Frequency and regularity effects in reading are task dependent: Evidence from ERPs

Simon Fischer-Baum[1,*], **Danielle S. Dickson**[2], and **Kara D. Federmeier**[2,3,4]

[1]Department of Psychology, Rice University

[2]Department of Psychology, University of Illinois, Urbana-Champaign

[3]Program in Neuroscience, University of Illinois, Urbana-Champaign

[4]Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign

Abstract

Many theories of visual word processing assume obligatory semantic access and phonological recoding whenever a written word is encountered. However, the relative importance of different reading processes depends on task. The current study uses event related potentials (ERPs) to investigate whether – and, if so, when and how – effects of task modulate how visually-presented words are processed. Participants were presented written words in the context of two tasks, delayed reading aloud and proper name detection. Stimuli varied factorially on lexical frequency and on spelling-to-sound regularity, while controlling for other lexical variables. Effects of both lexical frequency and regularity were modulated by task. Lexical frequency modulated N400 amplitude, but only in the reading aloud task, whereas spelling-to-sound regularity interacted with frequency to modulate the LPC, again only in the reading aloud task. Taken together, these results demonstrate that task demands affect how meaning and sound are generated from written words.

Keywords

ERPs; Lexical Frequency; Spelling-to-sound regularity; visual word processing; task-dependent processing

---

The past several decades have seen intense back-and-forth between competing theories of reading (e.g. Dual Route, Coltheart et al., 2001; Parallel Distributed Processing, Plaut et al., 1996). Although these theories disagree on the mechanisms that underlie written word processing, they agree on several basic assumptions. First, most theories agree that visual words are linked to both sound- and meaning-based representations. In experienced readers, meaning access can happen directly from orthography or be mediated by the activation of sound-based representations, which, in turn, are linked to meaning. Similarly, the sounds corresponding to the written word can be accessed either directly from orthography or mediated by print-to-meaning and meaning-to-sound connections. Second, most theories

---

[*]Corresponding author, Department of Psychology, MS-25, Rice University, P.O. Box 1891, Houston, TX 77251, Tel: 713-348-3841, simon.j.fischer-baum@rice.edu.

assume that both of these cognitive operations are obligatorily employed in parallel; whenever a written word is encountered, literate adults both search for the meaning of that word and generate its pronunciation, regardless of the goals of the task. The objective of the current paper is to address this second assumption: To what extent are the same cognitive operations deployed by the reading system regardless of the task context in which words are presented? It is certainly the case that the *import* of sound- and meaning-based representations differs as a function of task-related goals. However, it is unclear whether differences in task demands actually change the nature of the processes brought to bear during reading.

The two extreme views – either that reading processes are identical regardless of the task-context, or that reading processes are entirely dependent on the demands of the task – are easy to rule out. The Stroop effect, in which participants are slower to identify the ink color of an incongruous color word, demonstrates that meaning is generated from a written word even if doing so is disadvantageous (see MacCleod, 1991 for review). Similarly, the pseudohomophone effect, in which participants are slower and more errorful to reject a nonword that would be pronounced as a word (e.g. BRANE) in a lexical decision task, suggests that pronunciations are generated from written strings the even when doing so comes at a cost (e.g. Davelaar, Coltheart, Besner & Jonasson, 1978). Taken together, these results indicate that some amount of meaning and sound processing happens whenever a written word is presented, even if it is not required for the task.

It is also clear that reading processes are not fully identical across tasks. Reading tasks differ in the amount of attention given to written words – for example, the difference between skimming a text and reading it carefully. Some tasks, such as judging the font size, require participants to attend more to superficial visual aspects of word, whereas other tasks require deeper processing of either the meaning of the word (e.g., judging the semantic category of the word) or its sound (e.g., judging whether the word rhymes with a target). Several studies have used event-related brain potential (ERP) measures to assess the impact of tasks designed to direct attention to different aspects of stimuli on the electrical activity elicited have found enhancements (as indexed by component amplitude and measures of topography) of ERP components associated with the attended subprocesses. For example, N400 responses, which (as discussed below) are associated with lexico-semantic processing, were larger when the task/cues directed attention to semantics than when they directed attention to visual features. Thus, it is apparent that task demands can lead to different reading strategies, with consequences for the extent to which various reading subprocesses are engaged.

