Neuropsychologia 57 (2014) 179–190

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

More ways than one: ERPs reveal multiple familiarity signals in the word frequency mirror effect

Emma K. Bridger [a,*], Regine Bader [a,b], Axel Mecklinger [a]

[a] *Experimental Neuropsychology Unit, Department of Psychology, Saarland University, 66123 Saarbrücken, Germany*
[b] *Department of Cognitive and Clinical Neuroscience, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Square J 5, 68159 Mannheim, Germany*

A R T I C L E I N F O

Article history:
Received 5 December 2013
Received in revised form
9 February 2014
Accepted 12 March 2014
Available online 24 March 2014

Keywords:
Absolute familiarity
Event-related potentials
FN400
Mirror effect
N400
Recollection
Relative familiarity
Word frequency

A B S T R A C T

Recent dual-process models of the word frequency mirror effect place absolute familiarity, an item's baseline familiarity at a given time point, as responsible for false alarm differences and recollection for hit rate differences between high and low frequency items. One of the earliest dual-process propositions, however, posits an additional relative familiarity mechanism which is sensitive to recent presentation but relative to the absolute familiarity of a particular item (Mandler, 1980). In this study, it was possible to map these three mechanisms onto known event-related potential (ERP) effects in an old/new recognition task with high and low frequency words. Contrasts between ERPs elicited by high and low frequency new items were assumed to index absolute familiarity, and the distribution of this effect from 300 to 600 ms was topographically distinct from a temporally-overlapping midfrontally-distributed old/new effect which was larger for low than high frequency words, as would be expected from a relative familiarity mechanism. A later left parietal old/new effect, strongly linked to recollection, was only present for low frequency items. These frequency-sensitive amplitude differences for both old/new effects disappeared in a second recognition task in which old/new decisions were made under a time constraint, although the posterior absolute familiarity effect remained unaffected by the speeding of responses. The data support the assertion that three distinct recognition processes are affected by word frequency in recognition memory tasks, and the qualitatively distinct distributions associated with the two familiarity contrasts support the presence of two cognitively distinct familiarity mechanisms.

1. Introduction

Familiarity refers to one of two independent processes which can be used to support recognition memory judgments, as is postulated by dual process-models of recognition memory (see Yonelinas, 2002). Familiarity is usually described as a sense or feeling of oldness; an experience which is qualitatively distinct from recollection, which supports the reinstatement of explicit contextual details associated with the encoding episode (although reports in which familiarity appears to contribute to contextual-like retrieval are accumulating; see Kriukova, Bridger, & Mecklinger, 2013; Mollison & Curran, 2012 for recent examples; and Yonelinas, Aly, Wang, & Koen, 2010 for a review). One criticism recently leveled at dual-process accounts is that, whereas the phenomenon of recollection-based remembering is relatively well-described, familiarity is not as clearly characterized, making it easier under some conditions to define familiarity on the basis of what it is not (i.e. recollection) rather than what it is (Leynes & Zish, 2012; Voss, Lucas, & Paller, 2012). In this study, the issue of characterizing familiarity will be addressed by re-focusing on two memory phenomena which have contributed some of the most important evidence for dual-process models to date: event-related potential (ERP) correlates of recognition and the word frequency mirror effect (Glanzer & Adams, 1990).

The word frequency mirror effect describes the phenomenon by which, compared to high frequency words, low frequency items elicit more correct responses in recognition memory tasks both when they are old (an increase in hit rates) and new (a decrease in false alarm rates). This pattern is problematic for single-process signal-detection models, which presume the placement of an old/new decision criterion along a continuum of memory strength, because a simple strength mechanism cannot predict both the hit and false alarm rate without incorporating additional parameters which make these models unjustifiably complex (DeCarlo, 2007; Hintzman, 1994; Murdock, 1998). Dual-process models have dealt with this issue by positing that the hit and false alarm rates reflect the respective contributions of recollection and familiarity (Joordens & Hockley, 2000; Reder et al., 2000). A greater level of pre-experimental familiarity for high frequency items is assumed

* Corresponding author. Tel.: +49 681 302 58082.
E-mail address: e.bridger@mx.uni-saarland.de (E.K. Bridger).

http://dx.doi.org/10.1016/j.neuropsychologia.2014.03.007
0028-3932/

to make it more likely that high frequency new items will fall above the old/new criterion and thus be incorrectly classified as old. At the same time, low frequency items are inherently more distinctive, in part as a consequence of the low number of contexts in which they have previously been experienced. This increases the likelihood in which the experimental context associated with a low frequency item is recollected during an experimental test phase which drives the increased hit rate for these items. This pattern tallies with a number of data points including those which show that low frequency hits are disproportionately supported by judgments associated with a feeling of remembering, thought to correspond with recollection (Reder et al., 2000). Of particular interest for the current report are data from experimental designs which have capitalized on the assumption that familiarity occurs faster than recollection-based remembering. When participants are required to respond at an early time point by which it should not be possible to make recollection-based responses, the hit advantage for low frequency items has been shown to be removed whilst the false alarm difference remains unchanged (Balota, Burgess, Cortese, & Adams, 2002; Joordens & Hockley, 2000).

In their recent paper, Coane, Balota, Dolan, and Jacoby (2011) combined this dual-process model with another mechanism proposed to contribute to the word frequency mirror effect: a relative familiarity mechanism, which refers to the change in familiarity strength relative to its pre-experimental familiarity level after an item has been presented during the study phase of an experiment. A mechanism of this kind would necessarily elicit a greater feeling of familiarity for low than high frequency *old* items, because low frequency items would experience a greater change in strength following presentation than high frequency items, due to their relatively low pre-experimental familiarity level. This characterization corresponds with Mandler's (1980) original definition of incremental familiarity as it was termed in one of the earliest dual-process models of recognition. Mandler's perspective provides a simple mechanism by which familiarity would be larger for low than high frequency words which is described in the following equation:

$$\text{Relative familiarity} = d/(d+F)$$

where d is the incremental strength increase (following a single presentation), and F refers to the level of pre-experimental familiarity for a particular class of item (hereafter referred to as absolute familiarity in keeping with the terminology employed by Coane et al., 2011). The important aspect of this relative mechanism is that it necessarily presupposes the role of two distinct familiarity mechanisms, because: (i) it depends mathematically upon an index of absolute familiarity and (ii) it cannot explain the increase in FA rates for high compared to low frequency new items because relative familiarity should always be greater for low frequency items. Thus, if a relative familiarity component does exist, it cannot do so in the absence of differences in absolute familiarity.

To determine whether both absolute and relative familiarity contribute to recognition performance, Coane et al. (2011) adapted Jacoby's (1991) two-list exclusion task in which only a proportion of old words presented at test are to be endorsed as old depending upon their study context. The particular experimental set-up ensured that the successful exclusion of those old items from a non-targeted study context depended upon recollection, such that failures to exclude these items (exclusion errors) would be considerably greater under speeded response conditions in which response decisions need to be made before recollection is thought to be available. At these earlier response times, only early familiarity-type mechanisms should be available on which to base recognition judgments. If these early recognition processes include a relative familiarity mechanism, then the number of exclusion errors (old items which should have been excluded on the basis of their study context, but were not) should be significantly greater for low than high frequency old items. At the same time, if an additional absolute familiarity signal is present, the greater false alarm rate to high compared to low frequency new items should also be evident, in line with the standard mirror effect. Across two experiments with different response deadline implementations, Coane and colleagues observed significantly more exclusion errors for low than high frequency old words and a less reliable, but nonetheless broad trend for increased false alarms to new items for high than low frequency, as would be expected if two distinct familiarity mechanisms were contributing to response judgments in this paradigm. The finding that, under response conditions which ostensibly removed recollection, more exclusion errors were made to low than high frequency old items, provides the clearest behavioural demonstration that a relative familiarity-type signal contributes to recognition differences to high and low frequency words.

These outcomes are in line with the intriguing notion that word frequency manipulations can dissociate three distinct processes which contribute to recognition judgments: absolute familiarity for new items and relative familiarity and recollection for old items. It is worth bearing in mind, however, that this pattern was observed in Coane et al.'s study using a demanding exclusion/response-deadline task combination which differs somewhat from standard recognition conditions and which may in part be responsible for the fact that a robust increase in false alarms for high frequency items was not observed. One worthwhile approach towards providing convergent evidence for this pattern with more typical recognition task parameters would be to exploit the capacity of ERPs to index functionally distinct but temporally overlapping or contiguous processes. These characteristics have been successfully and robustly used to dissociate distinct ERP old/new effects (contrasts between ERPs elicited by correctly responded to old and new items) across a variety of recognition paradigms (Mecklinger, Brunnemann, & Kipp, 2011; Rugg & Curran, 2007; Wilding & Herron, 2006). Of most interest for the current topic are dissociations between an early old/new effect which usually elicits a midfrontal distribution from 300 to 500 ms post-stimulus and a later occurring old/new effect which is largest over left parietal sites around 500–700 ms. The earlier of the two effects, often referred to as the midfrontal old/new effect or the FN400, has been shown to operate in a way which is consistent with an index of familiarity. Examples of this are demonstrations that the effect varies with subjective familiarity strength (Woodruff, Hayama, & Rugg, 2006; Yu & Rugg, 2010) but does not distinguish old items and semantic lures (Curran, 2000; Nessler, Mecklinger, & Penney, 2001; Rugg & Curran, 2007). The later left parietal effect has been shown to correlate reliably with recollection-based responding and the amount of information recollected (Vilberg & Rugg, 2009) and is significantly larger for responses associated with a correct compared to an incorrect source attribution (Wilding & Rugg, 1996). These ERP effects cannot provide an exact 1:1 mapping of familiarity and recollection because other processes can elicit comparable functional modulations of these effects (see Voss et al., 2012, for considerations of this kind). With these restrictions in mind, these effects can nonetheless be usefully employed as putative neural correlates of familiarity and recollection (Rugg & Curran, 2007) given those reports in which the two effects have been shown to doubly dissociate in a manner corresponding with dual-process models of recognition memory (Jäger, Mecklinger, & Kipp, 2006; Stenberg, Hellman, Johansson, & Rosén, 2009; Woodruff et al., 2006).

