Clinical Neurophysiology 115 (2004) 1090–1103

www.elsevier.com/locate/clinph

Effects of word length and frequency on the human event-related potential

O. Hauk*, F. Pulvermüller

Medical Research Council, Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 2EF, UK

Accepted 4 December 2003

Abstract

Objective: We investigated the influence of the length and frequency of printed words on the amplitude and peak latencies of event-related potentials (ERPs). This served two goals, namely (I) to clarify their possible effects as confounds in ERP experiments employing word-stimuli, and (II) to determine the point in time of lexical access in visual word recognition.

Methods: EEG was recorded from 64 scalp sites while subjects ($n = 12$) performed a lexical decision task. Word length and frequency were orthogonally varied between stimulus groups, whereas variables including regularity of spelling and orthographic tri-gram frequency were kept constant.

Results: Long words produced the strongest brain response early on (~ 100 ms after stimulus onset), whereas those to short words became strongest later (150–360 ms). Lower ERP amplitudes were elicited by words with high frequency compared with low frequency words in the latency ranges 150–190 ms and 320–360 ms. However, we did not find evidence for a robust alteration of peak latencies with word frequency.

Conclusions: Length and frequency of word stimuli have independent and additive effects on the amplitude of the ERP. Studies on the precise time course of cognitive processes should consider their potentially confounding character. Our data support the view that lexical access takes place as early as 150 ms after onset of written word stimuli.

Keywords: Visual word recognition; Event related potential; Lexical decision; Lexical access; Word length; Word frequency

1. Introduction

The present experiment attempts to define the influence of two crucial parameters, length and normalized lexical frequency of printed words, on the event-related potential (ERP) response while other important variables are kept constant. An important reason for investigating the word frequency effect on ERPs, and to separate it from the length effect, is to determine the point in time of *lexical access*. Word frequencies can have an effect on brain physiology only after word representations have been activated. Therefore, the earliest word frequency effect on ERPs provides an upper limit for the latency of lexical access (Sereno et al., 1998). Furthermore, detailed knowledge about the neurophysiological effect of word frequency and length is necessary because these variables can act as confounds in studies of language processing in the brain (Pulvermüller, 1999).

The modulation of reaction times in a lexical decision task by word frequency or lexical familiarity is a well-established finding (Rubinstein et al., 1970; Scarborough et al., 1977; Whaley, 1978; Gernsbacher, 1984). Response latencies were generally found to be shorter for words with higher frequency. There is evidence that these latencies depend approximately linearly on the logarithm to base 10 of the word frequency (Whaley, 1978; Just and Carpenter, 1980). Investigations on the effect of word frequency are complicated by the fact that this parameter is negatively correlated with word length (Zipf, 1935; Whaley, 1978). To disentangle the effects of word length and frequency, these parameters need to be orthogonally varied. We report here on the first study in which length and frequency of words were varied independently and the effect of these variables on the ERP was analyzed.

Effects of stimulus or task parameters on ERP components can basically be classified into 3 categories. (1) Overall amplitude differences: The topography of

---

* Corresponding author. Tel.: +44-1223-355294x741; fax: +44-1223-359062.

E-mail address: olaf.hauk@mrc-cbu.cam.ac.uk (O. Hauk).

1388-2457/$30.00 © 2004 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.clinph.2003.12.020

the scalp distribution remains unaffected, but the signal strengths at all electrodes are modulated by a common factor. Effects of this sort can be detected by root-mean-square (RMS) or global-field-power (GFP) analysis. (2) Peak latency shifts: The maximum or minimum amplitude within specified time windows occurs at different latencies for different conditions. These can be computed for individual electrodes, mean values over electrode groups, or for RMS or GFP curves. (3) Topographic effects: The topography of the scalp distribution changes, i.e. signal strengths at different electrodes vary with different factors. Effects of this sort can be detected by analysis of variance (ANOVA) with appropriate normalization procedures (McCarthy and Wood, 1985).

Several previous studies reported findings on the variation of these parameters with word frequency and length. These studies and their main features concerning task, stimuli and methodology are summarized in Table 1. They differ significantly with respect to several aspects, for example the stimulus words' grammatical categories (nouns, verbs, adjectives, etc.), lexical frequency and length, the task employed (lexical decision, silent reading, memory task), measurement technique (MEG, EEG, for the latter which recording sites and reference electrodes were used), and number of recording channels. The previous results about effects of word length and frequency, as reported in the studies of Table 1, are summarized in Fig. 1. The summary is structured according to the latency ranges in which effects occurred, whether effects were found depending on word length or frequency, and whether overall amplitude effects or peak latency shifts were found.

Results on the influence of word length are sparse and inconsistent. Van Petten and Kutas (1990) found short words to produce more activity in the latency range 150–225 ms, and the opposite pattern between 250 and 600 ms. Brown et al. (1999) did not find any effects of word length on ERP amplitudes from 230 to 340 ms and from 350 to 550 ms. The study of Assadollahi and Pulvermüller (2001a) found the earliest effects of word length around 100 ms after word onset. Long words produced higher amplitudes than short words for several latency ranges between 60 and 220 ms, but larger responses to short words were seen between 370 and 800 ms. Osterhout and Bersick (1997) are the only authors who reported peak latency shifts in dependence of word length. They found a component between 250 and 450 ms that peaked later for long words.

The results concerning amplitude effects of word frequency are comparatively consistent. Generally, lower amplitudes for words with higher frequencies are reported (Rugg, 1990; Van Petten and Kutas, 1990; Brown et al., 1999; Assadollahi and Pulvermüller, 2001a). In contrast, Polich and Donchin (1988) found larger P3 amplitudes for rare than for common words. Van Petten and Kutas (1990) detected amplitude effects with word frequency in the latency range 350–500 ms, which corresponds to the 300–500 ms interval in which Rugg (1990) found frequency effects. Brown et al. (1999) reported frequency effects slightly earlier between 230 and 340 ms, but for short words only. The earliest effects were found in the MEG study by Assadollahi and Pulvermüller (2001a) between 120 and 170 ms, also only for short words. Both short and long words exhibited frequency effects around 200 ms.

Several studies focused on the analysis of peak latency shifts of evoked brain activity depending on word frequency (Polich and Donchin, 1988; Osterhout et al., 1997; King and Kutas, 1998; Embick et al., 2001). Without exception, they reported longer latencies for words with low frequency. The effects were mainly found for peaks occurring between 250 and 450 ms, except for Polich and Donchin (1988) who found them around 560 ms. The studies of Osterhout and Bersick (1997) and King and Kutas (1998) suffered from one major shortcoming: Word length, word class (content/function words), repetition rate and word frequency were correlated in their stimulus set. They did not vary these parameters orthogonally, but applied a regression analysis instead. This can reveal which parameter explains most of the variance in the data. However, if the parameters are highly correlated with each other, this approach makes it difficult to separate the effects of one parameter (e.g. word frequency) from those of the another (e.g. length). Little attention has been paid to topographical effects of word length and frequency. Only Rugg (1990) reported such effects for the latency ranges 300–500 ms and 500–800 ms. The main focus of our study therefore lies on amplitude and latency modulation of the ERP, though topographical factors are investigated where it seemed appropriate.

The excellent time resolution of ERP methodology furthermore enables us to detail the time course of visual word recognition. The most widely investigated ERP component related to semantic processing has been the N400, which exhibits maximal amplitude at around 400 ms. It is often interpreted as reflecting the integration of semantic information in a developing context (Kutas and Hillyard, 1980; Halgren et al., 2002). Behavioural as well as electrophysiological data, however, support the view that lexical access takes place much earlier, namely between 100 and 200 ms (Pulvermüller et al., 1995; Sereno et al., 1998). If words are presented together with orthographically and phonologically legal pseudowords, i.e. if the lexical decision cannot be based on the ground of pre-lexical features of the word form (e.g. illegal letter constellations), effects of word frequency can be expected to arise at the stage of lexical access or later (Gernsbacher, 1984; Balota and Chumbley, 1984). Our study, employing well-controlled stimuli, is therefore suited to provide further evidence to determine the time range in which lexical access occurs.