Between the two extremes of reading processes being completely determined by the task or completely insensitive to task, however, there are a number of alternative hypotheses for how task demands affect reading processes. The ERP results described above suggest that selective attention, as modulated by task demands, can alter the degree to which various subprocesses are brought on-line during the processing of visual word. A more controversial view is that task-context not only affects the relative weight of these subprocesses, but also the fundamental nature of the computation – *how* information about meaning or sound is extracted from a written word. One way to test this hypothesis is to look at whether task-

context alters the effects of core lexical variables. Specifically, in this paper, we look at task-related modulation of word frequency and spelling-to-sound regularity effects. As we discuss below, these different variables are thought to index the computations that underlie different reading subprocesses. If the effects of these variables are altered by task, then task demands alter not just *how much* certain linguistic features are processed but actually *how* processing unfolds.

A number of studies have already demonstrated that behavioral outcomes related to these core lexical variables depend, at least in part, on task context. Seidenberg, Waters, Barnes and Tanenhaus (1984) compared reaction times to the same set of words in both reading aloud and lexical decision tasks. In reading aloud, main effects of both frequency and regularity were observed, with faster reading times for both high frequency and regular words, and an interaction between the two lexical properties, with low-frequency, irregular words being especially slow and error prone. In contrast, in lexical decision, there were clear effects of frequency but no effects of regularity or frequency-by-regularity interactions. When stimuli are appropriately controlled, the presence of a spelling-to-sound regularity effects on reaction time and error rate depends on task. Balota and Chumbley (1984, 1990) compared the magnitude of frequency effects for the same set of words in different word processing tasks – lexical decision, reading aloud and a semantic categorization task in which participants verified whether a word (e.g. ROBIN) was a member of a restrictive semantic category (e.g. BIRD). Frequency effects were greatest in the context of lexical decision, smaller in reading aloud, and were either quite limited or non-existent in their semantic categorization task. The role that frequency plays in driving reaction times during word processing clearly depends on task.

These behavioral results are consistent with the hypothesis that task demands can alter how reading subprocesses unfold. But because response times and accuracy measures necessarily summate across both the initial access to meaning and sound information as well as later decision and response-related processing, an alternative explanation, that these differences in the effects of lexical variables arise, not because of differences in the basic processes involved in sound and meaning extraction, but later, when decisions are being made, cannot be ruled out. Indeed, this is how Seidenberg et al. (1984) interpret their task-related response time differences. The same meaning and sound representations are generated from a written word in lexical decision and reading aloud, but the reading aloud task requires participants' responses to make use of the sound information, while a response in the lexical decision task can be performed without reference to the word's pronunciation. Although phonological information is activated in the lexical decision task, participants can ignore it when producing their responses. Similarly, Balota and Chumbley (1984, 1990) assume a common print-to-meaning extraction process (that is somewhat sensitive to word frequency) across the three visual word recognition tasks and account for the task-related differences in reaction time that they observe by assuming that there are also later, task-specific decision processes that vary in how they make use of frequency information. Behavioral data alone cannot distinguish between accounts that posit task-related differences in how reading related subprocesses unfold and those that posit differences at task-specific decision stages.

Instead, distinguishing between these accounts requires the ability to track processing over time – as can be done with event-related brain potential (ERP) measures. In this way, we can see whether changes in the impact of core lexical variables on processing manifest early in word processing. Therefore, in the current experiment, we recorded EEG as participants read the same set words in the context of two reading tasks. One, which has previously been used to study single word processing (e.g. Laszlo & Federmeier, 2007, 2011), is a semantic categorization task, wherein participants are instructed to press a button to instances of common, American first names. Previous work using this same task has demonstrated effects of other lexical and semantic variables, including orthographic neighborhood size and number of lexical associates, on ERP components associated with lexico-semantic processing (i.e., the N400, described below), indicating that this task induces attention to word meaning (e.g., Laszlo & Federmeier, 2007, 2011, 2014). The other is a delayed reading aloud task, which requires accessing word sound information. We examine whether effects of word frequency and spelling-to-sound regularity – which, as discussed next, have been well established as playing an important role in lexical processing – differ as a function of task, and, if so, when and how these differences emerge.

