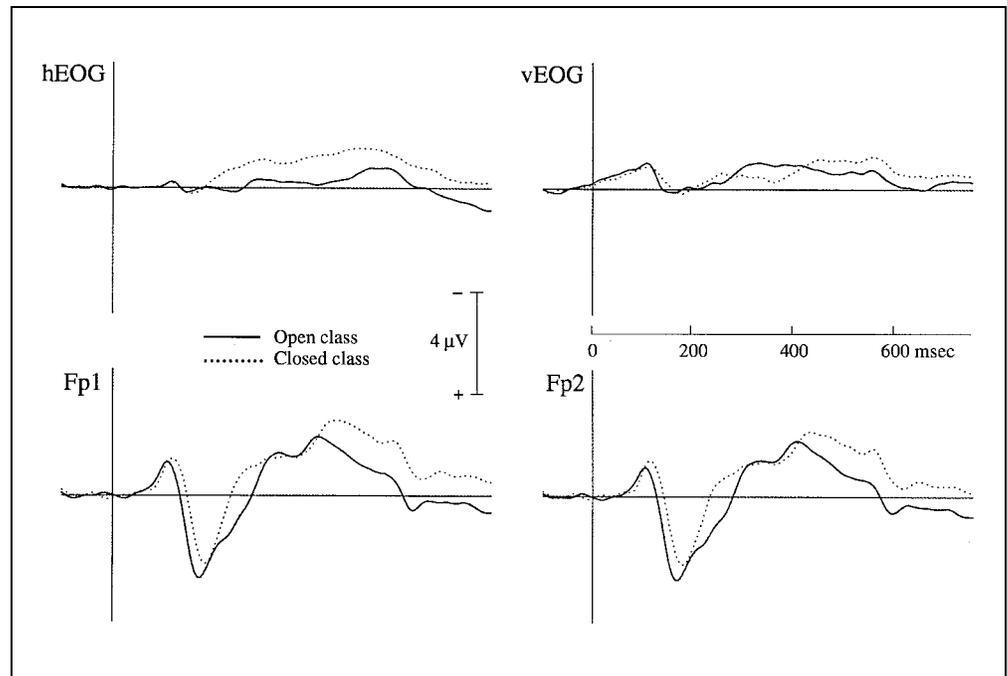


Figure 5. Averaged eye-movement activity (N = 13 subjects) elicited by the open- and closed-class categories, as well as grand average waveforms at frontal sites Fp1 and Fp2.



In all, then, 23 electrodes remained for analysis: 3 over midline sites and 10 over each of the hemispheres. In the analyses reported below, different subsets of electrodes were taken together to investigate the topographical distribution of the ERP effects. For purposes of brevity, we use the following labels: Anterior Left (AL; electrodes F7, F3, Tfc1, Fc1), Anterior Right (AR; electrodes F8, F4, Tfc2, Fc2), Posterior Left (PL; electrodes Tpc1, Pc1, T5, P3), and Posterior Right (PR; electrodes Tpc2, Pc2, T6, P4). In analyses with the factor Hemisphere, AL and PL, and AR and PR are collapsed.

On the basis of visual inspection of the waveforms, several latency windows were selected for statistical analysis. The separate epochs are specified in the results section. Analyses were performed on either mean amplitude values or on peak latencies. Both values were determined via computerized procedures: First, for each subject the waveform elicited by each word was normalized, point by point, with respect to the averaged activity in the 100 msec preceding that word. Then, in selected epochs, the mean amplitude was computed in the waveform for each vocabulary class or word type and for each subject. Peak latency was defined by the latency at which the maximal amplitude value occurred within the selected epoch. The data were entered into repeated measures analyses of variance with Vocabulary Class (open versus closed class) and Electrode Site as factors. Several subanalyses focused separately on either the open- or the closed-class category. These analyses each contained three levels of the factor Word Type (Nouns, Verbs, and Adjectives or Articles, Conjunctions, and Prepositions, respectively). The Greenhouse-Geisser correction was applied when evaluating effects with more

than 1 degree of freedom in the numerator (Greenhouse & Geisser, 1959; see Winer, 1971). The adjusted degrees of freedom and *p* values will be reported. Effects involving the factor Electrode Site are only reported for significant interactions with Vocabulary Class or Word Type. For these interactions, analyses of variance were computed after performing a *z*-score normalization procedure to equalize the overall mean amplitudes across experimental conditions. This procedure is described by Rösler, Heil, and Glowalla (1993) and is equivalent to the normalization procedure suggested by McCarthy and Wood (1985).

Acknowledgments

This research was supported in part by grant 400-56-384 from the Netherlands Organization for Scientific Research. We thank Peter Praamstra and Dick Stegeman (Dept. of Clinical Neurophysiology, Institute of Neurology, University Hospital Nijmegen) for their support.

Reprint requests should be sent to Colin Brown, "Neurocognition of Language Processing" Research Group, Max Planck Institute for Psycholinguistics, Wundtlaan 1, NL-6525 XD Nijmegen, The Netherlands.