A number of reports have previously investigated the impact of word frequency on ERP old/new effects. In a series of experiments, Rugg and colleagues (Rugg, 1990; Rugg, Cox, Doyle, & Wells, 1995; Rugg & Doyle, 1992) reported a significant late old/new effect (which likely corresponds to what is now referred to as the left parietal old/new effect) only for low frequency words, in line with the prediction derived from dual-process models that recollection is more likely to occur for these items. The pattern in the earlier time window in which familiarity-processing is likely to have been measured is less clear: whereas in one study there was no obvious interaction with word frequency (Rugg & Doyle, 1992), in another, the early effects were only significant for low frequency items (Rugg, 1990). This latter finding is consistent with an index of relative familiarity but it was not possible to address the topography of this effect due to the small number of electrodes employed. Moreover, participants were not required to make recognition decisions in this task complicating subsequent interpretations. These patterns were observed before the bulk of evidence in favour of a dual-process interpretation of ERP old/new effects had been accumulated, and this, combined with the lack of topographical comparisons, makes it difficult to determine the extent to which these effects map onto those typically reported in episodic recognition tasks of this kind. In a more recent report, Stenberg et al. (2009) manipulated frequency using Swedish names and observed a robust early midfrontal old/new effect for less frequent (rare) but not for highly frequent (common) names. This finding provides some of the clearest evidence that the midfrontal old/new effect operates in a way which is consistent with relative familiarity.

Not all early ERP old/new effects associated with familiarity-based responding have shown a midfrontal distribution. In one report, Mackenzie and Donaldson (2007) asked participants to learn a series of face-name associations before completing a recognition task in which they were required to say whether or not they remembered the name or any other information for old faces. The old/new effect which was associated with faces that were correctly recognized without the retrieval of any specific information (familiarity-based responding in the absence of recollection) exhibited a clear posterior distribution in the early time window. The authors suggested that this posterior effect might reflect the use of a distinct absolute familiarity mechanism for the recognition of faces compared to the relative familiarity effect observed in studies in which word stimuli have been employed. This reasoning comes about by considering differences in absolute familiarity for words and faces: whereas all words associated with known concepts will fall somewhere upon the continuum of absolute familiarity with a considerable degree of variance, faces of unknown individuals are all likely to be low in absolute familiarity and to fall on similar or the same points on the scale. MacKenzie and Donaldson suggested that the very low level of absolute familiarity for faces might have enabled changes in absolute familiarity to be diagnostic of prior occurrence. Further support for the possibility that a distinct familiarity mechanism might be employed for the recognition of stimuli with no pre-experimental familiarity comes from a different paradigm, in which Bader, Mecklinger, Hoppstädter, and Meyer (2010, see also; Wiegand, Bader, & Mecklinger, 2010) reported an early posterior old/new effect associated with the correct recognition of word pairs which had been introduced during study as novel compound words, such as smoke-apple ("a fruit maturing above flames"). Also noteworthy is the strong correspondence between these posterior old/new effects and the N400 correlate of semantic processing (Kutas & Federmeier, 2000). The N400 is a negative ERP component that peaks around 400 ms after stimulus onset and is most pronounced over posterior scalp sites. It is sensitive to the ease of processing semantic information and also larger for low than high frequency words when presented in isolation (Kutas & Federmeier, 2000; Van Petten & Kutas, 1990; Young & Rugg, 1992).

To summarize, there is considerable evidence within the ERP literature of an early midfrontal old/new effect which operates in line with an index of relative familiarity and which dissociates both temporally and functionally from a later parietal effect associated with recollection. It is also the case that a topographically distinct posterior old/new effect occurs in a time window which overlaps with that of the early frontal effect. The functional significance of this posterior effect in recognition tasks has yet to be specified but it is observed in conditions in which absolute familiarity might be more diagnostic of recent presentation and is similar in topography to the N400, which is sensitive to absolute word frequency. The goal of the current experiment was to bring together separate behavioural and ERP findings which implicate the presence of multiple familiarity signals, and to test for these signals by determining whether distinct ERP correlates of absolute familiarity and relative familiarity could be observed within a single recognition test containing a simple high and low word frequency manipulation. In a first study-test phase, participants completed a simple recognition test with standard old/new response requirements whilst EEG was recorded. In a second study-test phase, all parameters remained the same except that participants were required to make their old/new decisions within 750 ms of seeing the test item. This response manipulation was designed to reduce the extent to which recollection could contribute to recognition decisions in order to determine whether a hit advantage for low frequency old items would remain when recollection was no longer available, as would be expected if an early relative familiarity mechanism was also in operation for these words. A deadline of 750 ms was chosen because it is comparable to those used in a variety of behavioural reports in which a deadline manipulation has been shown to qualitatively influence the pattern of recognition responding in line with the selective removal of recollection-based responding (Bowles et al., 2007; Coane et al., 2011; Hintzman & Curran, 1994).

In line with the logic described above (and depicted in Fig. 1), absolute familiarity should be revealed by contrasting ERPs elicited by high and low frequency new items, whereas relative familiarity should be indexed by an old/new effect which is significantly larger for low than high frequency items and which is distinct from a later recollection old/new effect which should also be larger for low than high. If multiple early familiarity signals exist, then these contrasts should reveal ERP effects with distinct scalp topographies within overlapping early time windows, in line with the assumption that non-overlapping scalp distributions reflect qualitatively distinct neural generators or neural configurations and thus represent

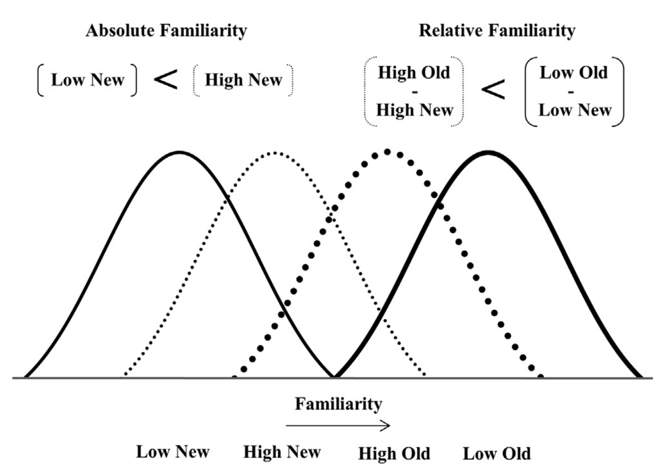

Fig. 1. An illustration of the mirror effect with underlying normal distributions with equal variance. The critical ERP contrasts of interest are depicted as follows: differences between new items wherein high frequency items are more positive should index absolute familiarity, whereas old/new item effects which are significantly bigger for low than high frequency items index relative familiarity.

functionally distinct cognitive processes (Rugg & Coles, 1995; Wilding, 2006).

2. Methods

2.1. Participants

Twenty-two native German speakers were recruited from the student population of Saarland University. All except one participant were right-handed as assessed by the Edinburgh Handedness Inventory (laterality quotient > 50; Oldfield, 1971) and had no known neurological problems. Informed consent was required, payment was provided at a rate of €8/h, and participants were debriefed after the experiment. The experiment was approved by the local ethics committee of Saarland University. Data from 4 students were excluded from the final analyses either because there were less than 16 artefact-free trials which could contribute to each critical ERP average ($n=1$) or had a discrimination value (p[hit] − p[false alarm]) at or below zero in the speeded response condition ($n=3$). The mean age of the remaining 18 participants (7 males) was 22 years (range = 18–28 years).

2.2. Stimuli and design

Stimuli were 480 high and low frequency German words all taken from the WebCelex database (Baayen, Piepenbrock, & Gulikers, 1995) and between 4 and 14 characters in length. The final high and low frequency lists each comprised 60 nouns, 90 adjectives and 90 verbs. Example high frequency words (and English translations) are TEIL ('part'), SPRECHEN ('to speak') and FREI ('free'), whereas representative low frequency examples are SICHEL ('sickle'), MELKEN ('to milk') and SALOPP ('sloppy'). Semantic repetitions across word categories were removed (e.g. the noun 'muscle' would be removed if the stimuli set already included the adjective 'muscular') to minimize overlap between nouns, verbs and adjectives. Table 1 reports the frequency values for high and low frequency words, both according to WebCelex as well as for an alternative database (available at www.dlexdb.de; Heister et al., 2011) which has recently been shown to better approximate lexical decision times than the older Celex corpus (Brysbaert et al., 2011). Each frequency list was split into 4 lists of 60, each matched for word length, frequency and proportion of nouns, verbs and adjectives. In each study phase, participants were presented with 60 high and 60 low frequency words, which were re-presented at test along with another 60 high and 60 low frequency words. Four counterbalanced lists ensured that items operated in both old and new and speeded and non-speeded conditions across participants. The programme ensured that the order of presentation of items at study and test comprised a distinct randomised order for each participant. Stimuli were presented in the centre of the screen in black on a grey background. The vertical visual angle subtended 1.1° and the horizontal visual angle was between 3.8° and 13.4° (at a viewing distance of 70 cm). All stimuli were presented using E-Prime 2.0 (Psychology Software Tools).