Lexical Frequency

Although effects of word frequency are common in visual word recognition tasks, there is still debate about what these effects tell us about how reading subprocesses operate. Frequency could directly affect perceptual processing of written words (e.g. Solomon & Postman, 1952), or it could influence the response bias for recognition processes (e.g. Morton, 1969). It could be an emergent property of serial lexical access (e.g. Forster, 1976) or an emergent property of the way that words are learned (e.g. Monsell, 1991). Across all of these accounts, however, word frequency is assumed to be an inherent part of how words are processed and thus a factor whose influence should not be sensitive to task-context. Instead, Norris (2006) offers a radically different view that starts from the assumption that the reader is trying to optimize performance in visual word processing tasks by performing the task as quickly as possible without making too many errors. Given these constraints, for many tasks, such as word identification or lexical decision, it is to the reader's benefit to be influenced by the prior probability that a specific word is going to be presented – i.e. that word's frequency. However, sensitivity to word frequency is not an intrinsic property of the processes involved in extracting meaning from print; the use of frequency information in meaning extraction depends on the goals of the task at hand. Such a view would be consistent with the results of Balota and Chumbley (1984, 1990) described above: since the demands of the tasks (lexical decision, reading aloud, certain semantic categorization tasks) differ, the role of word frequency in carrying out these tasks differs (although note that this is not Balota and Chumbley's interpretation of their data).

In the experiment described below, we look for effects of word frequency, not in reaction time, but in modulating different ERP components associated with word processing, in order to see whether task-context can affect how frequency influences lexico-semantic access and/or later, stimulus evaluation and decision making processes. Frequency effects have been relatively well-characterized in the ERP literature and affect many aspects of word processing. The frequency with which certain sub-word parts have been experienced – i.e.,

orthographic syllable frequency – modulate the P2 component, a visual sensory component that is also sensitive to attention (Barber, Vergara, & Carreiras, 2004). Frequency of whole words, rather than their perceptual features, has more often been linked to ERP effects beginning around 300 ms, especially in the form of amplitude modulations of the N400. The N400 is part of the normal evoked response to words and other meaningful stimuli. It has a very stable latency, peaking just before 400 ms in healthy young adults, but its amplitude is sensitive to factors that affect the ease of lexico-semantic access, with greater processing ease associated with smaller (more positive) N400s (reviewed in Kutas & Federmeier, 2011). Accordingly, visually presented words of higher frequency have been found to elicit smaller amplitude N400 responses than words of lower frequency, suggesting semantic access is easier for more frequently encountered items (Barber, Vergara, & Carreiras, 2004; Rugg, 1990; Van Petten & Kutas, 1990; Vergara-Martinez & Swaab, 2012).

After the N400, effects of frequency in ERPs have also been seen on positive-going components, such as the P300(b), elicited by processes related to stimulus evaluation and categorization, and, more generally, context updating (e.g., Kutas, McCarthy & Donchin, 1977; Donchin & Coles, 1988; see Polich, 2007 for review), as well as the more broadly defined late positive complex (LPC) that is often linked to reappraisal and conflict resolution (e.g., Van Petten et al., 1991; also see Van Petten & Luka, 2012 for a review). For example, Polich and Donchin (1988) examined word frequency effects in a lexical decision task and found that the P300 to low frequency words was delayed and reduced in amplitude, pointing to slower and more uncertain stimulus evaluation processes for categorizing low frequency items as words. Reports of LPC effects as a result of frequency manipulations have come from repetition paradigms wherein repeated low frequency words elicited long, slow positivities relative to their initial presentation, whereas no late differences were observed with high frequency words (e.g., Rugg, 1990; Rugg & Doyle, 1992).

N400 frequency effects have been reported under a variety of task conditions, including sentence reading (Van Petten & Kutas, 1990), lexical decision (Rugg, 1990; Barber, Vergara, & Carreiras, 2004), and category membership detection (Vergara-Martinez & Swaab, 2004), though the magnitude of these effects have not been directly compared across tasks. Thus, it remains unclear whether the influence of frequency on lexico-semantic access is task-sensitive. If, instead, the task-related differences in the effect of frequency on reaction time, like those reported by Balota and Chumbley (1984, 1990) arise only at later stages of processing related to the (re)appraisal of words in the context of the task demands, then we should see task-related modulations on the LPC rather than on the N400.