Notes

1. Scarcity was defined by the authors as the "linear transformation of log word frequency normalized to the size of the corpus. . . ." (p. 2).
2. From 550 msec onward, the off-potential elicited by the removal of the stimulus obscures the descending flank of this late negativity.
3. This was also the case for analyses on peak latency. The same holds for all analyses on the effect of length, with one

exception: Over anterior-right electrode sites, words with length 6 peaked earlier in frequency bin 2.2 to 2.8 than in bin 0.2 to 2.0 (184 vs. 195 msec, $F(1, 12) = 4.80$, $MSE = 4.71$, $p = 0.049$).

4. Note that in our analyses we included at most two tokens of the same open-class word type in the average waveform. If a repeated word was included, this was always the next presentation after the first one (with a default separation of at least eight intervening other words).

5. It is not clear whether the different sampling periods of the Celex and Brown databases have any implications for the validity of the frequency counts today. Although the estimates for the closed-class words probably do not suffer too much from the shortcomings of the Brown corpus, it is a matter for research whether this also holds for the open-class words.

6. This position bears some resemblance to the proposals of Bradley and colleagues (e.g., Bradley et al., 1980). However, in contrast to Bradley and others, we do not invoke separate storage systems or dedicated lexical access routes for closed-class items.

7. Although the presentation format and the stimulation rate differ from standard reading, previous research has shown that closely similar language-related brain potentials are elicited by both slow and fast (four words per second) single-word visual stimulation, as well as by naturally produced connected speech (cf. Hagoort & Brown, 1998; Kutas, 1993, 1997). Nevertheless, we acknowledge that the reading situation in the present experiment only approximates standard reading.

REFERENCES

- Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database* (CD-ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.
- Besson, M., & Kutas, M. (1993). The many facets of repetition: A cued-recall and event-related potential analysis of repeating words in same versus different sentence contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1115-1133.
- Bradley, D. C., & Garrett, M. F. (1983). Hemispheric differences in the recognition of closed- and open-class words. *Neuropsychologia*, *21*, 155-159.
- Bradley, D. C., Garrett, M. F., & Zurif, E. B. (1980). Syntactic deficits in Broca's aphasia. In D. Caplan (Ed.), *Biological studies of mental processes* (pp. 269-286). Cambridge, MA: MIT Press.
- Brown, C. M., & Hagoort, P. (1993). The processing nature of the N400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, *5*, 34-44.
- Brown, C. M., & Hagoort, P. (1999). On the electrophysiology of language comprehension: Implications for the human language system. In M. Crocker, M. Pickering, & C. Clifton (Eds.), *Architectures and mechanisms for language processing*. Cambridge: Cambridge University Press.
- Chiarello, C., & Nuding, S. (1987). Visual field effects for processing content and function words. *Neuropsychologia*, *25*, 539-548.
- Coltheart, M. (Ed.) (1987). *The psychology of reading: Attention and performance XII*. London: Erlbaum.
- Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*, 589-608.
- Forster, K. I. (1987). Form-priming with masked primes: The best match hypothesis. In M. Coltheart (Ed.), *The psychology of reading: Attention and performance XII* (pp. 127-146). London: Erlbaum.
- Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston, MA: Houghton Mifflin.
- Friederici, A. D. (1985). Levels of processing and vocabulary types: Evidence from on-line comprehension in normals and agrammatics. *Cognition*, *19*, 133-166.
- Garnsey, S. M. (1985). *Function words and content words: Reaction time and evoked potential measures of word recognition*. Unpublished Ph. D. thesis, Cognitive Science Technical Report URCS-29, University of Rochester.
- Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19-76). London: Academic Press.
- Gordon, B., & Caramazza, A. (1982). Lexical decision for open- and closed-class words: Failure to replicate differential frequency sensitivity. *Brain and Language*, *15*, 143-160.
- Gordon, B., & Caramazza, A. (1985). Lexical access and frequency sensitivity: Frequency saturation and open/closed class equivalence. *Cognition*, *21*, 95-115.
- Greenhouse, S., & Geisser, D. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95-112.
- Hagoort, P., & Brown, C. M. (1998). Semantic and syntactic ERP effects of listening to speech compared to reading. (submitted)
- Hagoort, P., Brown, C. M., & Grootjusen, J. (1993). The Syntactic Positive Shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, *8*, 439-483.
- Henderson, L. (1987). Word recognition: A tutorial review. In M. Coltheart (Ed.), *The psychology of reading: Attention and performance XII* (pp. 171-200). London: Erlbaum.
- Hillyard, S. A. (1973). The CNV and human behavior: A review. *Electroencephalography and Clinical Neurophysiology Supplement*, *33*, 161-171.
- Jasper, H. H. (1958). Report to the committee on methods and clinical examination in electroencephalography. Appendix: The ten-twenty system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 371-375.
- Kellenbach, M. L., & Michie, P. T. (1996). Modulation of event-related potentials by semantic priming: Effects of color-selective attention. *Journal of Cognitive Neuroscience*, *8*, 155-173.
- King, J. W., & Kutas, M. (1998). Neural plasticity in the dynamics of human visual word recognition. *Neuroscience Letters*, *244*, 1-4.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, R. I.: Brown University Press.
- Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Processes*, *8*, 533-572.
- Kutas, M. (1997). Views on how the electrical activity that the brain generates reflects the functions of different language structures. *Psychophysiology*, *34*, 383-398.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203-205.
- Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory and Cognition*, *11*, 539-550.
- Kutas, M., & King, J. W. (1996). The potentials for basic sentence processing: Differentiating integrative processes. In T. Inui & J. L. McClelland (Eds.), *Information integration in perception and communication: Attention and performance XVI* (pp. 501-546). Cambridge, MA: MIT Press.
- Kutas, M., & Van Petten, C. K. (1994). Psycholinguistics elec-