Summarizing, our study intends to clarify which ERP effects in visual word recognition for content words can be attributed to word length and frequency, and to determine the time range of lexical access. To accomplish our goal, we carefully considered the following aspects of our design: We employed a lexical decision task as one of the simplest and

Table 1
Summary of previous results on effects of word length and frequency on the human ERP

Author, year	Stimuli	Parameter range and database	Task	Method	No. of channels
Polich and Donchin, 1988	Mono-syllabic, Nouns/Verbs, H/L	F: < 0– > 1.5 (Kuc-Fr) L: mean 4.5 letters	Lexical decision	EEG (LM)	3
Rugg, 1990	OC, H/L, matched for length	F: 0– > 2 (Kuc-Fr) L: not reported	Lexical decision	EEG (LM)	5
Van Petten and Kutas, 1990 (Expt. 1)	OC, 6 frequency categories, partially matched for length	F: ~0.54–2.9 (Fr-Kuc) L: ~5.4 (Fr-Kuc)	Silent reading of sentences	EEG (LfM)	13
Van Petten and Kutas, 1990 (Expt. 2)	OC, 2 frequency categories, matched for length	F: < 1.5– > 1.5 L: 3– > 8 letters	Silent reading of sentences	EEG (LM)	10
Osterhout et al., 1997	OC and CC, variable in length and frequency	F: ~2.25–3.8 (Fr-Kuc) L: ~2–6 letters	Silent reading of scrambled prose (Expt. 2)	EEG (LfM)	13
King and Kutas, 1998	OC and CC, variable in length and frequency	F: 0–5 (Fr-Kuc) L: not reported	Silent reading of unrelated sentences	EEG (LM)	26
Sereno et al., 1998	H/L regular and exception words, matched for length	F: ~0.9–2.4 (Fr-Kuc) L: 4/5/6 letters	Lexical decision	EEG (AR)	64
Brown et al., 1999	OC and CC, variable in length and frequency	F: ~1.3–3.3 (CELEX) L: 3–8 letters	Silent reading of a story	EEG (LfM)	23 (out of 29)
Assadollahi and Pulvermüller, 2001a	OC and CC, orthogonal design [St, Lg]*[H, L], 4 stimuli per category	F: 1.1–2.3 (CELEX) L: ~4.5–6.5 letters	Memory task	MEG	5–30 channels (out of 148)
Embick et al., 2001	OC, 6 frequency categories	F: -0.7–2.8 (CELEX) L: 3–7 letters	Lexical decision, words repeated 3 times	MEG	17 (out of 64) in left hemisphere
Present study	OC, orthogonal design [St, Lg]*[L, M, H]	F: 0.63–2.14 (CELEX) L: 4.1–6.3	Lexical decision	EEG (AR)	64

If a study reported results for words within context and out of context in separate experiments, only the results for the latter were included. OC/CC, open class/closed class; L/M/H, low/medium/high frequency; St/Lg, Short, Long; LM, linked mastoids reference; LfM, left mastoid reference; AR, average reference. F, frequency; L, length. Frequencies are given as logarithm to base 10. Kuc-Fr, Fr-Kuc and CELEX refer to lexical databases of Kucera and Francis (1967), Francis and Kucera (1982) and Baayen et al. (1993), respectively.

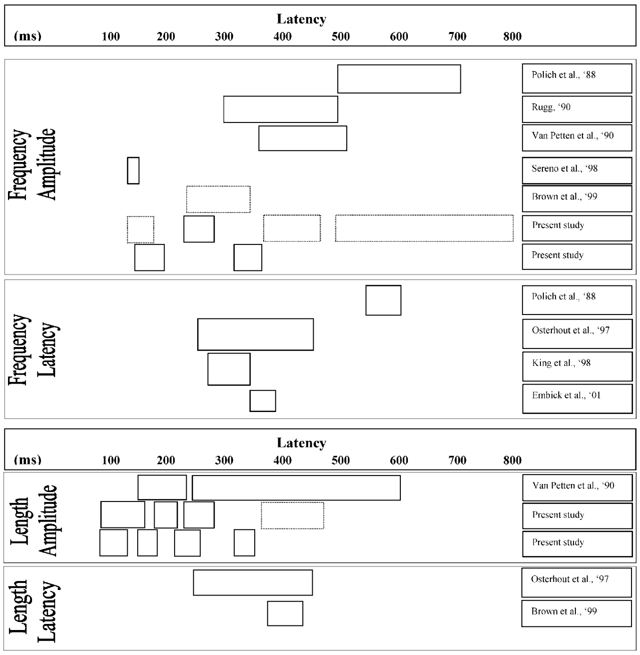

Fig. 1. Summary of effects of word length and frequency on ERP data as revealed by the studies listed in Table 1. The references to the studies are given at the right side of the figure. The squares indicate a significant effect reported for the corresponding latency range. Dotted squares indicate that effects were obtained only for a subset of stimuli, such as word frequency effects that were present for short words only.

most widely used tasks to tap into the processing stage of lexical access. (1) We orthogonally varied word frequency and word length to separate the effects of both parameters. In addition, this allowed us to test for possible interactions of these parameters. (2) Three-word frequency groups were used (high/medium/low frequency) to check whether effects are due to extreme cases or can rather be considered as 'continuous' over the whole frequency range. (3) Our stimuli were controlled for the dimensions word length, word form frequency, lemma frequency, regularity of spelling and orthographic tri-gram frequency. Since behavioural data were shown to correlate with the logarithm to base 10 of the word frequencies, the logarithms were also matched. (4) We analysed our ERP data with respect to amplitude, peak latency shifts, and topography where appropriate.

We will use our results to draw careful conclusions on the interpretation of neuroimaging studies of language processing in the brain.

2. Methods

2.1. Subjects

Twelve right-handed monolingual native speakers of English entered the analysis (10 female, 2 male). Their mean age was 21.4 years (SD 3.6). All had normal or corrected-to-normal vision and reported no history of neurological illness or drug abuse. Handedness was determined according to Oldfield (1971), revealing a mean laterality quotient of 89 (SD 14). Informed consent was obtained from all subjects and they were paid for their participation. This study was approved by the Cambridge Psychology Research Ethics Committee.

2.2. Stimuli

The stimulus words were subdivided into 6 categories ([short, long] × [low, medium, high frequency]). Short words were mono- and long words bi-syllabic. All words could be used as nouns, though not exclusively, which is difficult to accomplish in English. Each category contained 69 items, resulting in 414 stimulus words altogether. An equal number of pseudowords was included. Pseudowords were generated for each word category by exchanging letters within words or between words within each category. Therefore, pseudowords were as similar to the real words as possible, and were matched for mean length and letter frequency as well. Word categories were matched according to mean word length, mean word form and lemma frequencies taken from the CELEX database (Baayen et al., 1993).

Word frequency is measured as the number of occurrences of a word within a given corpus of written and/or spoken items (Kucera and Francis, 1967; Baayen et al., 1993). The most commonly used frequency measure is the word form frequency, which distinguishes between inflectional variants such as 'car' and 'cars.' In this example, however, both items share their word stem and root and thus much of their semantic contents, and might have a common lexical entry or 'lemma.' If it is lexical entry retrieval that is reflected in the behavioural and electrophysiological data, the lemma frequency would be the more adequate parameter to match for. To avoid any problem of this kind, we matched both word form and lemma frequency in our study.

Care was taken to keep also the standard deviations of these parameters similar between categories, since different variabilities between categories may cause differences in peak width, latency and amplitude in the ERP (Pulvermüller, 1999). In addition, logarithms of word frequencies were matched because they were shown to relate linearly to behavioural responses (Whaley, 1978; Just and Carpenter, 1980). The main properties of our stimuli are listed in Table 2.