2.3. Procedure

Each experimental session began with the fitting of the EEG cap (see parameters below). The experimental tasks began with a short practice phase to familiarize participants with the overall task construction, principally the study-test phase distinction. This practice task comprised a study phase with 10 high and low frequency words, a distractor task identical to that in the experiment proper (see below) and a short test phase without any speeded instructions. Only half of the studied practice words were re-presented at this point, whilst the remainder were employed in a final practice test phase before the actual test phase. This was necessary in order to remind participants of the exact test requirements immediately before the critical test. During study, participants were informed that they would see a series of words presented on the screen and that they were to say each word aloud. Each study trial began with a fixation cross for 500 ms which was replaced by a word presented for 1500 ms. The screen was then blanked for a 1500 ms response interval, before the next trial began. Participants took a self-paced break midway through the study phase.

After the study phase, participants completed a short distractor task, in which they were required to count backwards silently in threes from a randomly presented number for 30 s. Participants entered the final number reached into the computer at the end of this interval. The test phase was preceded by a short practice phase including the remaining words presented during the practice study phase. Test trials began with a fixation cross for 500 ms, which was replaced by the target word for 850 ms and a blank screen for 1400 ms, during which participants were required to make a simple old/new binary judgment. The next trial began 800 ms later. The mapping of old: new responses onto left: right ('c': 'm') buttons was counterbalanced across participants. Participants took self-paced breaks every 40 trials. Piloting revealed that response times remained significantly shorter in the non-speeded condition for participants who began with the speeded condition. To avoid this and to keep the timing of responding distinct in the two conditions, participants always completed the non-speeded study-test phase first, and were not aware of the speeded response requirements until after the second study phase. All parameters were identical in the non-speeded and speeded blocks, except that participants were asked to make their response whilst the test item was on-screen in the speeded block (850 ms). If participants made a response after 750 ms, they heard a warning tone and these responses were excluded from all further analyses. As was the case in the first study-test block, participants were given a short practice phase to acquaint them with this requirement.

Table 1
Length and frequency characteristics for high and low frequency words. Frequency values represent the Mannheim lemma frequency (per million) from the Celex database and normalized lemma frequency (per million) for dlexdb. Three high frequency and 17 low frequency words were not available in the dlexdb database. Ranges (min − max) are shown in parentheses.

	Length	Celex	dlexDB
High	6.66 (4–14)	221.16 (82–605)	192 (16–1114)
Low	6.77 (5–9)	1.97 (1–4)	2 (0–10)
p-value	.45	< .001	< .001

2.4. Electrophysiological recording parameters and analyses

Continuous EEG was recorded from 58 scalp locations based on the extended international 10–20 system (Jasper, 1958). EEG was acquired referenced to the left mastoid and re-referenced offline to the average of the mastoid signals. All offline and pre-processing of the data was conducted using EEProbe (ANT Software). EEG signals were band-pass filtered from DC-70 Hz and digitized at a sampling rate of 500 Hz. Electro-oculographic activity (EOG) was assessed using signals recorded from four additional electrodes above and below the right eye (vertical EOG) and on the outer canthi (horizontal EOG). Electrode impedances were kept below 5 kΩ. Offline, a digital band-pass filter (.03–30 Hz) was applied and epochs were created beginning 100 ms prior to and ending 1000 ms after the onset of stimulus presentation. Waveforms were corrected relative to the 100 ms pre-stimulus baseline period. EOG blink and movement artifacts were corrected using a modified linear regression algorithm (Gratton, Coles, & Donchin, 1983) embedded in the EEProbe Software package. The mean number of artifact-free trials contributing to individual subject grand averages for each of the four non-speeded conditions was as follows: high frequency hits = 24 (range 17–35), high frequency correct rejections = 31 (17–44), low frequency hits = 27 (18–46), low frequency correct rejections = 33 (19–48). The mean numbers of trials for each condition in the speeded response condition were as follows: high frequency hits = 24 (16–37), high frequency correct rejections = 27 (20–33), low frequency hits = 24 (16–36), low frequency correct rejections = 29 (18–43). To provide an indicator of effect sizes, partial eta squared is reported alongside outcomes from MANOVAs and ANOVAs whilst Cohen's *d* is reported for *t*-test outcomes, for analyses of both behavioural and ERP data.

3. Results

3.1. Behaviour

Table 2 shows the proportion of hits and false alarms for high and low frequency words separately for the non-speeded and speeded conditions. As expected, the pattern in the non-speeded condition comprises a typical word frequency mirror effect, with a greater proportion of hits for low than high frequency items, whilst the reverse pattern is evident for false alarms. The hit advantage for low frequency items is reduced considerably in the speeded condition, however. A 2 × 2 repeated measures ANOVA on hit responses with factors of Frequency (high, low) and Response Deadline (non-speeded, speeded) revealed a main effect of Frequency ($F(1,17)=9.49$, $p=.007$, $\eta_p^2=.358$), a marginally significant effect of Response Deadline ($F(1,17)=4.24$, $p=.055$, $\eta_p^2=.200$) and a significant interaction ($F(1,17)=6.17$, $p=.024$, $\eta_p^2=.266$). Bonferroni-corrected follow-up *t*-tests revealed that low frequency hits were significantly more likely in the non-speeded than the speeded condition ($t(17)=3.12$, Bonferroni-corrected $p=.024$, $d=.823$), and that this low frequency hit advantage (relative to high frequency hits) was significant in the non-speeded ($t(17)=3.91$, Bonferroni-corrected $p=.004$, $d=.814$) but

Table 2
Mean proportions of hits and false alarms (FAs) for high and low frequency words separated according to response deadline condition. Standard deviations are in parentheses.

Frequency	Non-speeded				Speeded			
	Hits	FAs	Pr	Br	Hits	FAs	Pr	Br
High	.55 (.14)	.19 (.09)	.36 (.11)	.31 (.15)	.52 (.16)	.30 (.14)	.22 (.11)	.39 (.17)
Low	.66 (.13)	.09 (.07)	.57 (.13)	.23 (.17)	.54 (.16)	.19 (.12)	.35 (.13)	.30 (.19)

Table 3
Mean reaction times to hits and correct rejections (CRs) for high and low frequency words separated according to response deadline condition. Standard deviations are shown in parentheses.

Frequency	Non-speeded		Speeded	
	Hits	CRs	Hits	CRs
High	906 (193)	909 (182)	592 (26)	586 (39)
Low	871 (161)	887 (181)	601 (30)	587 (36)

not the speeded condition ($p = .541$). The pattern of high frequency hits did not change across the two response deadline conditions ($p = .567$). There was thus no evidence of a significant hit advantage in the speeded condition which would have been consistent with a behavioural indicator of relative familiarity in the speeded condition. The pattern of false alarms across both response deadlines showed the expected pattern, with the increase in high frequency false alarms (relative to low frequency false alarms) also evident in the speeded condition, where the number of false alarms showed an overall increase. This pattern was confirmed by a second ANOVA on the proportion of false alarms which revealed only main effects of Frequency ($F(1,17) = 48.82$, $p < .001$, $\eta_p^2 = .742$) and Response Deadline ($F(1,17) = 14.78$, $p = .001$, $\eta_p^2 = .465$).

Response bias was indexed using the Br parameter (False Alarms/[1 − Pr]) specified by Snodgrass and Corwin (1988) for which .50 represents a neutral bias with values less than this representing a conservative bias. The average Br values for all critical conditions are shown in Table 2. A 2 × 2 ANOVA with factors of Response Deadline and Frequency, revealed a main effect of Frequency ($F(1,17) = 8.64$, $p = .009$, $\eta_p^2 = .337$) and a marginally significant effect of Response Deadline ($F(1,17) = 4.23$, $p = .055$, $\eta_p^2 = .199$), in line with a more conservative response bias for low than high frequency items and in the non-speeded than speeded condition.

Table 3 shows the mean reaction times for hits and correct rejections for high and low frequency words separately for the non-speeded and speeded conditions. In line with the response deadline cut-off, correct responses were on average 300 ms faster in the speeded condition. In the non-speeded condition, responses were also approximately 30 ms faster for low than high frequency items. A 2 × 2 × 2 ANOVA with factors of Response Deadline, Frequency and Item-Type (hit, correct rejection) revealed a main effect of Response Deadline ($F(1,17) = 62.91$, $p < .001$, $\eta_p^2 = .787$), a marginally significant effect of Frequency ($F(1,17) = 3.89$, $p = .065$, $\eta_p^2 = .186$) and an interaction between Response Deadline and Frequency ($F(1,17) = 9.28$, $p = .007$, $\eta_p^2 = .353$). Subsequent analyses separated for the non-speeded and speeded conditions revealed a main effect of Frequency only in the non-speeded condition ($F(1,17) = 6.40$, $p = .022$, $\eta_p^2 = .274$). On average, participants made 2.92 timeouts in the speeded response condition. An ANOVA on speeded timeouts with factors of Frequency and Item-Type revealed main effects of both factors (Frequency: $F(1,17) = 7.67$, $p = .013$, $\eta_p^2 = .311$; Item-Type: $F(1,17) = 4.922$, $p = .040$, $\eta_p^2 = .225$) because participants made significantly more timeouts to low frequency ($M = 3.47$, SD = 1.85) than high frequency ($M = 2.36$, SD = 2.56) and to new ($M = 3.31$, SD = 2.53) than old items ($M = 2.53$, SD = 1.98).