- trified: Event-related brain potential investigations. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 83–143). San Diego: Academic Press.
- Kutas, M., Van Petten, C., & Besson, M. (1988). Event-related potential asymmetries during the reading of sentences. *Electroencephalography and Clinical Neurophysiology, Supplement, 39*, 325–330.
- Linebarger, M. C., Schwartz, M. F., & Saffran, E. M. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition, 13*, 361–392.
- Loon-Vervoorn, van, W. A. (1985). *Voorstelbaarheidswaarden van Nederlandse woorden*. Lisse: Swets & Zeitlinger.
- McCallum, W. C. (1988). Potentials related to expectancy, preparation and motor activity. In T. W. Picton (Ed.), *Human event-related potentials: Handbook of electroencephalography and clinical neurophysiology* (rev. ed., vol. 3, pp. 427–534). Amsterdam: Elsevier.
- McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology, 12*, 203–208.
- Mohr, B., Pulvermüller, F., & Zaidel, E. (1994). Lexical decision after left, right, and bilateral presentation of content words, function words and non-words: Evidence for inter-hemispheric interaction. *Neuropsychologia, 32*, 105–124.
- Neville, H. J., Coffey, S. A., Holcomb, P. J., & Tallal, P. (1983). The neurobiology of sensory and language processing in language impaired children. *Journal of Cognitive Neuroscience, 5*, 235–253.
- Neville, H. J., Mills, D. L., & Lawson, (1992). Fractionating language: Different neural subsystems with different sensitive periods. *Cerebral Cortex, 2*, 244–258.
- Nobre, A. C., & McCarthy, G. (1994). Language-related ERPs: Scalp distributions and modulations by word type and semantic priming. *Journal of Cognitive Neuroscience, 6*, 233–255.
- Oldfield, R. C. (1971). The assessment of the analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*, 97–113.
- Osterhout, L., Bersick, M., & McKinnon, R. (1997). Brain potentials elicited by words: Word length and frequency predict the latency of an early negativity. *Biological Psychology, 46*, 143–168.
- Osterhout, L., & Holcomb, P. J. (1995). Event related potentials and language comprehension. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiological studies of human cognitive function* (pp. 171–215). Oxford: Oxford University Press.
- Pulvermüller, F., Lutzenberger, W., & Birbaumer, N. (1995). Electro cortical distinction of vocabulary types. *Electroencephalography and Clinical Neurophysiology, 94*, 357–370.
- Rösler, F., Heil, M., & Glowalla, U. (1993). Monitoring retrieval from long-term memory by slow event-related brain potentials. *Psychophysiology, 30*, 170–182.
- Rugg, M. D. (1995). ERP studies of memory. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind* (pp. 132–170). Oxford: Oxford University Press.
- Rugg, M. D., & Doyle, M. C. (1992). Event-related potentials and recognition memory for low- and high-frequency words. *Journal of Cognitive Neuroscience, 4*, 69–79.
- Segui, J., Mehler, J., Frauenfelder, U., & Morton, J. (1982). The word frequency effect and lexical access. *Neuropsychologia, 20*, 615–627.
- Swaab, T. S., Brown, C. M., & Hagoort, P. (1997). Spoken sentence comprehension in aphasia: Event-related potential evidence for a lexical integration deficit. *Journal of Cognitive Neuroscience, 9*, 39–66.
- Swinney, D. A., Zurif, E. B., & Cutler, A. (1980). Effects of sentential stress and word class upon comprehension in Broca's aphasics. *Brain and Language, 10*, 132–144.
- Ter Keurs, M., Brown, C. M., Hagoort, P., & Stegeman, D. F. (1999). Electrophysiological manifestations of open- and closed-class words in patients with Broca's aphasia with agrammatic comprehension: An event-related brain potential study. *Brain*.
- Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory and Cognition, 18*, 380–393.
- Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open- and closed-class words. *Memory and Cognition, 19*, 95–112.
- Winer, B. J. (1971). *Statistical principles in experimental design*. New York: McGraw-Hill.