We found only 2–4 items per category to be spelled irregularly, which made a more detailed analysis of regularity of spelling unnecessary. We further quantified the matching of our stimulus material performing an ANOVA analysis with the factors LENGTH and FREQUENCY on all relevant parameters. We did not find any unintended significant differences with respect to the parameters number of letters, word form frequency, logarithm to base 10 of word form frequency, lemma frequency, logarithm to base 10 of lemma frequency and

Table 2
Listing of stimulus parameters taken into consideration in the selection of words

	Word length	Low frequency	Medium frequency	High frequency
NL	Short	4.1 (0.84)	4.1 (0.87)	4.1 (0.81)
	Long	6.3 (0.80)	6.2 (0.91)	6.2 (0.83)
WF	Short	5 (2)	30 (15)	159 (104)
	Long	5 (2)	31 (19)	180 (203)
Log WF	Short	0.67 (0.23)	1.42 (0.23)	2.13 (0.25)
	Long	0.60 (0.24)	1.40 (0.27)	2.14 (0.27)
LF	Short	8 (4)	45 (22)	207 (133)
	Long	7 (5)	42 (27)	222 (223)
Log LF	Short	0.84 (0.27)	1.59 (0.24)	2.24 (0.25)
	Long	0.77 (0.23)	1.54 (0.28)	2.25 (0.26)
OTF	Short	9.22 (7.12)	10.4 (8.32)	10.99 (7.65)
	Long	8.6 (5.84)	8.79 (6.57)	9.55 (7.46)
ONS	Short	6.04 (4.88)	8.78 (6.39)	8.87 (5.68)
	Long	1.09 (1.62)	1.48 (1.71)	1.58 (2.32)

Stimuli were grouped into long and short words, and words with low, medium, and high word frequency (number of word occurrences per million in database). Average values for stimulus groups ($n = 69$ per group) are given along with their standard deviations in brackets. Parameters are: Number of letters (NL), Word form frequency (WF), Logarithm to base 10 of WF (Log WF), Lemma frequency (LF), Logarithm to base 10 of LF (Log LF), Orthographic tri-gram frequency (OTF), Orthographic neighborhood size (ONS).

orthographic tri-gram frequency (OTF). The ANOVA on orthographic neighbourhood size (ONS), however, revealed main effects LENGTH ($F(1,68) = 192.16$, $P < 0.001$) and FREQUENCY ($F(2,136) = 7.93$, $P < 0.001$) as well as an interaction LENGTH × FREQUENCY ($F(2,136) = 14.24$, $P < 0.05$). The main effect LENGTH was due to the smaller ONS for long words. Planned comparison tests revealed that the effects involving word frequency were mainly due to significantly lower ONS values for short low-frequency words compared to those with short high and medium frequency. We will discuss at a later stage whether this is a plausible confound of any effects found in our data.

2.3. Procedure

A lexical decision task was applied. White letter strings were presented on a grey background on a computer screen. Each stimulus was presented tachistoscopically for 100 ms. The stimulus onset asynchrony (SOA) varied between 2 and 3 s. A fixation cross was shown in the centre of the screen during the whole experiment.

Subjects were instructed to press one button of a computer keyboard with the index finger of one hand in response to a real word, and another button with the middle finger of the same hand in response to a pseudoword. The response hand was alternated among subjects. The stimulus delivery and response collection were controlled by the Experimental Run Time System software (ERTS, BeriSoft, Germany).

Four sequences of stimuli were created that were randomly assigned to different subjects. These sequences were further divided into 4 blocks each. Each block lasted about 9 min and contained two breaks of 10 s duration. The sequence of blocks was changed among subjects. Subjects were instructed to minimize eye and body movements throughout the experiment, and to restrict them to the break periods.

2.4. Data recording

The electroencephalogram (EEG) was measured in an electrically and acoustically shielded EEG chamber at the MRC Cognition and Brain Sciences Unit in Cambridge, UK. Data were recorded from 64 Ag/AgCl electrodes mounted on an electrode cap (QuickCap, Neuromedical Supplies, Sterling, USA) using SynAmps amplifiers (NeuroScan Labs, Sterling, USA), arranged according to the extended 10–20 system. Sampling rate was 500 Hz, and a 0.1–30 Hz bandpass filter was applied. AFz was used as recording reference for the EEG channels, and data were converted off-line to average reference. The EOG was recorded bipolarly through electrodes placed above and below the left eye (vertical) and at the outer canthi (horizontal). After the experimental session, eye movements were recorded for eye artefact correction according to Berg and Scherg (1994). Subjects were instructed to blink and to move their eyes to the left, right or up and down, as indicated by signs appearing on the computer screen.

2.5. Data analysis

2.5.1. Pre-processing of ERP data

The continuously recorded data were divided into epochs of 1 s length, starting 200 ms before stimulus onset. Trials with peak-to-peak potential differences larger than 100 μV in at least one EEG channel or 120 μV in at least one EOG channel were rejected, and an eye artefact correction algorithm was applied (Berg and Scherg, 1994). In the averaged data, for each channel the mean amplitude of a 100 ms pre-stimulus interval was subtracted at all time points.

2.5.2. Amplitude analysis

Most of the studies listed in Table 1 reported mere amplitude effects of word length and frequency. We therefore investigated overall amplitude effects by an ANOVA on root mean square (RMS) values computed over all EEG channels for each condition. This method attenuates effects that are only present at few electrode sites. For this reason, we ran additional analyses on selected electrode clusters, which also allowed us to introduce a topographical variable where appropriate. The most robust statistical results can be expected if the analysis is restricted to recording sites with high signal-to-noise ratio, and if the number of variables is kept low. Thus we selected electrode clusters around prominent peaks in the ERP distribution

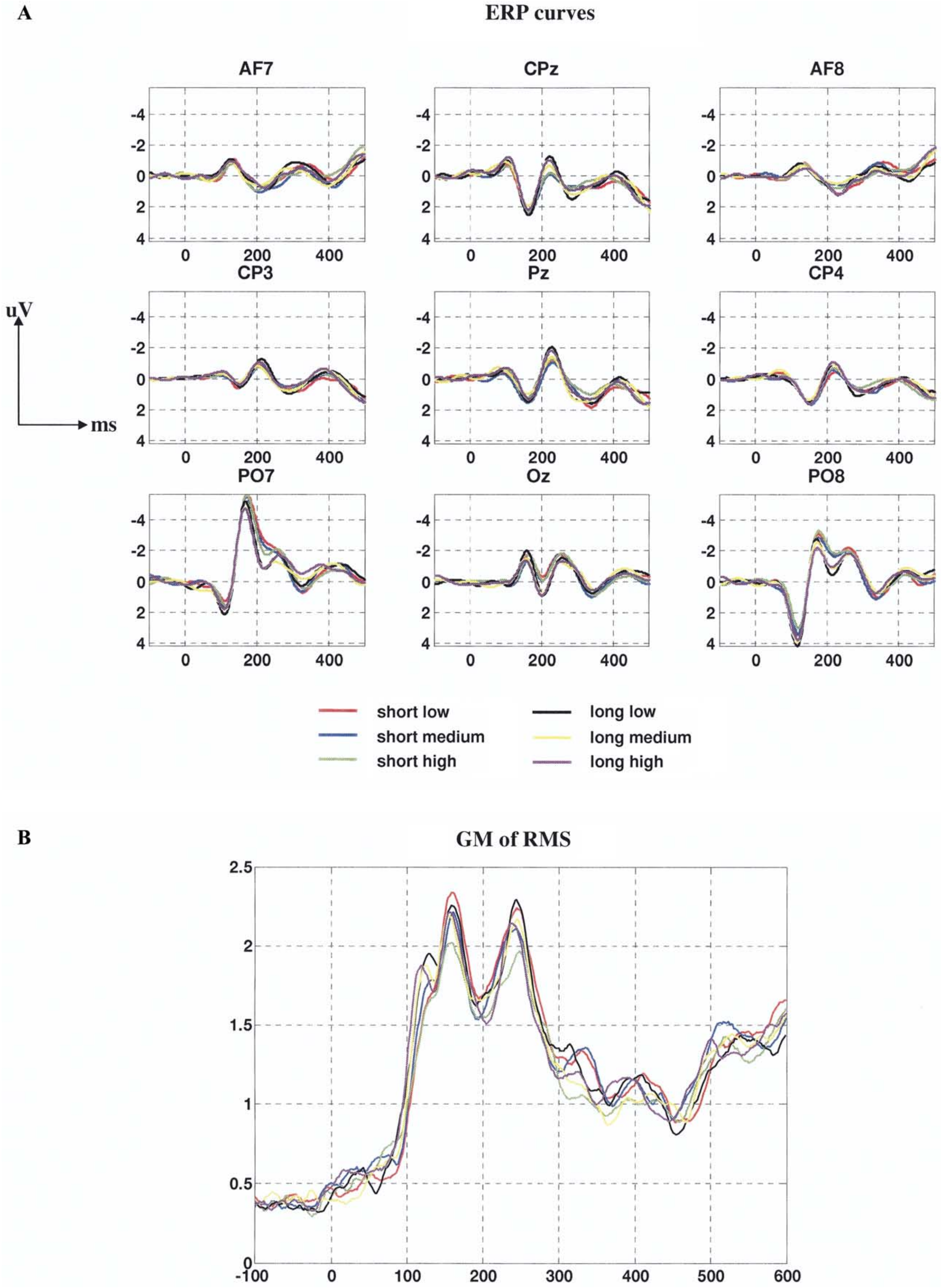

Fig. 2. Time course of electrophysiological data. (A) ERP curves (average reference) are shown for prominent electrode locations. (B) The grand mean of RMS values across all electrodes is plotted over time.