3.2. ERPs

Fig. 2 shows the grand average old/new effects for high and low frequency items in the non-speeded condition. Old/new effects for both classes of frequency begin from around 250 ms and remain until the end of the epoch but are markedly larger for low frequency items. Differences between high and low frequency correct rejections, the proposed index of absolute familiarity, are also apparent from approximately 300 until 600 ms, where high frequency new items are more positive going than low frequency new items. This latter difference extends across most sites but has a posterior maximum. From 300 ms, old/new effects show the characteristic morphological negativity and old/new differences for both contrasts are larger over frontal sites in line with the midfrontal old/new effect. The old/new effects retain their frontal maximum later in the epoch but from 500 to 800 ms, a large posterior positivity, the putative index of recollection, is evident for low but not high frequency hits. At the end of the recording epoch, both old and new high frequency items become relatively negative compared with low frequency items.

Fig. 3 shows these same contrasts for the speeded condition, where again old/new effects are evident from 250 ms onwards, but do not show such a marked difference in amplitude for the two classes of word frequency. As in the non-speeded condition, there are also differences between high and low frequency correct rejections in this contrast; high frequency correct rejections are more positive than low frequency from 400 to 600 ms at posterior sites, but this difference reverses in polarity after 600 ms. This late relative negativity for high frequency new items is maximal over central and frontal sites. As in the non-speeded condition, the old/new effects in the early 350–550 ms window show a frontal maximum although the effects are comparable in amplitude for high and low frequency items. In direct contrast to the non-speeded condition, the posterior positivity for hits around 500–800 ms is equivalent for both high and low frequency items. Similarly, robust old/new effects appear at frontal sites until the end of the recording epoch, and these are comparable in amplitude for both classes of frequency, although low frequency words at these sites are generally more positive than high frequency words.

Mean amplitude data were taken from the 15 electrodes depicted in Figs. 2 and 3. MANOVAs were employed because of their greater statistical power, and in line with the guidelines specified by Dien and Santuzzi (2005), data from five electrodes in each chain of electrodes at frontal (F), central (C) and parietal (P) sites, were thus averaged to create three levels of an anterior–posterior (AP) location factor. The five electrodes in each chain were left midlateral (3), left superior (1), midline (z), right superior (2) and right midlateral (4). This measure was used in order to reduce the number of levels of location factors and degrees of freedom by collapsing across adjacent electrodes which commonly covary. Old/new MANOVAs comprised factors of Response Deadline (non-speeded, speeded), Frequency (high, low), Item-Type

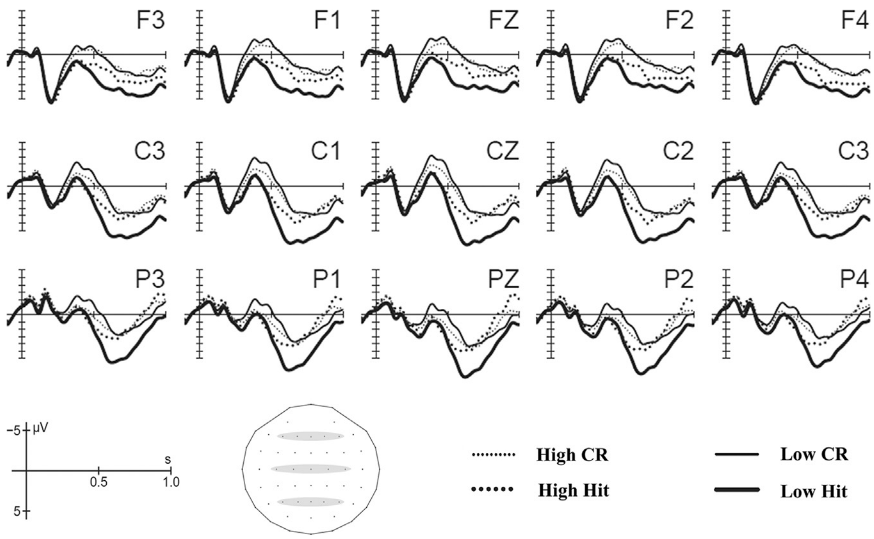

Fig. 2. Grand average ERPs elicited by hits and correct rejections (CR) for the non-speeded response condition. ERPs are separated according to word frequency. Data are shown for the 15 electrode locations used in all analyses at frontal (F3, F1, Fz, F2, F4), central (C3, C1, Cz, C2, C4), and parietal (P3, P1, Pz, P2, P4) scalp sites. A 12 Hz low pass filter has been applied to ERP waveforms for the purpose of illustration only.

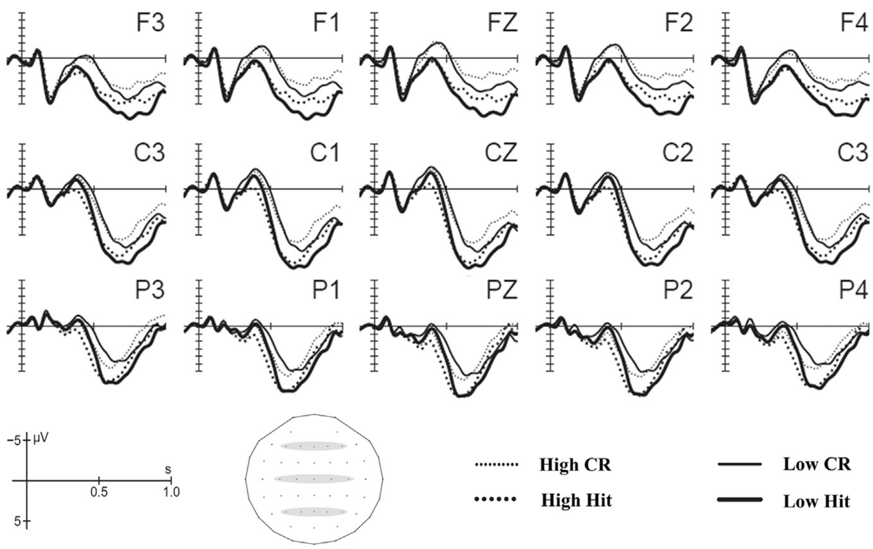

Fig. 3. Grand average ERPs elicited by hits and correct rejections (CR) for the speeded response condition. ERPs are separated according to word frequency.

(hits, correct rejections) and AP (frontal, central, posterior). For the quantification of absolute familiarity, new item comparisons included factors of Response Deadline (non-speeded, speeded), Frequency (high, low) and AP (frontal, central, posterior). The two 200 ms time windows employed for the old/new contrasts were defined a priori so as to encompass the known peaks of the two effects (Mecklinger, 2006; Rugg & Curran, 2007) as well as to correspond with time windows used in recent reports (e.g. Kriukova, Bridger, & Mecklinger, 2013) and mean amplitudes for these analyses were thus taken from 350 to 550 and 550–750 ms post-stimulus. For the analyses of the ERPs to correctly rejected new items, for which the time course was not clearly defined a priori, analyses were conducted in a 300–600 ms time window selected via visual inspection of the grand average data, as well as in the 350–550 and 550–750 ms time windows to aid comparison with the old/new effect analyses.

3.2.1. *Old/new ERP contrasts*

3.2.1.1. 350–550 ms. A main effect of Item-Type ($F(1,17)=35.52$, $p<.001$, $\eta_p^2=.676$) was observed in the 350–550 ms window, but this was moderated by an interaction between Response Deadline, Frequency and Item-Type ($F(1,17)=6.82$, $p=.018$, $\eta_p^2=.286$) and marginally significant interactions between Item-Type and AP ($F(2,16)=3.25$, $p=.065$, $\eta_p^2=.289$) and between Response Deadline, Frequency, Item-Type and AP ($F(2,16)=3.57$, $p=.052$, $\eta_p^2=.308$). Subsequent MANOVAs were thus separately conducted in the two response deadline conditions. The Item-Type main effect ($F(1,17)=45.23$, $p<.001$, $\eta_p^2=.727$) was moderated by an interaction with Frequency ($F(1,17)=11.24$, $p=.004$, $\eta_p^2=.398$) in the non-speeded condition. The Item-Type by AP interaction was also marginally significant ($F(1,17)=3.46$, $p=.056$, $\eta_p^2=.302$). Separate contrasts for each frequency condition in the non-speeded condition revealed main effects of Item-Type for both levels of Frequency (high: $F(1,17)=13.79$, $p=.002$, $\eta_p^2=.448$; low: $F(1,17)=55.20$, $p<.001$, $\eta_p^2=.765$) and an Item-Type by AP interaction for low frequency items ($F(2,16)=4.18$, $p=.035$, $\eta_p^2=.343$). Two (Item-Type) by two (adjacent levels of AP) MANOVAs for the low frequency items revealed an interaction only when central and posterior sites were compared ($F(1,17)=4.90$, $p=.041$, $\eta_p^2=.224$) in line with the maximum of this effect over frontal and central sites (see Figs. 2 and 4). Old minus new difference values were calculated