Spelling-to-Sound Regularity

A second core lexical variable we investigate in the experiment below is spelling-to-sound regularity. A word's spelling-to-sound regularity measures the mismatch between its actual pronunciation and that expected given typical correspondences between spelling and sound. As discussed above, reaction time effects of spelling-to-sound regularity are typically found in reading aloud tasks, but are attenuated or absent in lexical decisions tasks, particularly once other lexical variables (bigram frequency, number of neighbors) are taken into account. One possible explanation for this difference is that the computations involved in extracting

sound from the written word are different in reading aloud than they are in lexical decision. But an alternative explanation, which cannot be ruled out with behavioral data alone, is that these task-related effects reflect differences in what information is used to make decisions in the different tasks.

Regularity typically affects reading aloud in the form of frequency-by-regularity interactions: low-frequency irregular ("exception") words are read slower and with more errors than low-frequency regular words, whereas regularity does not affect reaction time (or error rate) for high frequency words (cf. Jared, 1997). Both dual-route theories of reading (e.g. Coltheart et al., 2001) and parallel Distributed Processing (PDP) theories of reading (e.g. Harm & Seidenberg, 2004) explain these frequency-by-regularity interactions, at least in part, by assuming a mismatch in the pronunciations generated by different reading subprocesses. Dual-route theories account for this frequency-by-regularity interaction by assuming that a sublexical processing route generates a plausible pronunciation of a word via knowledge about which sounds are likely to correspond to which letters (e.g. P tends to be read as /p/), whereas a lexical route generates the specific pronunciation associated with a written word (e.g. PINT is pronounced "paɪnt"). These pronunciations meet at a common level of phonological representation, at which point reconciliation between the inputs from the routes occurs. For exception words, the pronunciations generated by the lexical and the sublexical routes differ, leading to slower and more errorful processing of irregular words. The interaction between regularity and frequency is explained by assuming that frequency affects the speed of processing in the lexical routes. With high-frequency words, lexical route processing is fast, and articulation is initiated before a sublexical pronunciation is generated. With low-frequency words, the lexical route is slow and a sublexical pronunciation is generated before articulation can be initiated. Therefore, only low-frequency irregular words experience a conflict between lexical and sublexical pronunciations that requires resolution (e.g. Paap & Noel, 1991).

For PDP theories, print-to-sound mappings are learned over a range of grain-sizes, from single-letter to single-sound mappings (e.g. P→/p/) to whole written word to whole spoken word mappings (e.g. PINT→/paɪnt/). The nature of learning causes whole word mappings to be more common with higher frequency words (Plaut et al., 1996). Therefore, pronunciations for low frequency words are more likely to be generated from single-sound mappings, and thus are more likely to yield the wrong pronunciation. As the print-to-meaning-to-sound processing will result in the correct pronunciation regardless of regularity, low-frequency irregular words are again unique in that there will be a mismatch between the phonological representations generated by the print-to-sound pathway and those generated by the print-to-meaning-to-sound pathway (Harm & Seidenberg, 2004).

Although there has been considerable interest in accounting for frequency-by-regularity interactions, to our knowledge, no ERP study has systematically manipulated these factors, at least in the reading aloud task in which this interaction is typically observed.[1] There have been a number of studies that show that N400 amplitude can be modulated by phonological properties of the orthographic stimulus. Pseudohomophones elicit a reduced N400 amplitude relative to orthographically-matched pseudowords in lexical decision (Briesemeister et al., 2009) and elicit equivalently reduced N400 as their lexical homophones in predictable

sentence contexts, when compared to words that violate semantic expectations and pseudowords (Newman & Connolly, 2004). These results indicate that prior to semantic access some amount of phonological processing has occurred. However, regular and irregular words differ not on the ease of generating some pronunciation, but on the ease of generating the *correct* pronunciation. Given regular and irregular words carefully matched for properties such as orthographic neighborhood size, known to modulate N400 amplitude, we do not predict effects of spelling-to-sound regularity on the N400. Instead, exception words seem more likely to elicit downstream effects, possibly manifesting on the LPC, when the mismatch between alternative pronunciations is detected.

LPC effects of regularity should be observed in the reading aloud task, which requires participants to settle on a single phonological representation for each word form. Although the behavioral evidence suggests that an effect might only occur for low-frequency words, it is possible that the interaction in the end-state RT measure is actually due to a combination of effects on different cognitive processes – e.g., a frequency effect on the N400 combined with a regularity effect on the LPC. If, instead, the interaction is due to a cognitive process that is required uniquely for low-frequency exception words, then we might observe an LPC effect for this condition alone.