(see Fig. 5). For earlier time windows (80–260 ms), we computed mean voltages around electrodes PO7 (PO7, O1, P7 and PO5) in the left and PO8 (PO8, O2, P8 and PO6) in the right hemisphere, which entered the analysis as a factor LATERALITY. For a later time window (320–360 ms), 9 electrodes around Pz (Pz, P1, P2, Cz, C1, C2, CPz, CP1 and CP2) were chosen.

2.5.3. Peak latency analysis

Peak latencies were analysed on the basis of RMS values over all channels and on ERPs at single electrodes. For latency ranges in which prominent RMS peaks could be found in individual data sets, peak latencies were determined for each condition in each subject. The peak latency was defined as the time point at which the RMS reached its maximum within the given time range. The corresponding values entered an ANOVA on the factors LENGTH and FREQUENCY. For some latency ranges, one or two subjects did not show detectable RMS peaks and were therefore left out of the analysis. Where this is the case, it will be reported in Section 3.

Peak latencies for ERPs at single electrodes were determined for grand mean data only. Only electrodes that exhibited local extrema within the given time ranges were considered, i.e. curves with ‘peaks’ only at the borders of the time range were neglected. Peak latencies were then defined as the time points at which the corresponding ERP curve reached its extreme value within the corresponding time range.

3. Results

3.1. Behavioural data

Reaction times distinguished between word frequencies, but not between word lengths. The only significant result was a main effect for FREQUENCY (low: 697 ms, medium: 668 ms, high: 650 ms; $F(2,22) = 27.60$, $P < 0.001$). Planned comparison tests revealed that all differences between frequency ranges were significant (low/medium: $F(1,11) = 30.86$, $P < 0.001$; medium/high: $F(1,11) = 6.36$, $P < 0.05$; low/high: $F(1,11) = 53.07$, $P < 0.001$).

The same pattern of results was found for error rates. Only the main effect FREQUENCY was significant (low: 11.3%, medium: 6.2%, high: 3.3%; $F(2,22) = 25.39$, $P < 0.001$). Planned comparison tests revealed significant differences between all frequency groups: low/medium, $F(1,11) = 17.18$, $P < 0.01$; medium/high, $F(1,11) = 18.64$, $P < 0.01$; low/high, $F(1,11) = 33.48$, $P < 0.001$.

3.2. ERP data

ERP curves for selected electrodes and RMS curves over all channels are presented in Fig. 2. The first deflection in the RMS curve appears shortly after 100 ms (Fig. 2B). The ERPs show maximal values with positive polarity at posterior electrodes PO7 and PO8 (Fig. 2A). The spline map in Fig. 5A shows two distinct peaks around these sites. The next deflection in the RMS curve is present shortly after 150 ms, with negative polarity at left and right posterior electrode sites (Figs. 2B and 5B). It is followed by a peak around 250 ms with a broader voltage distribution (Fig. 5C), that has negative polarity at posterior electrode sites, and a positive peak at a fronto-central location. A smaller RMS peak occurs after 300 ms. It is characterized by a positive peak at centro-parietal electrode sites, and a weak negative voltage distribution in frontal areas (Fig. 5D). A peak with even smaller amplitude is visible at around 400 ms, followed by a larger deflection around 500 ms. Because no significant effects were found for the two latter time windows (see below), we are not presenting the corresponding topographies for these latency ranges.

Time ranges around the peaks described above were chosen such that topographies did not substantially vary within these time windows. Later latencies were not considered either because of the low amplitudes of the signals and the corresponding low signal-to-noise ratios, or because they coincided or succeeded the button press response (average response time ~670 ms). The corresponding time windows were 80–125, 150–190, 210–260, 320–360, 380–440 and 450–550 ms.

3.3. Results of statistical ERP analysis

The results of the ANOVA analyses for several latency ranges are reported in Table 3. RMS values and mean voltages for electrode clusters, as described above, corresponding to significant ANOVA effects are presented in Figs. 3 and 4, respectively. Spline interpolated grand-mean topographies are shown in Fig. 5.

Statistically significant effects of word length occurred already in the latency range 80–125 ms, longer items showing higher amplitude than shorter ones (Figs. 3A and 4A). Further length effects were found between 150 and 190 ms (Fig. 4B), 210–260 ms (Fig. 4D) and 320–360 ms (Fig. 3C), where short words showed stronger signals than long ones. Frequency effects occurred in the latency ranges 150–190 ms (Figs. 3B and 4C) and 320–360 ms (Figs. 3D and 4E). In both of these latency ranges, overall signal strength decreased with increasing word frequency. No significant interactions between word length and frequency were found.

Below, additional results of post hoc tests or marginally significant results that are not reported in Table 3 will be described in more detail:

In the earliest time window, we found a length effect for electrode clusters around PO7 and PO8. The main effect LATERALITY only approached significance

Table 3
Summary of ANOVA results documenting neurophysiological differences as a function of word length and frequency

	80–125 ms	150–190 ms	210–260 ms	320–360 ms
Frequency amplitude	–	RMS: L > M > H $F(2,22) = 3.89$, $P < 0.05$	–	RMS: L > M > H $F(2,22) = 3.88$, $P < 0.05$ Elec: L > M > H $F(2,22) = 3.87$, $P < 0.05$
Length amplitude	RMS: Lg > St $F(1,11) = 18.42$, $P < 0.01$ Elec: Lg > St $F(1,11) = 5.72$, $P < 0.05$	Cluster: St > Lg $F(1,11) = 10.22$, $P < 0.01$	Cluster: St > Lg $F(1,11) = 8.94$, $P < 0.05$	RMS: St > Lg $F(1,11) = 9.96$, $P < 0.01$

RMS, Effect found in root-mean-square data; Cluster, Effect found in electrode cluster data; L/M/H, low/medium/high frequency words; St/Lg, short/long words. Hyphens indicate the absence of significant effects. The latency ranges 380–440 ms and 450–550 ms did not disclose any significant effects.

($F(1,11) = 4.08$, $P = 0.07$), the right cluster being more positive than the left (Fig. 4A).

Between 150 and 190 ms, there was a robust frequency effect on RMS values. RMS values decreased with increasing word frequency (Fig. 3B). ERPs to high and low frequency words were significantly distinct ($F(1,11) = 8.94$, $P < 0.05$), but only a marginally significant effect emerged for the low/medium comparison $F(1,11) = 3.68$, $P = 0.08$), and no significant effect at all was present for medium/high frequency words. The word frequency effect on electrode clusters only approached significance ($F(2,22) = 2.71$, $P = 0.09$). There was a main effect LATERALITY, the left electrode clusters showing larger amplitude than the right one ($F(1,11) = 14.52$, $P < 0.01$; Fig. 4C).

In the time window 210–260 ms, planned comparison tests explained the interaction LENGTH × LATERALITY (see Table 3) as a larger dissociation of short and long words in the left hemisphere (Fig. 4D): for the left electrode cluster, short words produced significantly more negativity than long ones ($F(1,11) = 14.10$, $P < 0.01$), but this contrast only approached significance in the right hemisphere ($F(1,11) = 4.10$, $P = 0.07$). The tests on LATERALITY for short and long words separately did not reach significance.