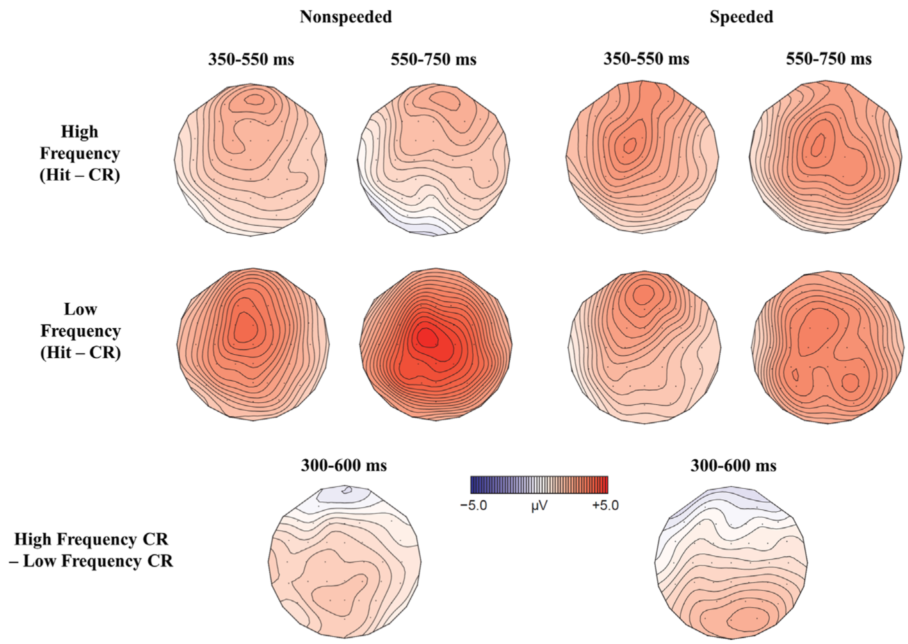

Fig. 4. Topographic maps showing the scalp distributions of the differences between ERPs associated with the critical contrasts. Contrasts from the non-speeded condition are shown on the left side and speeded contrasts are depicted on the right. Hits minus correct rejections (CR) comprise the upper row whereas new item contrasts are shown underneath. Maps are computed on the basis of the mean difference scores taken from the respective time windows (300–600, 350–550 or 550–750 ms) and are all depicted along the same scale.

separately for the high and low frequency conditions over frontal and central sites. At both levels of AP, low frequency old/new effects were significantly bigger than high frequency old/new effects (F: $t(17)=3.24$, $p=.005$, $d=.660$; C: $t(17)=3.37$, $p=.004$, $d=.832$). These analyses thus reveal a fronto-centrally distributed early old/new effect that was robustly larger for low than high frequency items in the non-speeded condition.

For the speeded condition, there was also a main effect of Item-Type ($F(1,17)=16.73$, $p=.001$, $\eta_p^2=.496$) and a marginally significant three-way interaction between Frequency, Item-Type and AP ($F(2,16)=3.61$, $p=.051$, $\eta_p^2=.311$). Separate Item-Type by AP analyses conducted separately for the two Frequency conditions revealed main effects of Item-Type (high: $F(1,17)=13.86$, $p=.002$, $\eta_p^2=.449$; low: $F(1,17)=11.18$, $p=.004$, $\eta_p^2=.397$) but no interactions between Item-Type and AP in either instance. Follow-up Frequency by Item-Type analyses conducted separately for each level of the AP factor revealed significant main effects of Item-Type at all three AP levels (frontal: $F(1,17)=15.15$, $p=.001$, $\eta_p^2=.471$; central: $F(1,17)=13.34$, $p=.002$, $\eta_p^2=.440$; parietal: $F(1,17)=15.62$, $p=.001$, $\eta_p^2=.479$) but no interactions between Frequency and Item-Type. Paired *t*-tests between high and low old/new differences revealed no significant differences at any level of AP (all $ps>.257$). The original marginal frequency, Item-Type by AP interaction is thus likely to reflect the generally more positive ERPs for high than low frequency items at posterior but not anterior sites. Robustly larger early old/new effects for low than high frequency items were thus not evident for the speeded condition in this time window.

3.2.1.2. 550–750 ms. A main effect of Item-Type ($F(1,17)=45.65$, $p<.001$, $\eta_p^2=.729$) was also observed in the 550–750 ms window, and was moderated by interactions between Frequency and Item-Type ($F(1,17)=11.93$, $p=.003$, $\eta_p^2=.412$), Response Deadline, Frequency and Item-Type ($F(1,17)=5.64$, $p=.030$, $\eta_p^2=.249$) and a four-way Response Deadline, Frequency, Item-Type by AP interaction ($F(2,16)=7.73$, $p=.004$, $\eta_p^2=.491$). Follow-up MANOVA in the non-speeded condition revealed a main effect of Item-Type ($F(1,17)=19.52$, $p<.001$, $\eta_p^2=.534$), a Frequency by Item-Type interaction ($F(1,17)=20.39$, $p<.001$, $\eta_p^2=.545$) and a marginally significant interaction between Frequency, Item-Type and AP ($F(2,16)=3.57$, $p=.052$, $\eta_p^2=.309$). Separate contrasts for each frequency condition revealed a main effect of Item-Type for low frequency items ($F(1,17)=34.43$, $p<.001$, $\eta_p^2=.669$) which was moderated by an Item-Type by AP interaction ($F(2,16)=3.89$, $p=.042$, $\eta_p^2=.327$). Two (Item-Type) by two (adjacent levels of AP) MANOVAs for the low frequency items revealed a trend for an interaction only when frontal and central sites were compared ($F(1,17)=3.10$, $p=.096$, $\eta_p^2=.154$) in line with the maximum of this effect over central and posterior sites (see Figs. 2 and 4). Only the Item-Type by AP interaction was significant for high frequency items ($F(2,16)=4.48$, $p=.028$, $\eta_p^2=.359$) and two (Item-Type) by two (adjacent levels of AP) MANOVAs revealed an interaction only when central and posterior sites were compared ($F(1,17)=8.29$, $p=.010$, $\eta_p^2=.328$). These outcomes reveal a striking distinction between high and low frequency old/new effects in this time window: whereas a robust old/new effect was present at all sites and largest over central and posterior sites for low frequency items, the high frequency old/new effect was only significant at frontal sites.

In the speeded condition, only the main effect of Item-Type was significant ($F(1,17)=24.00$, $p<.001$, $\eta_p^2=.585$) in line with a broadly distributed old/new effect which was not modified by frequency type.

3.2.2. New item ERP contrasts

As noted above, inspection of Figs. 2 and 3 also reveals a divergence between ERPs elicited by new items from approximately 300–600 ms in both response deadline conditions. This effect takes the form of a relative positivity for high frequency new items which is most obvious at posterior sites and is in line with the predicted ERP correlates of absolute familiarity. In order to test for the significance of this effect, additional MANOVAs on ERPs elicited by correct rejections were initially conducted with factors of Response Deadline, Frequency and AP in the same time windows as the preceding analyses. A main effect of Frequency (F(1,17)=4.83, p=.042, η_p^2=.221) and an interaction between Frequency and AP (F(2,16)=5.05, p=.020, η_p^2=.387) were observed in the early time window (350–550 ms). Follow-up analyses at each level of AP only revealed a main effect of Frequency at parietal sites (F(1,17)=16.55, p=.001, η_p^2=.493). There were no significant effects of Frequency in the 550–750 ms time window. A final MANOVA was conducted on mean amplitude data from 300 to 600 ms which encompasses the entirety of the positive-going ERP effect for high frequency new items over posterior sites (see Figs. 2 and 3). A main effect of Frequency (F(1,17)=5.01, p=.039, η_p^2=.228) was again moderated by an interaction with AP (F(2,16)=5.20, p=.018, η_p^2=.394). Follow-up analyses at each level of AP, revealed a main effect of Frequency at parietal sites (F(1,17)=19.41, p=.001, η_p^2=.533) and a marginally significant main effect at central sites (F(1,17)=3.88, p=.065, η_p^2=.186) only, again reflecting the posterior emphasis of this effect throughout the majority of this time window.

3.2.3. *Analysis of scalp distributions*

The current ERP data thus reveal both a new item difference over posterior sites which did not vary across the speeded and non-speeded condition and a pattern of ERP old/new effects which did interact with response deadline condition. Under non-speeded response requirements, an early frontally-focused old/new effect was considerably larger for low than high frequency items; a pattern which remained in the subsequent later time window where the low frequency old/new effect showed a robust parietal old/new effect, which was not evident for high frequency items. This pattern is consistent with that predicted by dual-process models of the mirror effect, insofar as the ERP old/new effect generally thought to be associated with recollection was present for low but not for high frequency words and these models assume that recollection occurs only or to a considerably greater degree for low frequency words by virtue of their distinctiveness (e.g. Reder et al., 2000). Moreover, evidence that the earlier frontally distributed old/new effect was larger for low frequency than high frequency items is in line with the notion that this effect is an index of relative familiarity. Of additional interest is the fact that although the late parietal old/new effect was largest for low frequency items and not observed for high frequency in the non-speeded condition, it was of comparable amplitude for both frequency-types in the speeded test condition.

In order to provide further support for the distinction between these three effects, a series of topographic analyses were conducted. Analyses of this kind are usually motivated to determine whether the scalp distribution of particular effects differ, in line with the claim that qualitatively distinct patterns of scalp topography can be used to infer the engagement of at least partially non-overlapping brain regions and thus functionally distinct cognitive processes (Wilding, 2006; Rugg & Coles, 1995). The current analyses were conducted on the difference scores for each of the effects which were then rescaled using the max–min method of normalization to remove amplitude differences of the effects (McCarthy & Wood, 1985), and always comprised factors of Subtraction Contrast (two levels) and Electrode (all 58 levels were included). A significant interaction between these factors for amplitude normalized data is indicative of distinct scalp topographies and in line with this only the presence or absence of such interactions are reported here. The high number of levels precluded the use of a MANOVA, so all analyses were repeated measures ANOVAs with Greenhouse–Geisser corrections for sphericity, and corrected degrees of freedom are reported where necessary. For old/new subtractions, new ERPs were always subtracted from old, and for the new item subtractions, low frequency new was subtracted from high frequency new.