The effect of regularity would seem to play a much smaller role in our semantic categorization task. Even if readers do generate pronunciations for written words in this task, as suggested by the pseudohomophone findings, it is unnecessary to settle on a single pronunciation. Of course, it is possible that reader's brains could appreciate the difference between regular and exception words – and show an LPC effect – even though the phonological information is rendered less salient; such a finding would suggest that readers attempt to reconcile phonological representations of words irrespective of their task relevance.

In summary, building on behavioral results that have found task-dependent effects of frequency and regularity on reaction times, the current ERP study examines the timecourse and functional specificity of these effects. By analyzing the ERPs to words that vary in lexical frequency and spelling-to-sound regularity, we can determine when and during what type of process each variable has an effect. By comparing the ERPs to the same words in the context of different tasks, we can then determine whether the role that frequency and regularity play in processing written words depends on task demands.

---

[1]Sereno, Rayner and Posner (1998) did examine ERP effects of spelling-to-sound regularity in lexical decision, finding early effects in at least some of their participants. However, there are several reasons to be cautious in interpreting this result. First, the authors report a behavioral frequency by regularity interaction in a lexical decision task. When low-level orthographic statistics are controlled for (e.g. bigram frequency, orthographic neighborhood size), this interaction typically is not observed in lexical decision (Seidenberg et al., 1984). Sereno and colleagues do not report controlling for these variables. Therefore, it is possible that their early effects arose because of differences in low-level orthographic statistics (i.e., differences in bigram frequency) and not because of differences in spelling-to-sound regularity. Second, their effects were not robust in anaylses that included all of their participants. Instead, the effect was only found in the subset of participant who showed a strong behavioral effect of regularity. This approach may have artificially inflated the likelihood of finding the reported result

Method

Participants

Data were analyzed from twenty-four young adults recruited from the University of Illinois community and compensated with either course credit or cash. All were right-handed native monolingual English speakers (12 female, 12 male) with normal or corrected-to-normal vision, no early exposure to a second language, history of brain trauma, or current use of psychoactive medications. Mean age was 19 years old (Range: 18–26). Mean laterality quotient on the Edinburgh handedness inventory (Oldfield, 1971) was 0.85 (Range: 0.50–1.00, with 1.0 indicating strongly right-handed and −1.0 strongly left-handed). Eleven participants reported having a left-handed biological family member.

Materials

The critical stimuli consisted of 120 words, which varied in lexical frequency (divided into high and low frequency categories) and spelling-to-sound regularity (including both regular and irregular words). Table 1 gives the lexical properties of the four critical word classes. Word classes were matched for length, mean bigram frequency, and neighborhood size. Lexical frequency counts were taken from the Hyperspace Analogue to Language Corpus (Lund & Burgess, 1996) and bigram frequency and the different measures of orthographic neighborhood size were taken from the English Lexicon Project's database (Balota et al., 2007; Yarkoni, Balota & Yap, 2008). Spelling-to-sound regularity was calculated based on the predictability metric proposed by Berndt, Reggia and Mitchum (1987). In addition, there were 80 nonwords in the proper name detection experiment, 40 consonants strings and 40 pseudowords. These nonwords were matched to each other for bigram frequency, length and number of orthographic neighbors, but were not matched to the word stimuli on any of these dimensions. Because the words and nonwords are not matched on these critical variables, the nonwords are excluded from the analyses described below.

Procedure

Each participant completed two sequential sessions that were blocked for task, such that each participant participated in both the proper name detection and the delayed reading aloud task. The same set of critical words appeared in each task. Task order and response hand (for the detection task) were counterbalanced across participants. Analyses conducted with task order as a factor yielded no main effects or interactions, so task order was collapsed for the reported analyses.

Participants were seated 100cm from the computer monitor and were instructed to remain still and relaxed. Directly preceding critical trials, participants completed a short practice block to familiarize themselves with each task.

The proper name detection task followed procedures from recent studies by Laszlo and Federmeier (2007, 2011, 2014). During the proper name detection task, participants were instructed to respond with a button-press, as fast as possible without losing accuracy, to common American proper name targets (50 total). In addition to the critical words, filler stimuli for this task included pronounceable pseudowords and unpronounceable letter

strings. All items were presented in capital letters, in white Arial font on a black screen. Each appeared for 500ms with an interstimulus interval of 3000ms. This interval consisted of 1000ms of a blank screen immediately following item presentation, 1500ms of a white fixation cross, during which subjects were allowed to blink, and 500ms of a red fixation cross to serve as a warning that the next item was about to appear.