Between 320 and 360 ms, overall activity was lower for long words than for short ones, and it decreased with increasing word frequency (Fig. 3D). Planned comparison tests revealed a significant difference between high and low frequency words ($F(1,11) = 4.94$, $P < 0.05$), and

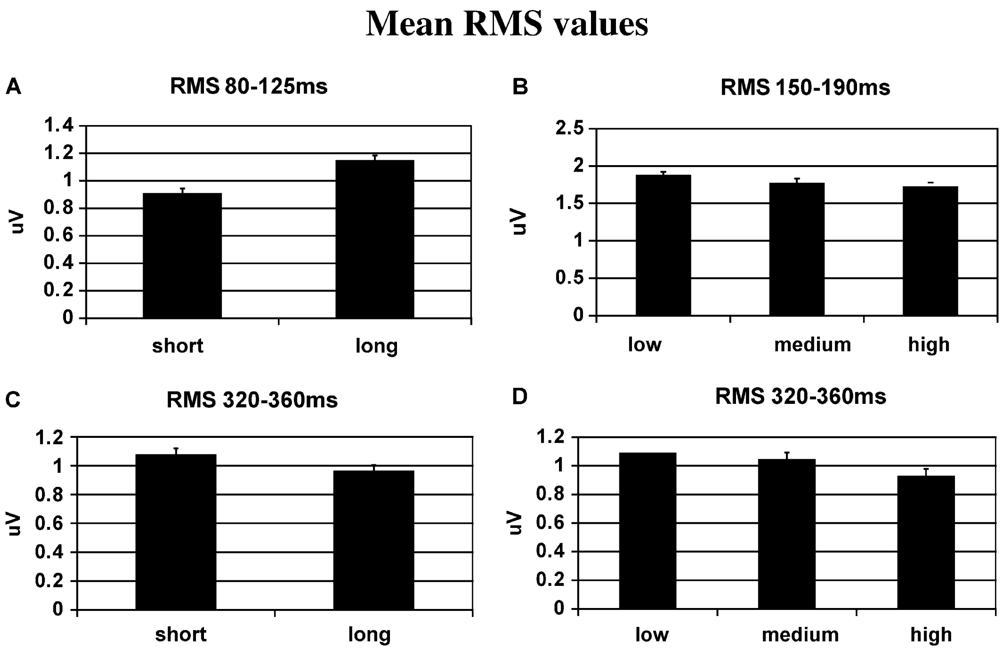

Fig. 3. Mean RMS values across subjects corresponding to significant main effects obtained in the ANOVA analysis. Error bars show the standard error of the mean after between subject variability has been removed, appropriate for repeated-measures comparisons (Loftus and Masson, 1994).

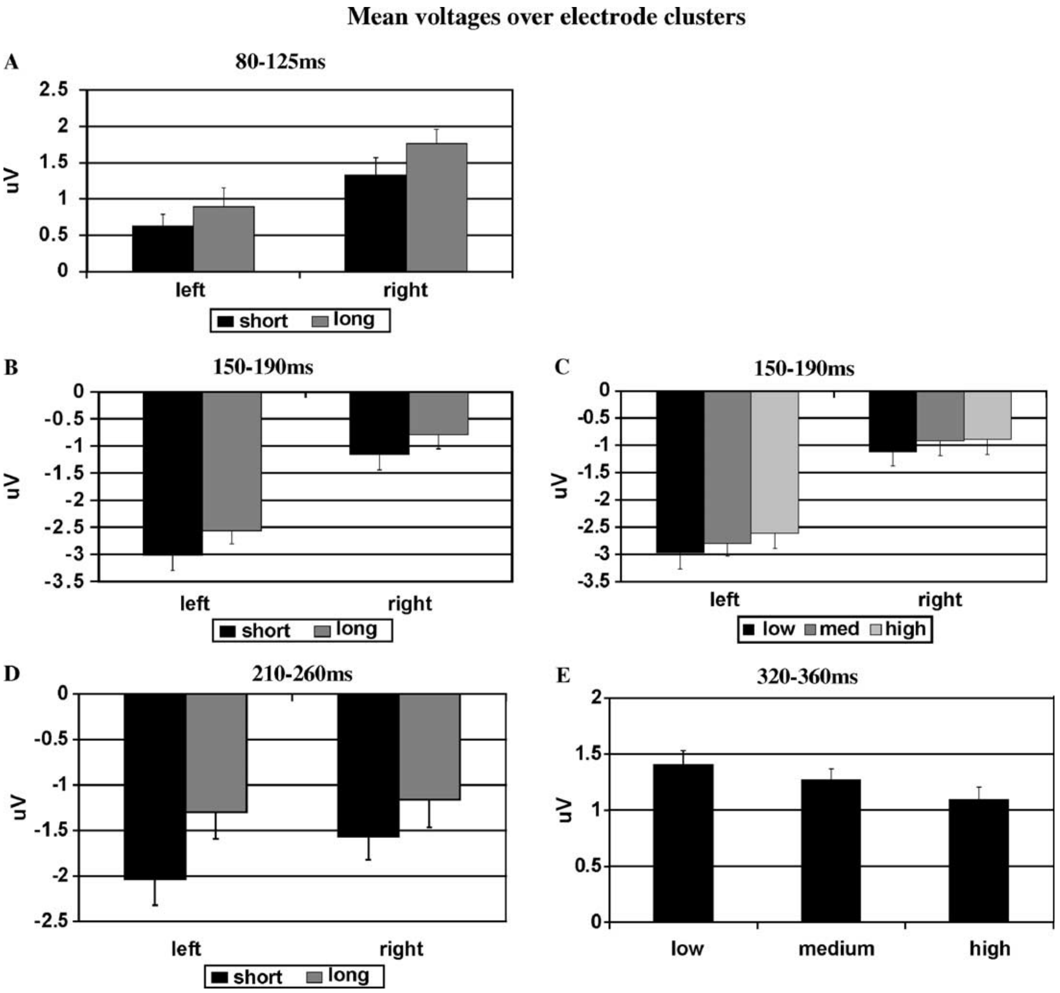

Fig. 4. Mean voltage values over electrode clusters, corresponding to significant main effects or trends of an ANOVA analysis. For the time ranges within 80–260 ms, clusters of 4 electrodes each were selected around electrodes PO7 (left hemisphere) and PO8 (right hemisphere). In the latency range 320–360 ms, a cluster of 9 electrodes around electrode Pz was chosen. Effects of word length on the ERP are displayed in diagrams on the left side of the figure, those of word frequency on the right. Error bars show the standard error of the mean after between-subject variability has been removed, appropriate for repeated-measures comparisons (Loftus and Masson, 1994).

a marginally significant difference between medium and high frequency words ($F(1,11) = 3.76$, $P = 0.08$). The mean amplitude of the Pz-cluster decreased with increasing word frequency (Fig. 4E). Planned comparison tests found a significant difference between high and low frequency words ($F(1,11) = 6.18$, $P < 0.05$), and between medium and high frequency words ($F(1,11) = 5.03$, $P < 0.05$).

No effects of word frequency and length on the amplitude of the ERP was found for the time windows 380–440 ms and 450–550 ms.

3.4. Peak latencies

In Fig. 6 we present peak latencies for RMS curves as well as for individual electrode sites. Only data for time ranges which exhibited significant effects or trends are shown.

RMS peak latencies could be detected for most of our subjects and were subjected to statistical analysis. We will report the number of subjects that entered the corresponding analyses (n) in parentheses.

- 100–150 ms ($n = 10$): The ANOVA on RMS peak latencies in this early time window revealed a main effect of LENGTH ($F(1,9) = 5.24$, $P < 0.05$), which consisted of a small 2.4 ms advantage for long words compared with short words (Fig. 6A).
- 150–190 ms ($n = 10$), 190–240 ms ($n = 11$), 260–360 ms ($n = 11$): The peak latency analysis in these time ranges did not reveal any significant effects.
- 450–600 ms: The main effect of FREQUENCY approached significance in this late time window ($F(2,20) = 3.01$, $P = 0.07$). Peak latency decreased with increasing word frequency (Fig. 6B).