In a first analysis on the non-speeded data, high and low old/new effects in the early 350–550 time window were compared to determine whether the assumption that these effects differed in amplitude but not distribution was justified. There was no significant interaction between Subtraction Contrast (high old/new subtraction, low old/new subtraction) and Electrode (F(3.28,55.68)=1.16, p=.337). This same contrast in the subsequent 550–750 ms time window was significant, however, (F(4.27,72.62)=5.23, p=.001, η_p^2=.235), in line with the differential engagement of late retrieval processes for high and low frequency items in this response deadline condition. To determine whether the early frontal old/new effect differed from the posterior new item effect in the same time window (350–550 ms), paired comparisons between high old/new vs. new item difference and low old/new vs. new item difference revealed significant Subtraction Contrast by Electrode interactions (high old/new vs. new: F(4.19,71.29)=3.61, p=.009, η_p^2=.175, low old/new vs. new: F(4.22,71.80)=6.33, $p<.001$, η_p^2=.271). These effects remained significant when the new item difference from the entire 300–600 ms window was tested (high old/new vs. new: F(4.10,69.73)= 3.44, p=.012, η_p^2=.168, low old/new vs. new: F(3.98,67.72)=5.92, $p<.001$, η_p^2=.258). The final comparisons focused on the scalp distributions of old/new effects across adjacent time windows (350–550 ms, 550–750 ms), in order to determine whether the current data indicate the engagement of distinct functional processes over time. Whereas the early vs. late effect contrast elicited a marginally significant interaction with Electrode location for low frequency items (F(2.89,49.15)=2.44, p=.078, η_p^2=.125), this was not the case for high frequency items (F(4.62,78.46)=1.25, p=.296, η_p^2=.068).

In the speeded response condition, both the early and late effects showed comparable amplitudes for high and low frequency and there was no clear evidence that old/new effects were influenced by frequency. Although this outcome differs from the non-speeded condition, the pattern of topographic analyses in the early time window was nonetheless comparable for the two conditions. There was no evidence of a difference in distribution for the high and low early old/new effects (F(3.64,61.80)=.50, p=.718), whereas both these early effects differed from the new item effect when this was taken for the 350–550 ms (high old/new vs. new: F(3.83,65.15)=3.07, p=.024, η_p^2=.153; low old/new vs. new: F(3.79,64.35)=5.75, p=.001, η_p^2=.253) and 300–600 ms (high old/new vs. new: F(3.85,65.50)= 3.20, p=.020, η_p^2=.158; low old/new vs. new: F(3.79,64.51)=6.07, $p<.001$, η_p^2=.263) time windows. In the late time window, however, there was no difference between the distribution of the high and low old/new effects (F(2.96,50.35)=1.11, p=.354) in line with the engagement of comparable processes for high and low frequency items in this response deadline condition. In a pattern more comparable with the non-speeded condition, the early and late old/new effects differed over time for low frequency (F(4.57,77.69)=2.99, p=.019, η_p^2=.150) but not for high frequency items (F(4.03,68.58)=.79, p=.537).

4. General discussion

Based upon previous work and theoretical consideration of the word frequency mirror effect, the ERP data were expected to reveal two distinct types of familiarity, namely absolute and relative familiarity. In line with this prediction, the operationally defined contrasts shown in Fig. 1 revealed temporally overlapping but qualitatively distinct early ERP effects. This pattern could be most clearly observed in the non-speeded condition: an early old/new effect with a midfrontal maximum was significantly larger for low than high frequency items and thus behaved in a manner concordant with an index of relative familiarity. This effect was topographically distinct from the difference between high and low frequency new items, assumed to index absolute familiarity, which was posteriorly distributed between 300 and 600 ms. This pattern provides evidence for the existence of functionally distinct early familiarity signals. Before discussing the functional interpretations of this finding, we turn to the impact of constraining the time window in which participants could make old/new decisions on both performance and the critical ERP effects.

4.1. Speeded response demands influence ERP old/new effects

Both the current ERP and behavioural data differed robustly across the two response deadlines. The standard mirror effect was observed under non-speeded conditions whereas the greater hit rate for low frequency items was selectively removed under response deadline conditions. The selective elimination of the hit advantage whilst leaving the false alarm difference intact, replicates previous reports which have employed a response deadline manipulation (Balota et al., 2002) and is in line with the assumption that the hit advantage for low frequency words depends upon a later-occurring recognition process than that responsible for the false alarm rate. The assumption made throughout is that this later process is recollection and that this cannot be employed under speeded response conditions. The ERP index of the engagement of recollective processing data in the current data, the left parietal old/new effect, cannot be straightforwardly incorporated into this account, however, for two reasons. Whereas the effect was only present for low frequency items in the non-speeded condition, as indicated by the different topographic distributions for high and low frequency old/new effects in the late time window, there was no evidence that the amplitude of this effect was reduced when a response deadline was imposed. Moreover, the effect was equivalent in amplitude for high and low frequency items under speeded conditions, in line with the comparable topographies in the 550–750 ms time window for this response deadline. This pattern is most striking if one considers that the left parietal old/new effect was in fact larger for high frequency items in the speeded than the non-speeded condition.

It is not possible to determine on the basis of old/new differences alone whether the process indexed by this effect actually contributed to the correct speeded judgments and it may be the case that participants experienced recollection but were unable to utilize the information it provided before they responded. One way to ascertain whether the process indexed by this effect actually contributed to old/new decisions is to analyse old items in the time window of interest separated according to whether they were given an old (hit) or new (miss) response. If the effect is larger for hits than misses this would imply that the process of interest contributes to making an old response. It was possible to test this post-hoc, using data from 11 participants who had a sufficient number of artifact-free trials ($n \geq 13$; see Addante, Ranganath, & Yonelinas, 2012; Gruber & Otten, 2010, for examples of reports in which a similar minimum trial number criterion has been employed) for misses in the speeded condition. A MANOVA with factors of Frequency (high, low), Item-Type (hit, miss) and AP (frontal, central, posterior) in the 550–750 ms revealed a main effect of Item-Type ($F(1,10)=17.10$, $p=.002$, $\eta_p^2=.63$) only, in line with the greater ERP amplitude for hits than misses in this time window. The larger ERP effects for hits than misses suggests that the current deadline of 750 ms was not sufficient to remove recollection for both item types in the speeded condition. This particular deadline was chosen to correspond with the principal 'fast' deadline of 800 ms implemented by Coane et al. (2011), but it may be that much shorter response deadlines of around 500 ms (cf. Balota et al., 2002) are necessary to ensure that recollection is entirely excluded. A further related point pertains to the possibility that the 750 ms deadline may have accentuated participants' focus on recognition processes occurring at this point, which might go towards explaining the increase in amplitude of the left parietal old/new effect for high frequency items from the non-speeded to speeded condition, causing it to become equivalent in amplitude to that for low frequency items.

Although the miss analysis clearly indicates that the speeded response condition did not elicit a reduction in the amplitude of the recollection effect, it nonetheless had a significant impact on the ERP old/new effects more generally, most obviously by removing any interactions between ERP old/new amplitude difference and frequency. Thus, the larger early and late old/new effects for low frequency words that were observed in the non-speeded condition were replaced by old/new effects of comparable amplitudes for high and low frequency items. Here we consider the possibility that the engagement of equivalent recognition processes for high and low frequency words, rather than the selective reduction of recollection for low frequency items, is the reason for the absence of the hit difference in the speeded response condition. An ensuing question is therefore: why would recognition processing of high and low frequency words be equivalent in the speeded condition? One possibility is that the initial non-speeded recognition test provided participants with sufficient insight into the phenomenological differences between high and low frequency items that they were able to some extent to overcome these differences, perhaps by minimizing encoding differences in the second study phase. Reder, Paynter, Diana, Ngiam, and Dickison (2007) recently updated their computational model of the mirror effect to incorporate the influence of attentional differences at encoding by including the additional assumption that working memory resources for encoding are limited and that low frequency items demand a greater proportion of these resources by virtue of their low absolute familiarity. It is conceivable that participants may have been able to reallocate these resources more equitably to high and low frequency items in the second study phase and that this may have reduced overall differences at retrieval. Irrespective of the exact cause of the unexpected pattern in the speeded task, by ensuring all participants initially completed the task with standard response instructions, the data from the non-speeded response condition can be considered representative of the way in which recognition processes differ for high and low frequency words under standard response requirements. These non-speeded ERP data show a striking correspondence with predictions derived from dual-process models, although it does also appear to be the case that particular task demands – in this case, reducing response time and/or familiarizing participants with a high/low recognition task – can overwrite this pattern and eliminate functional differences in the way in which high and low frequency items are recognized.

The considerations outlined above highlight the difficulty with which behavioural evidence for relative familiarity can be extracted. Another factor which may contribute to this is the possibility that lexical decision times are longer for low than high frequency words (e.g. Scarborough, Cortese, & Scarborough, 1977). The greater likelihood of a timeout for low than high frequency time-outs in the current speeded condition does indeed indicate that lexical processing was generally slower for low frequency words in the current experiment. This might suggest that, although relative familiarity is on average greater for these items, slower lexical processing might offset this relative benefit under speeded response conditions, leaving performance equivalent for high and low frequency items. It may be that more complex designs such as that derived from the two-list exclusion task employed by Coane et al. (2011) are the best way to provide evidence of the influence this process has on actual responding to high and low frequency items. The absence of robust behavioural measures of relative familiarity does not necessarily undermine the fact that an early old/new effect associated with familiarity behaved in a way consistent with a relative familiarity mechanism under standard mirror effect conditions. The current ERP data, however, do serve as a necessary cautionary reminder when making functional interpretations about the contribution of recollection across response deadline conditions based upon behavioural data alone.