During the reading aloud task, participants were confronted with words and instructed to read them out loud after a delay; delayed naming was used to prevent recordings from being contaminated by articulatory movement artifacts. Speed of speech response was de-emphasized to place higher priority on accuracy, and the experimenter manually advanced to the next trial. Items were presented with the same visual characteristics (font and size) as in the name detection task, again for 500ms. Following presentation of the word, there was a blank screen for 1000ms, followed by a "?" to indicate that participants should begin articulation. Prior to the next item, a white cross was presented for 1000ms to indicate that blinking was permissible and then a red warning cross was presented for 500ms to alert the participant that the next stimulus was about to appear.

EEG Data Collection and Analysis

EEG was recorded using 26 silver/silver-chloride electrodes mounted in an elastic cap, with electrode impedances kept below 5kΩ. Caps were placed using standard fiduciary markers (nasion and inion for frontal-posterior distance, earlobes for determining midline), and the electrodes were distributed evenly over the scalp. The vertical electrooculogram (EOG) signal was monitored with an electrode on the infraorbital ridge and horizontal EOG was monitored with electrodes placed on the outer canthus of each eye. The data was referenced on-line to the left mastoid and re-referenced offline to an average of the right and left mastoids. A separate frontocentral electrode acted as ground. EEG signal was sampled at 250 Hz and subjected to an analog bandpass of 0.02 to 100Hz after online amplification by Sensorium amplifiers.

Raw waveforms were assessed for inclusion on a trial-by-trial basis with artifact thresholds separately calibrated by visual inspection for each subject. Trials with blink artifacts were either rejected (for 19 subjects) or corrected (for 5 subjects who had between 10–40% of their trials contaminated by blink artifacts, and thus had a sufficient quantity to create a corrective filter following Dale, 1994). Trials were additionally rejected after detection of other artifacts (horizontal eye gaze, movement, signal drift, amplifier blocking).

For the detection task, critical word trials in which subjects erroneously reported a name were also removed; this occurred at a rate of 2%. For the speaking task, trials were removed when participants failed to produce a correct response or if they began to articulate too early. Performance was marked online by an experimenter and was re-checked offline from recordings by a listener who was naïve to the hypotheses. On average, less than 2% of trials were lost due to early articulation or errors. The few speaking errors were primarily with low-frequency irregular words (74% of all errors), followed by high-frequency irregulars (15%), low-frequency regulars (9%) and high-frequency regulars (2%). Errors on regular items were primarily orthographically similar words ("grab" for GARB), whereas errors with irregular items were primarily regularizations ("pɪnt" for PINT). After removal of error

and artifact trials, the four critical bins contained an average of 28 out of 30 possible items (7.7% trial loss). Across all participants and conditions, the critical bins had at least 20 trials.

ERPs were created by extracting 256 data points (a software default) to create epochs spanning 100 ms prior to item onset until 920ms after item onset. The 100 ms prior to stimulus onset served as a baseline and was subtracted prior to averaging. ERP mean amplitudes were measured after application of a digital bandpass filter of 0.2 to 20Hz.

Results

Grand average waveforms at all electrodes to all four critical conditions in each task are shown in Figure 1. ERPs in all conditions manifested the morphology expected for visual word processing. Early components include, at posterior sites, a positivity peaking around 100 ms (P1), a negativity peaking around 170 ms (N1), and a positivity peaking around 240 ms (P2), and, at frontal sites, a negativity peaking around 100 ms (N1) and a positivity peaking around 200 ms (P2). Early components are followed in all conditions by a negativity, broadly distributed and largest over central electrode sites, peaking between 350 and 400 ms (N400), followed by a posterior positivity (LPC).

N400

N400 mean amplitudes were measured over all electrode channels in a time window of 300 to 450 ms (N400 peak latency, calculated by obtaining the average timepoint between 200 and 500 ms at which the waveforms across all channels reached a minimum value, was 365 ms). Mean amplitudes were assessed using a repeated measures omnibus ANOVA with two levels of task (detection and speaking), two levels of frequency (high and low), two levels of regularity (regular and irregular), and 26 levels of electrode site. Interactions with electrode site are not reported, as they were of no theoretical interest beyond ascertaining that the waveforms manifested the well-characterized distribution of the N400. There were no main effects of task ($F(1,23) = 0.10$; *ns*) or regularity ($F(1,23) = 2.69$; *ns*). There was a main effect of frequency ($F(1,23) = 15.56$; $p < 0.001$), with reduced N400 amplitude to high frequency words compared to low frequency words. This effect was moderated by an interaction between frequency and task ($F(1,23) = 7.31$; $p < 0.05$). No other interactions were significant: frequency and regularity, ($F(1,23) = 0.33$; *ns*), task and regularity ($F(1,23) = 0.05$; *ns*), or the three-way interaction ($F(1,23) = 0.72$; *ns*).