Peak latencies at individual recording sites appeared to be too variable over subjects to perform reliable statistical

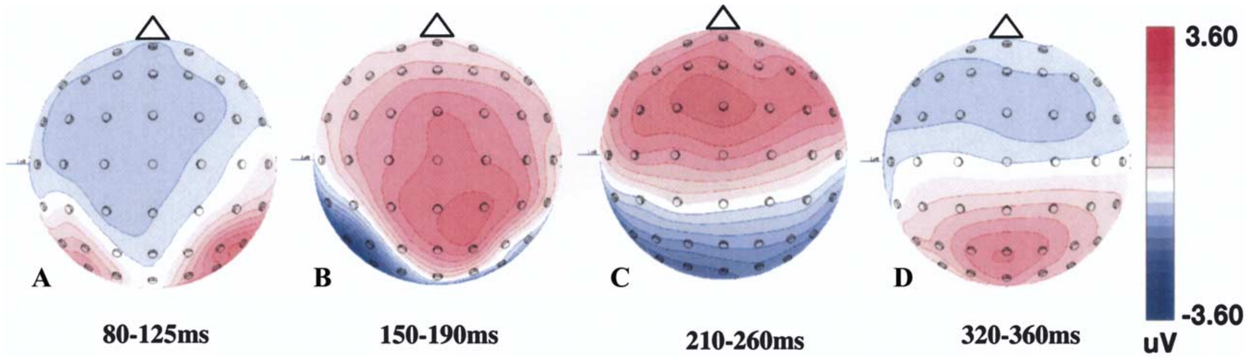

Fig. 5. Spline interpolated grand mean topographies (average reference) for time ranges showing significant effects. The topographies were averaged over all subjects and conditions. Maps are in top view (the nose is up). The colour and contour scaling is the same for all maps.

analysis. Fig. 6C,D show peak latencies for grand mean ERPs at selected electrode positions. The selection of electrode sites and time ranges was guided by our search for a negative correlation between word frequency and peak latencies of early components, as reported by Osterhout and Bersick (1997), King and Kutas (1998) and Embick et al. (2001), and for the P3-component as in Polich and Donchin (1988). For the early time ranges between 100 and 200 ms, peak electrodes PO7 and PO8 are displayed (Fig. 6C). For higher latencies, Pz (Fig. 6C), Cz (Fig. 6D) and Fz (Fig. 6D) were chosen, as well as more lateralized electrodes CP3 and CP4 (Fig. 6C,D) and F3 and F4 (Fig. 6D).

In general, differences between conditions were small and did not follow an easily deducible rule. We illustrate this finding in Fig. 6C (150–350 ms). Most electrodes do not show a monotonous increase or decrease with word frequency at all. Only Pz and CP3 show a slight decrease of peak latencies with increasing word frequency around 300 ms, but only for long words. Electrode CP4 shows larger latency variations than CP3 and Pz, but these do not depend systematically on word frequency, and are partly in

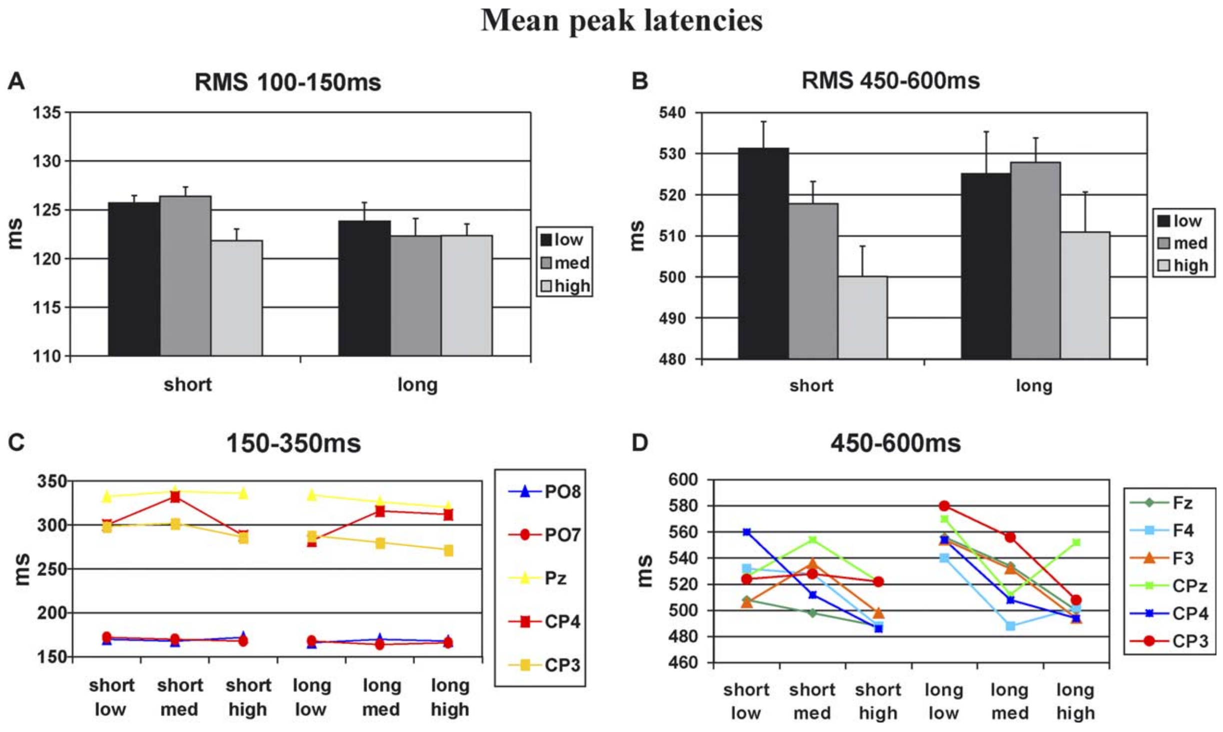

Fig. 6. Mean latencies of prominent peaks. (A,B) Mean peak latencies of the individual RMS curves (with standard errors) for each condition. The only effect that reached significance was a 2.4 ms advantage for long compared with short words at the earliest peak ~120 ms (A). Error bars show the standard error of the mean after between subject variability has been removed, appropriate for repeated-measures comparisons (Loftus and Masson, 1994). (C,D) Peak latencies for grand mean ERP curves at selected electrodes in an early (C) and late (D) time ranges. If electrode labels appear twice within one diagram, they belong to peaks within different time ranges.

the opposite direction of the results reported in the previous studies mentioned above.

In the RMS peak analysis, only the time range after 450 ms showed at least a marginally significant decrease of peak latency with increasing word frequency. This pattern can be observed for several individual electrodes in this latency range, like CP4 and Fz for both short and long words, and CP3 and F3 for long words only (Fig. 6D). However, the electrode where maximum signal was obtained in this interval, CPz, does not show this pattern.

In summary, this data set did not confirm a consistent shift of ERP response latencies with word frequency for components earlier than 500 ms. In most cases, no effect of word frequency was visible, and if so, results from neighbouring electrodes were contradictory. At long latencies above 500 ms, we found a tendency of ERP peak latencies to reflect word frequency similar to Polich and Donchin (1988), but the effect was only marginally significant.

4. Discussion

We investigated the impact of word length and frequency on the human ERP in a visual lexical decision paradigm. Length and frequency were orthogonally varied (long and short words with low, medium, and high word frequency). The word evoked potential changed significantly at different lags:

1. The earliest influence of word frequency on the amplitude of the ERP occurred between 150 and 190 ms. Amplitude decreased with increasing word frequency. A similar pattern was found at 320–360 ms.
2. Word length affected the ERP already in the time range 80–125 ms. Long words produced more activity compared with short ones. In addition, we found slightly shorter RMS latencies for long words in this time window.
3. Additional amplitude effects of word length were detected in the latency ranges 150–190 ms, 210–260 ms, 320–360 ms. In these time intervals, short words produced stronger responses than long ones.

The present data set could not confirm several results reported in earlier studies, as will be discussed in more detail below:

1. There were no interactions between word frequency and word length, neither in the behavioural nor in the ERP data.
2. Reliable peak latency shifts depending on word frequency could not be detected, neither for peaks in the RMS, nor for ERPs at individual electrodes.
3. We did not find reliable amplitude or latency effects later than 360 ms.