4.2. The relationship between absolute and relative familiarity

Two ERP effects with distinct functional characterizations were observed in overlapping early time windows in the current experiment. As outlined and reasoned below, we favour interpretations which posit these effects to be distinct kinds of familiarity. This is not the first finding in which a functionally distinct posteriorly distributed ERP effect has been observed in the same time window as the early midfrontal old/new effect, however, and it is necessary to consider the extent to which the current posterior effect corresponds with those reported previously. In several reports, posterior recognition-related ERP effects have been related to implicit recognition processing (e.g. Rugg, Mark, Walla, Schloerscheidt, Birch, & Allan, 1998; Voss & Paller, 2009). For example, Rugg et al. observed a posterior positivity from 300 to 500 ms for ERPs to misses relative to those elicited by correctly rejected items. Conversely, Voss and Paller (2009; see also Ryals, Yadon, Nami, & Cleary, 2011) report a slightly earlier (around 300 ms) relative negativity over posterior sites for correct guesses. These findings comprise correlates for at least two distinct processes associated with previous presentation which are not consciously available. Critically, it is unlikely that the early posterior effect observed in the current paradigm corresponds with either kind of implicit process because it is observed when contrasting new items for which episodic implicit memory should be comparable. Another possible characterization of implicit memory which would be more likely to systematically vary between high and low frequency new items, however, is implicit conceptual fluency. The current contrasts cannot unequivocally inform whether the processing indexed by this contrast is consciously available. It is nonetheless interesting to note the correspondence between the amplitude of this effect and the false alarm rate, which might indicate that the strength of this signal relates to the likelihood that participants respond old to a new item (see end of this section for more on this argument). Reasoning of this kind would suggest that the processing reflected by this signal is explicitly available to participants.

The polarity and spatial distribution of the effect elicited by the absolute familiarity contrast is consistent with what would be expected based upon previous work on the N400 index of semantic access for words presented in isolation (Kutas & Federmeier, 2000). One possibility is that the processes which give rise to the absolute familiarity effect are based on structures and operations which strongly overlap with those which give rise to the N400 index of semantic accessibility. Fundamental to most accounts of the frequency effect is the assumption that high frequency words, either because of a higher baseline level of activation and/or because they are associated with a greater variety of contexts, are on average more accessible than low frequency words (Cook, Marsh, & Hicks, 2006; Reder et al., 2007). This would then be consistent with the current data, insofar as the N400 is always less pronounced for more semantically accessible items and posterior ERPs to high frequency correct rejections in the relevant time window were significantly less negative than those to low frequency correct rejections. The posterior distribution of this effect also corresponds with functional interpretations put forward in reports in which an early old/new effect with a posterior distribution was observed (Bader et al., 2010; Mackenzie & Donaldson, 2007). In both these reports, it was considered feasible that the absence of pre-experimental absolute familiarity engendered by the particular stimuli employed might enable this familiarity mechanism to be more diagnostic of previous presentation than relative familiarity. Another way of understanding this might be to incorporate the power function with which Reder et al. (2007) have modelled absolute familiarity, which ensures that first exposures to an item contribute much more to the accumulation of absolute familiarity than later exposures. This would mean that for items which are presented for the very first time during the study phase of an experiment, absolute familiarity may respond in a way which is relatively diagnostic of previous occurrence in a simple study-test recognition task, but for previously encountered items (even those that are encountered relatively infrequently) this same mechanism would not be able to sufficiently discriminate recent (study phase) from general occurrences. In other words, absolute familiarity could only provide a diagnostic signal of previous occurrence for such items on their first (or perhaps second) ever repetition. A signal of this kind would be useful in tasks such as those used by Mackenzie and Donaldson (2007) and Bader et al. (2010) in which all stimuli are novel during learning, but it should not be diagnostic in the current task where all items are previously known (albeit with varying levels of familiarity). An additional relative familiarity mechanism which is more sensitive to the recent occurrence of previously known items would be required for these items. We take this to be the role of the process measured by the current early midfrontal old/new effect.

There is ongoing debate concerning the most appropriate functional interpretation of the midfrontal old/new effect (see Rugg & Curran, 2007; Voss et al., 2012). The current data indicate that this midfrontal effect is distinct from an automatic semantic mechanism (albeit, likely dependent upon it; see below) and is consistent with the notion that it reflects familiarity driven by the recent presentation of a previously known item. In previous reports (Bridger et al., 2012; Mecklinger, Frings, & Rosburg, 2012, see also Rugg & Curran, 2007), we have favored an account which assumes that the midfrontal effect – and the process it reflects – is multiply determined. This could, for example, be as a consequence of differentially combining perceptual fluency and/or conceptual information (e.g. Leynes & Zish, 2012) or changes in the reliance of different kinds of information depending upon task demands (Ecker, Zimmer, & Groh-Bordin, 2007; Küper, Groh-Bordin, Zimmer, & Ecker, 2012; Rosburg, Johansson, & Mecklinger, 2013). Considered in light of the current data, this idea perhaps best resonates with Mandler's original model of relative familiarity which posits at least two components: absolute familiarity and an increase in 'integration of an item' following recent presentation (termed constant d). Given that this model predicates that relative familiarity should increase as the value of absolute familiarity (F) decreases (assuming that d remains constant: see formula in Section 1), a worthwhile enterprise might be to test such a model by determining, either across stimuli or participants, whether the relationship between absolute familiarity and d predicts relative familiarity estimates. Such an endeavour is, however, problematic for a number of reasons not least because of the difficulty of theoretically and practically operationalizing d. Another problem which likely precludes the possibility of using ERP signals to test such a model is the high degree of covariance between ERP signals (multicollinearity) which would confound any functional interpretations. The logic employed throughout here emphasizes the interdependence between the two mechanisms, however, and one aspect of the current data which might be perceived as consistent with the idea that absolute familiarity contributes to relative familiarity is the slightly staggered time course of the two effects, in which the absolute familiarity effect was observable 50 ms before the midfrontal effect was influenced by frequency in the non-speeded condition.

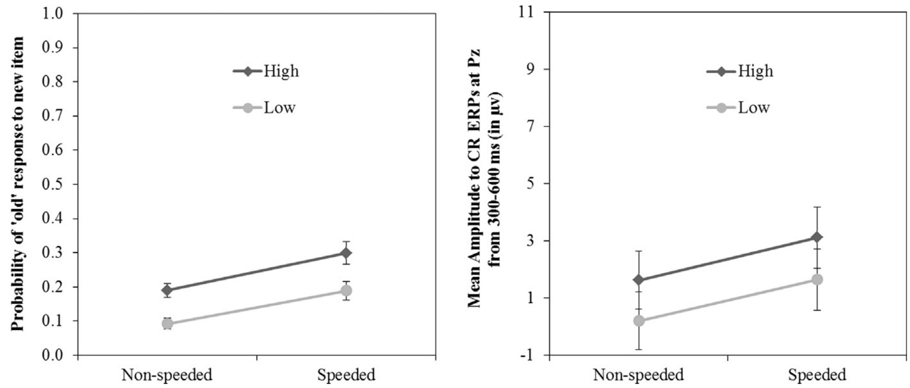

Fig. 5. Line plots showing the correspondence between false alarm rates and the ERP index of absolute familiarity, across frequency and response deadline condition. Left Panel: Mean false alarm rates. Right Panel: Mean amplitude of the ERP at electrode Pz between 300 and 600 ms post-stimulus. All error bars represent ± 1 standard error of the mean.

ERP evidence of a relative familiarity mechanism was only observed in the non-speeded condition, whereas the difference between new item ERPs was not influenced by response deadline and remained comparable in size (at least over posterior sites) in the speeded condition. This disconnect across response deadlines reveals a crucial qualification of the preceding conclusion: variation in absolute familiarity may not inevitably lead to changes in relative familiarity, and other factors, such as attentional allocation during encoding, are likely to contribute to the extent to which recollection and familiarity differ for high and low frequency items at test. The absolute familiarity effect was alone out of the current effects in that it was the only ERP effect which was not influenced by response deadline, an observation which corresponds with the fact that the greater false alarm rate for high than low frequency words remained constant across response deadlines (although the false alarm rate did increase generally in the speeded condition). Fig. 5 shows the strong correspondence between the pattern of false alarms and ERP amplitude over posterior sites in the early time window, across frequency and response deadlines. This pattern is compatible with Coane and colleagues' assumption that this absolute familiarity signal to new items is what drives the difference in false alarm rates between high and low frequency items, and suggests a task-invariant process. By the same token, the old/new effects also correspond with performance, insofar as high and low frequency old/new effects were comparable in amplitude when the hit rate difference between high and low frequency was removed (in the speeded condition) but were larger for low frequency items when there were more hits for these items (in the non-speeded condition).

4.3. Summary

In summary, the present ERP data provide convergent evidence for the presence of two topographically distinct familiarity signals with overlapping time windows. A slightly earlier onsetting posteriorly distributed effect was observed when new high and low frequency items were contrasted to reveal differences in absolute familiarity, whereas a midfrontally distributed old/new effect was significantly larger for low than high frequency items under standard recognition conditions, as would be commensurate with a relative familiarity index. Together the data provide convergent support for recent claims based upon behavioural data (Coane et al., 2011), that word frequency can modulate three recognition-related processes (absolute familiarity, relative familiarity and recollection). They also go some way towards a clearer characterization of the term familiarity by emphasizing the presence of two early familiarity processes with distinct but interwoven functional mechanisms.