Follow-up analyses were conducted within each task, collapsing across regularity. In the speaking task, there was a significant effect of frequency ($F(1,23) = 34.67$, $p < 0.001$) such that high frequency items (mean = 1.12 µV, SE= 0.14) elicited smaller N400 amplitudes than did low frequency items (mean = 0.04 µV, SE = 0.14). However, the same analysis performed for data from the detection task revealed no effect of lexical frequency (low frequency: mean = 0.44 µV, SE = 0.15; high frequency: mean = 0.56 µV, SE = 0.13, $F(1,23) = 0.20$; *ns*). Direct comparisons across the tasks revealed no difference in the amplitude of the response to low frequency words ($F(1,23) = 1.45$; *ns*) and a marginal trend for reduced responses to high frequency words in the speaking compared to the name detection task ($F(1,23) = 3.36$; $p = 0.08$). Figure 2 shows the mean amplitude for high and low frequency items, averaged over all 26 channels, in the naming and name detection experiments.

We found no effect of experimental variables on N400 amplitude in the proper name detection task. One concern with these null findings might be that the task – detecting a proper name – does not rely on normal word recognition processes. One way to address this concern is to demonstrate that N400 amplitude is modulated by lexical level variables other than frequency and regularity. Previous studies using this same task have found that items with higher orthographic neighborhood (N) sizes elicit larger N400s than items with lower N do (Laszlo & Federmeier, 2007; Laszlo & Federmeier, 2011; Lazslo & Federmeier, 2014; cf, Vergara-Martinez & Swaab, 2012). While the stimuli used in this experiment were not designed to test effects of neighborhood size, the words used in this experiment naturally varied along this dimension. After performing a median split by N (greater than or less than or equal to five orthographic neighbors) on all words in our study, we measured N400 amplitude over all electrodes from 300 to 450 ms (as above) and replicated this well-documented finding: N400 amplitude was greater for high N items than for low N items ($F(1,23) = 4.25$; $p = 0.05$).[2] The fact that N400 amplitude was sensitive to number of orthographic neighbors in the proper name detection task makes it difficult to attribute the null effects of frequency and regularity to a failure of the task to engage with normal word recognition processes.

Late Positive Complex (LPC)

To capture the typical, posterior distribution of the LPC and to distinguish it from functionally separable frontal positivities sometimes observed in language tasks (e.g., Van Petten & Luka, 2012), the LPC was measured over the posterior 15 electrodes, using a late time window extracted from the last 300ms of our epoch (620–920 ms) to isolate it from residual N400 effects. Again, there were no main effects of task ($F(1,23) = 0.62$; *ns*) or regularity ($F(1,23) = 0.39$; *ns*). There was a main effect of frequency ($F(1,23) = 4.31$; $p < 0.05$), such that low frequency items elicited larger LPC responses than high frequency items. This effect was moderated by a three-way interaction of task by frequency by regularity ($F(1,23) = 4.37$; $p < 0.05$). None of the two-way interactions were significant: task and frequency ($F(1,23) = 0.54$; *ns*), task and regularity ($F(1,23) = 1.45$; *ns*), or frequency and regularity ($F(1,23) = 1.37$; *ns*).

Follow-up analyses were conducted within each task. In the reading aloud task, there were no main effects of frequency ($F(1,23) = 2.66$; *ns*) or regularity ($F(1,23) = 1.33$; *ns*), but the interaction was significant ($F(1,23) = 4.32$; $p < 0.05$). Pairwise examinations revealed no effects of regularity for high frequency words (regular: mean = 0.70 μV, SE = 0.16 vs. irregular: mean = 0.49 μV, SE = 0.14, $F(1,23) = 0.28$; *ns*), but a significant effect of regularity for low frequency words (regular: mean = 0.75 μV, SE = 0.14 vs. irregular: mean = 1.61 μV, SE = 0.16, $F(1,23) = 5.64$; $p < 0.05$). In summary, in the reading aloud task, low frequency irregular items elicited a late positive component that was larger than in all other conditions. In the detection task, there were no main effects (frequency, ($F(1,23) = 2.87$; *ns*); regularity ($F(1,22) = 0.29$; *ns*), and there was no interaction of frequency by regularity