Our results differed depending on the analysis procedure employed. For example, some amplitude effects were observed for electrode clusters but not in the RMS analysis (see word length effects 150–190 and 210–260 ms). This is easily explained by the fact that the RMS attenuates effects that are present at only a few electrode sites. However, the opposite pattern was also found: The word frequency effect between 150 and 190 ms and the word length effect between 320 and 360 ms occurred only in the RMS analysis, though the former effect was marginally significant. This indicates that the effect was distributed over larger areas and equally present at numerous electrodes, although possibly too weak to be detected at only a few of them separately. We therefore conclude that the approaches we presented (RMS, electrode clusters, and topographical analysis) provide complementary information. Some differences between findings reported in the studies listed in Tables 1 and 2 might be due to too restricted analysis procedures. In the following, we will discuss our findings in the light of these previous findings.

Which factors might have confounded our results? As mentioned, the parameters word form frequency, lemma frequency, word length, orthographic tri-gram frequency, regularity of spelling, and the lexical category of the stimuli were controlled and can therefore be excluded as possible confounds. The only parameter for which we found significant differences between word categories was orthographic neighbourhood size (ONS). However, the ERP effects did not follow the same pattern as the ONS variable, i.e. we did not find an interaction of word length and frequency A recent study by Holcomb et al. (2002) suggests that ONS influences the amplitude of the N400 after 350 ms, but not the earlier components for which frequency effects were evident in the present data set. We therefore conclude that ONS was not a serious confound in our study.

4.1. Word length effects

Word length was reflected in the ERP from ~100 ms up to ~400 ms. This is consistent with the results of Van Petten and Kutas (1990) and Assadollahi and Pulvermüller (2001a). Whereas long words elicited stronger early neurophysiological responses than short ones, the inverse pattern, stronger responses to short than long words, was seen between 150 and ~400 ms. Furthermore, word length modulated the laterality of the ERP between 210 and 260 ms, the difference between short and long words being larger in the left hemisphere.

These results indicate that the length of words is an important factor determining the amplitude of brain activity elicited by visual word stimuli. The length factor should therefore be controlled in studies of word evoked brain activity. Furthermore, the relatively stronger ERP responses of long words at ~100 ms and short words at ~300 ms imply that the signal-to-noise ratio (SNR) differs as a function of both word length and latency. In order to obtain maximal SNR, it is therefore advisable to restrict the variance of the word length variable in studies of word-evoked neurophysiological activity. Mixing of words with different lengths may blur early neurophysiological effects.

From our data it cannot be decided if the effects attributed to word length reflect differential cortical processing with respect to (a) physical parameters of the stimuli (e.g. luminance), (b) the number of graphemes or phonemes, or (c) the number of syllables. These variables covariated in the present study, as they did in the EEG and MEG studies reviewed in this paper. In the following, we will therefore put the emphasis on word frequency effects. A more fine-grained analysis of the word length effect should be addressed by future research.

4.2. Early vs. late effects of word frequency

We found the earliest effect of word frequency in the N1 latency range between 150 and 190 ms. Similar effects have been reported previously by Sereno et al. (1998) at 132 ms, and by Assadollahi and Pulvermüller (2001a) between 120 and 160 ms. These studies and our own taken together provide converging evidence for the modulation of early electrophysiological brain responses by word frequency. This is evidence that lexical access from written word stimuli is an early process that follows stimulus presentation by less than 200 ms. The later frequency effect at 320–360 ms, as well as late frequency effects in previous studies (Polich and Donchin, 1988; Van Petten and Kutas, 1990; Rugg, 1990; Assadollahi and Pulvermüller, 2001b), are unlikely to reflect lexical access, but may rather indicate re-processing of word-related information or post-lexical processes.

We did not find any effects at later latencies, i.e. after 360 ms, as reported by Rugg (1990), Polich and Donchin (1988), and Van Petten and Kutas (1990). There was a tendency for peak latencies after 400 ms to be shorter for high frequency words. Consistent with our results, significant main effects of word frequency were also absent in the majority of studies on ERP amplitude and latencies occurring more than 400 ms after stimulus onset (Smith and Halgren, 1987; Sereno, 1998; Brown et al., 1999; King and Kutas, 1998; Embick et al., 2001). In their single-case study, Assadollahi and Pulvermüller (2001a) reported a frequency effect for the time window 500–800 ms for short words only, but this finding could not be replicated in their subsequent group study (Assadollahi and Pulvermüller, 2001b). Thus, the question arises why some studies found these very late effects. Although there are, of course, different possible reasons related to differences in stimulus selection, task, data recording and analysis, we would like to point out that two of the studies revealing very late frequency effects included words with very low frequency (equal to or less than 1 per million) into their stimulus set (Polich and Donchin, 1988; Rugg, 1990). Therefore, it is possible that the very late frequency effect depends on the presence of very rare items in the stimulus set.

The effects of word length and frequency were additive in our data, i.e. no interaction between these variables was found. Assadollahi and Pulvermüller (2001a), however, reported such an interaction. Short words exhibited frequency effects about 100 ms earlier than long ones. Neither our behavioural nor our ERP data suggest such an interaction of word length and frequency.

4.3. Relationship between behavioural and electrophysiological results

The pattern of our electrophysiological results on word frequency is in agreement with our behavioural results: The shorter reaction times and lower error rates for words of higher frequency are reflected by smaller ERP amplitudes. The finding that ERP amplitudes increased with decreasing word frequency is consistent with most of the previous studies investigating ERPs (Rugg, 1990; Van Petten and Kutas, 1990; Brown et al., 1999) and also with the psycholinguistic literature (Whaley, 1978; Balota and Chumbley, 1984). The increase of P3 amplitude with increasing word frequency reported by Polich and Donchin (1988) could not be confirmed by the present data set. As pointed out earlier, some of the differences across studies might result from differences in stimulus selection.

There were no effects of word length in the behavioural data. Nevertheless, clear length effects were present in several time slices of our electrophysiological data (80–125 ms, 210–260 ms, 320–360 ms). Unfortunately, none of the ERP studies listed in Table 1 reported behavioural data related to word length. Psycholinguistic studies on lexical decision times reported a significant positive correlation between reaction times and word length, though this effect was small compared to that of word frequency (Whaley, 1978; Balota and Chumbley, 1984). In a recent study by Lavidor and Ellis (2002), no word length effects on lexical decision times were reported when 4–8 letter words were presented centrally, but only when they were presented in the left-visual field. In another experiment, orthogonally varying the factors word length (4–6 letters) and orthographic neighbourhood size (ONS), it was found that word length only affected performance on low ONS words in the left-visual field. Other studies did not report word length effects on lexical decision latencies (Mohr et al., 1994). The size and nature of the effect of word length on lexical decision times can therefore not be considered as well established. Consistent with the psycholinguistic literature, our data suggest that it is small – if present at all – in comparison with the effect of word frequency.

It is possible that the early word length effect in our ERP data is due to differences in physical length, i.e. luminance, between word categories. A similar interpretation was suggested by Assadollahi and Pulvermüller (2001b), who found a similar effect. The difference in number of letters in our study (4.1 vs. 6.2) may not have been large enough to significantly affect processing speed. However, ERP measurements were sufficiently sensitive to the corresponding difference between stimulus groups. This indicates that neuroimaging experiments impose additional requirements with respect to stimulus matching as compared with behavioural techniques.

4.4. ERP latency shifts with word frequency

Contrary to earlier reports (Osterhout et al., 1997; King and Kutas, 1998; Embick et al., 2001), we could not detect strong evidence for ERP latency shifts with word frequency. Two earlier studies that also used a lexical decision task and investigated peak latencies reported latency shifts for the P3 component around 560 ms (Polich and Donchin, 1988), and for an earlier left-lateralized component around 360 ms (Embick et al., 2001). Other studies using silent reading paradigms also found latency shifts with word frequency (King and Kutas, 1998; Osterhout et al., 1997), though Brown et al. (1999) could not confirm this finding.