Acknowledgements

We thank Marie Schwartz and Katharina Jung for assistance with stimuli preparation and data collection. The authors note that there are no conflicts of interest.

References

Addante, R. J., Ranganath, C., & Yonelinas, A. P. (2012). Examining ERP correlates of recognition memory: Evidence of accurate source recognition without recollection. *NeuroImage*, *62*(1), 439–450.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database [webcelex]*. Philadelphia, PA, USA: University of Pennsylvania, Linguistic Data Consortium.

Bader, R., Mecklinger, A., Hoppstädter, M., & Meyer, P. (2010). Recognition memory for one-trial-unitized word pairs: Evidence from event-related potentials. *NeuroImage*, *50*(2), 772–781.

Balota, D. A., Burgess, G. C., Cortese, M. J., & Adams, D. R. (2002). The word-frequency mirror effect in young, old, and early-stage Alzheimer's disease: Evidence for two processes in episodic recognition performance. *Journal of Memory and Language*, *46*(1), 199–226.

Bowles, B., Crupi, C., Mirsattari, S. M., Pigott, S. E., Parrent, A. G., Pruessner, J. C., et al. (2007). Impaired familiarity with preserved recollection after anterior temporal-lobe resection that spares the hippocampus. *Proceedings of the National Academy of Sciences*, *104*, 16382–16387.

Bridger, E. K., Bader, R., Kriukova, O., Unger, K., & Mecklinger, A. (2012). The FN400 is functionally distinct from the N400. *NeuroImage*, *63*, 1334–1342.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, *58*(5), 412–424.

Coane, J. H., Balota, D. A., Dolan, P. O., & Jacoby, L. L. (2011). Not all sources of familiarity are created equal: The case of word frequency and repetition in episodic recognition. *Memory & Cognition*, *39*(5), 791–805.

Cook, G. I., Marsh, R. L., & Hicks, J. L. (2006). The role of recollection and familiarity in the context variability mirror effect. *Memory & Cognition*, *34*, 240–250.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, *28*(6), 923–938.

DeCarlo, L. T. (2007). The mirror effect and mixture signal detection theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(1), 18–33.

Dien, J., & Santuzzi, A. M. (2005). Application of repeated measures ANOVA to high-density ERP datasets: A review and tutorial. In: T. C. Handy (Ed.), *Event-related potentials: A methods handbook* (pp. 1–73). Cambridge, MA: MIT Press.

Ecker, U. H., Zimmer, H. D., & Groh-Bordin, C. (2007). The influence of object and background color manipulations on the electrophysiological indices of recognition memory. *Brain Research*, *1185*, 221–230.

Glanzer, M., & Adams, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(1), 5–16.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*, 468–484.

Gruber, M. J., & Otten, L. J. (2010). Voluntary control over prestimulus activity related to encoding. *The Journal of Neuroscience*, *30*(29), 9793–9800.

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., et al. (2011). dlexDB – eine lexikalische Datenbank für die psychologische und linguistische Forschung [dlexDB – a lexical database for psychological and linguistic research]. *Psychologische Rundschau*, *62*, 10–20.

Hintzman, D. L. (1994). On explaining the mirror effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(1), 201–205.

Hintzman, D. L., & Curran, T. (1994). Retrieval dynamics of Recognition and Frequency Judgments: Evidence for Separate Processes of Familiarity and Recall. *Journal of Memory and Language*, *33*, 1–18.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513–541.

Jäger, T., Mecklinger, A., & Kipp, K. H. (2006). Intra- and inter-item associations doubly dissociate the electrophysiological correlates of familiarity and recollection. *Neuron*, *52*(3), 535–545.

Jasper, H. H. (1958). Report of the Committee on Methods of Clinical Examination in Electroencephalography. *Electroencephalography and Clinical Neurophysiology*, *10*, 371–375.

Joordens, S., & Hockley, W. E. (2000). Recollection and familiarity through the looking glass: When old does not mirror new. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1534–1555.

Kriukova, O., Bridger, E.K, & Mecklinger, A. (2013). Semantic relations differentially impact associative recognition memory: Electrophysiological evidence. *Brain and Cognition*, *83*, 93–103.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science*, *4*, 463–470.

Küper, K., Groh-Bordin, C., Zimmer, H. D., & Ecker, U. K. H. (2012). Electrophysiological correlates of exemplar-specific processes in implicit and explicit memory. *Cognitive, Affective* & *Behavioral Neuroscience*, *12*(1), 52–64.

Leynes, A. P., & Zish, K. (2012). Event-related potential (ERP) evidence for fluency-based recognition memory. *Neuropsychologia*, *50*(14), 3240–3249.

Mackenzie, G., & Donaldson, D. I. (2007). Dissociating recollection from familiarity: Electrophysiological evidence that familiarity for faces is associated with a posterior old/new effect. *NeuroImage*, *36*, 454–463.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*, 252–271.

McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology*, *62*, 203–208.

Mecklinger, A. (2006). Electrophysiological measures of familiarity memory. *Clinical EEG and Neuroscience*, *37*(4), 292–299.

Mecklinger, A., Brunnemann, N., & Kipp, K. (2011). Two processes for recognition memory in children of early school age: An event-related potential study. *Journal of Cognitive Neuroscience*, *23*(2), 435–446.

Mecklinger, A., Frings, C., & Rosburg, T. (2012). Response to Paller et al.: The role of familiarity in making inferences about unknown quantities. *Trends in Cognitive Sciences*, *16*(6), 315–316.

Mollison, M.V, & Curran, T. (2012). Familiarity in source memory. *Neuropsychologia*, *50*(11), 2546–2565.

Murdock, B. B. (1998). The mirror effect and attention-likelihood theory: A reflective analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(2), 524–534.

Nessler, D., Mecklinger, A., & Penney, T. B. (2001). Event related brain potentials and illusory memories: The effects of differential encoding. *Cognitive Brain Research*, *10*, 283–301.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember-know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 294–320.

Reder, L. M., Paynter, C., Diana, R. A., Ngiam, J., & Dickison, D. (2007). Experience is a double-edged sword: A computational model of the encoding/retrieval trade-off with familiarity. In: B. Ross, & A. Benjamin (Eds.), *The psychology of learning and motivation*, Vol. 48 (pp. 271–312). San Diego: Academic.

Rosburg, T., Johansson, M., & Mecklinger, A. (2013). Strategic retrieval and retrieval orientation in reality monitoring studied by event-related potentials (ERPs). *Neuropsychologia*, *51*(3), 557–571.

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory* & *Cognition*, *18*(4), 367–379.

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

Rugg, M. D., Cox, C. J. C., Doyle, M. C., & Wells, T. (1995). Event-related potentials and the recollection of low and high frequency words. *Neuropsychologia*, *33*, 471–484.

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, *11*, 251–257.

Rugg, M. D., & Doyle, M. C. (1992). Event-related potentials and recognition memory for low- and high-frequency words. *Journal of Cognitive Neuroscience*, *4*(1), 69–79.

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, *392*, 595–598.

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and Repetition Effects in Lexical Memory. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(1), 1–17.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*, 34–50.

Stenberg, G., Hellman, J., Johansson, M., & Rosén, I. (2009). Familiarity or conceptual priming: Event-related potentials in name recognition. *Journal of Cognitive Neuroscience*, *21*(3), 447–460.

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory* & *Cognition*, *18*(4), 380–393.

Vilberg, K. L., & Rugg, M. D. (2009). Functional significance of retrieval-related activity in lateral parietal cortex: Evidence from fMRI and ERPs. *Human Brain Mapping*, *30*(5), 1490–1501.

Voss, J. L., & Paller, K. A. (2009). An electrophysiological signature of unconscious recognition memory. *Nature Neuroscience*, *12*(3), 349–355.

Voss, J. L., Lucas, H. D., & Paller, K. A. (2012). More than a feeling?: Pervasive influences of memory without awareness of retrieval. *Cognitive Neuroscience*, *3* (3–4), 193–226.

Wiegand, I., Bader, R., & Mecklinger, A. (2010). Multiple ways to the prior occurrence of an event: An electrophysiological dissociation of experimental and conceptually driven familiarity in recognition memory. *Brain Research*, *1360*, 106–118.

Wilding, E. L. (2006). The practice of rescaling scalp-recorded event-related potentials. *Biological Psychology*, *72*, 325–332.

Wilding, E. L., & Herron, J. E. (2006). Electrophysiological measures of episodic memory control and memory retrieval. *Clinical EEG and Neuroscience*, *37*(4), 315–321.

Wilding, E. L., & Rugg, M. D. (1996). An event-related potential study of recognition memory with and without retrieval of source. *Neuropsychologia*, *37*, 441–454.

Woodruff, C. C., Hayama, H. R., & Rugg, M. D. (2006). Electrophysiological dissociation of the neural correlates of recollection and familiarity. *Brain Research*, *1100*, 125–135.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441–517.

Yonelinas, A. P., Aly, M., Wang, W.-C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, *20*(11), 1178–1194.

Young, M. P., & Rugg, M. D. (1992). Word frequency and multiple repetition as determinants of the modulation of event-related potentials in a semantic classification task. *Psychophysiology*, *29*(6), 664–676.

Yu, S. S., & Rugg, M. D. (2010). Dissociation of the electrophysiological correlates of familiarity strength and item repetition. *Brain Research*, *1320*, 74–84.