One can argue that the absence of a statistically significant latency difference between low and high frequency words in the present study may be due to a too small sample size ($n = 12$). To sustain this argument, we should at least have observed a trend in our data, since our sample size is comparable to most of the studies that found latency shifts. Instead, as demonstrated in Fig. 6, we either did not find any systematic latency shifts at all, or the small trends that were observable were smaller than the unsystematic variation at other electrodes.

In order to quantify that our sample size was sufficient to detect latency shifts reported by earlier studies, we performed a power analysis of our statistical approach. We focused on the time range around 300 ms, where several previous studies reported shorter ERP latencies for high compared to low frequency words (Osterhout et al., 1997; King and Kutas, 1998; Embick et al., 2001). We took the difference of the means reported by Embick et al. (2001) between their lowest and highest frequency items (45 ms) and found that for $n = 12$ and a standard error of the difference of 17 ms as estimated from our RMS peak latency analysis, the statistical power to detect this latency shift for an $\alpha = 0.05$ is 99% (Howell, 2002). Since we did not see any significant effect, we conclude that such an effect was not present in our data set.

It may be that stimulus selection accounts for this discrepancy. Osterhout and Bersick (1997) and King and Kutas (1998) included extremely frequent words (> 1000 occurrences per million). Embick et al. (2001) and King and Kutas (1998) included very rare words (< 1 occurrences per million). We did not include very rare items, because they might be completely unknown to some of the subjects, thus behaving like meaningless pseudowords to these individuals. If exceptionally common words, as those used by King and Kutas (1998) and Osterhout and Bersick (1997), are crucial for ERP latency shifts, it must be noted that the highest frequency range mostly comprises grammatical function words, implying that word class (content or open class vs. function or closed class words) is a potential confound. It is possible that the latency shift of ERP components with word frequency depends on the inclusion of exceptionally common or exceptionally rare lexical items.

4.5. Implications for psycholinguistic theories of word processing

Our data address two important questions with respect to lexical access: First, which is the time range in which lexical access occurs? Second, how is the lexicon organized in the human brain? Behavioural data suggest that effects of word frequency or lexical familiarity on lexical decision times reflect the recall of information from the mental lexicon (Gernsbacher, 1984). Balota and Chumbley (1984), however, report larger effects in their lexical decision compared to category verification and pronunciation tasks, and conclude that word frequency effects are produced at the decision stage after lexical access. Bentin and Peled (1990) argued that “processing of the word’s meaning is the default action of the word perception mechanism.” However, they found the largest differences between words and pseudowords in a semantic decision task, marginally significant differences in a lexical decision task, and no differences in a rhyme task. Their interpretation is that the degree to which different aspects of the word representations are activated can depend on the information necessary to complete the task. In any case, our word frequency effects on ERP data can provide an upper time limit for the time range of lexical access: Since we found these effects already at 150 ms after stimulus onset, which is in line with data from previous studies (Pulvermüller et al., 1995; Sereno et al., 1998; Assadollahi and Pulvermüller, 2001b), we argue that lexical access must happen at or already before this time point.

The amplitude modulation of the ERP by word frequency points towards neuronal plasticity within the brain network underlying visual word recognition: The more often a word is encountered, the more efficient the synaptic connections representing this word in the network become, such that less activation is necessary to retrieve the corresponding word. This would be expected if the network representing a word was established following a correlation principle, as in the framework of Hebbian cell assemblies suggested by Pulvermüller (1999).

Acknowledgements

We are grateful to Gabriele Holz for her assistance in data acquisition, Maarten van Casteren and Mike Ford in the selection of stimulus material, and Ian Nimmo-Smith and Matt Davis for their advice on statistical issues. We would also like to thank two anonymous reviewers for their valuable suggestions.

References

Assadollahi R, Pulvermüller F. Neuromagnetic evidence for early access to cognitive representations. Neuroreport 2001a;12(2):207–13.

Assadollahi R, Pulvermüller F. Word length and frequency in early lexical access: neuromagnetic evidence (Society for Psychophysiological Research, Abstracts of the 41st Annual Meeting). Psychophysiology 2001b;38(Suppl 1):S21.

Baayen RH, Piepenbrock R, van Rijn H. The CELEX lexical database (CD-ROM). Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania; 1993.

Balota DA, Chumbley JI. Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. J Exp Psych Hum 1984;10(3):340–57.

Bentin S, Peled BS. The contribution of task-related factors to ERP repetition effects at short and long lags. Mem Cognit 1990;18:359–66.

Berg P, Scherg M. A multiple source approach to the correction of eye artifacts. Electroenceph clin Neurophysiol 1994;90:229–41.

Brown CM, Hagoort P, ter Keurs M. Electrophysiological signatures of visual lexical processing: open- and closed-class words. J Cogn Neurosci 1999;11(3):261–81.

Embick D, Hackl M, Schaeffer J, Kelepir M, Marantz A. A magnetoencephalographic component whose latency reflects lexical frequency. Cogn Brain Res 2001;10:345–8.

Francis WN, Kucera H. Frequency analysis of English usage: lexicon and grammar. Boston, MA: Houghton Mifflin; 1982.

Gernsbacher MA. Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness and polysemy. J Exp Psychol 1984;113:256–81.

Halgren E, Dhond RP, Christensen N, Van Petten C, Marinkovic K, Lewine JD, Dale AM. N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. Neuroimage 2002;17(3):1101–16.

Holcomb PJ, Grainger J, O'Rourke T. An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. J Cogn Neurosci 2002;14(6):938–50.

Howell DC. Statistical methods for psychology, 5th ed. Pacific Grove, CA: Duxbury; 2002.

Just MA, Carpenter PA. A theory of reading: from eye fixations to comprehension. Psychol Rev 1980;87(4):329–54.

King JW, Kutas M. Neural plasticity in the dynamics of human visual word recognition. Neurosc Lett 1998;244:61–4.

Kucera H, Francis WN. Computational analysis of present-day American English. Providence, RI: Brown University Press; 1967.

Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 1980;207:203–5.

Lavidor M, Ellis AW. Word length and orthographic neighborhood size effects in the left and right cerebral hemispheres. Brain Lang 2002;80:45–62.

Loftus GR, Masson MEJ. Using confidence-intervals in within-subject designs. Psychon Bull Rev 1994;1:476–90.

McCarthy G, Wood CC. Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. Electroenceph clin Neurophysiol 1985;62:203–8.

Mohr B, Pulvermüller F, Zaidel E. Lexical decision after left, right and bilateral presentation of content words, function words and non-words: evidence for interhemispheric interaction. Neuropsychologia 1994;32:105–24.

Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 1971;9:97–113.

Osterhout L, Bersick M, McKinnon R. Brain potentials elicited by words: word length and frequency predict the latency of an early negativity. Biol Psychol 1997;46:143–68.

Polich J, Donchin E. P300 and the word frequency effect. Electroenceph clin Neurophysiol 1988;70:33–45.

Pulvermüller F. Words in the brain's language. Behav Brain Sci 1999;22(2):253–79.

Pulvermüller F, Lutzenberger W, Birbaumer N. Electrocortical distinction of vocabulary types. Electroenceph clin Neurophysiol 1995;94:357–70.

Smith ME, Halgren E. Event-related potentials during lexical decision: effects of repetition, word frequency, pronounceability, and concreteness. Electroenceph clin Neurophysiol Suppl 1987;40:417–21.

Rubinstein H, Garfield L, Millikan JA. Homographic entries in the internal lexicon. J Verb Learn Verb Behav 1970;9:487–94.

Rugg MD. Event-related brain potentials dissociate repetition effects on high- and low-frequency words. Mem Cognit 1990;18(4):367–79.

Scarborough DL, Cortese C, Scarborough HS. Frequency and repetition effects in lexical memory. J Exp Psychol Hum 1977;3(1):1–17.

Sereno SC, Rayner K, Posner MI. Establishing a time-line of word recognition: evidence from eye movement and event-related potentials. Neuroreport 1998;9:2195–200.

Van Petten C, Kutas M. Interactions between sentence context and word frequency in event-related brain potentials. Mem Cognit 1990;18(4):380–93.

Whaley CP. Word-nonword classification time. J Verb Learn Verb Behav 1978;17:143–54.

Zipf GK. Psycho-biology of language. Boston, MA: Houghton Misslien; 1935.