---

[2]Because our stimuli were not designed to be split by N, there were length differences in the words that were high N (4 letters average) and low N (4.6 letters average). Such length differences affect early perceptual components, so we used a post-stimulus baseline to align waveforms in the earliest time-window from 0 to 100 ms in order to ensure that reported differences on the N400 were not contaminated by earlier effects of perceptual properties.

(F(1,22) = 0.41; *ns*). Figure 3 shows the difference in amplitude for four word types, averaged over the 15 posterior channels.

P300 in the detection task

Visual inspection of the waveforms in the detection task suggested the possibility of an effect of lexical frequency in the time period in between the N400 and LPC effects. An analysis between 450 and 600ms revealed a significant difference (F(1,23) = 14.18; $p <$ 0.01) between high frequency (mean = 2.00 μV, SE = 0.12) and low frequency words (mean = 1.25 μV, SE = 0.14). As can be seen in Figure 4, the timing, distribution, and morphology of this response is similar to the P300 response elicited by the targets (proper names) in this task. Indeed, an analysis of peak latency finds no difference (F(1,23) = 0.02; *ns*) between the latency of the small peak in response to high frequency words (514 ms, SE = 2.53) and the considerably larger amplitude (and expected) P300 response to target names (515 ms, SE = 2.35). This response seems to be due to participants experiencing minor confusions (in the absence of overt false alarms) between some high frequency words and the proper name targets.

General discussion

The results of this experiment clearly show that the impact of lexical frequency and spelling-to-sound regularity on the neural activity associated with visual word processing varies as a function of task demand – does so from fairly early stages of processing. Intention to name increased the influence of frequency and regularity on word processing compared to a proper name detection task. In particular, in delayed reading aloud, semantic access (as measured by the N400) was facilitated by high frequency compared to both low frequency words. Spelling-to-sound regularity did not affect N400 amplitudes in either task. However, in the delayed reading aloud task, a clear frequency by regularity interaction was observed on the late positive complex (LPC). Low frequency irregular words patterned differently than the other conditions, eliciting an enhanced LPC response. In contrast, the spelling-to-sound regularity and frequency of the same set of words did not modulate N400 or LPC responses in the proper name detection task. It seems likely that the task-related adjustments in neural processing that we document here are related – perhaps causally – to downstream changes in response speed patterns across tasks that have been documented previously for word frequency (Balota & Chumbley, 1984, 1990) and spelling-to-sound regularity (Seidenberg et al., 1984).

The current experiment is able to adjudicate between alternative explanations of these previous results. Although the behavioral results are consistent with the idea that task-context can modulate the role that core lexical variables play in the processes by which meaning and sound are extracted from written words (e.g. Norris, 2006), they can also be explained by assuming differences in task-specific decision processes (Balota & Chumbley, 1984, 1990; Seidenberg et al., 1984). Our results are consistent with the former account. As discussed in detail below, the differential effects of frequency on the N400 across task suggest that task demands can modulate how meaning is accessed from words, and the

differential effects of regularity and frequency on the LPC across tasks further suggests that some aspects of how pronunciations are derived from visual words are also task-sensitive.

N400, semantic access, and the effect of frequency but not regularity

A number of previous studies have examined the effects of lexical frequency on the N400 (e.g. Barber, Vergara, & Carreiras, 2004; Laszlo & Federmeier, 2011; Rugg, 1990; Van Petten & Kutas, 1990; Vergara-Martinez & Swaab, 2012). Several of these studies report clear frequency effects on N400 amplitude. In the context of a lexical decision task, Rugg (1990; see also Barber, Vergara & Carreiras, 2004) showed that higher frequency words presented in isolation evoked smaller N400s than lower frequency words. A similar effect of frequency was reported by Van Petten and Kutas (1990) in a sentence reading task, specifically to the sentence's first open class word.

The N400 frequency effect we observed in the reading aloud task is consistent with these previous studies: we observed reduced N400s to higher frequency words. Following a generally accepted interpretation of the N400 component, reduced N400 amplitude reflects easier semantic access (see, e.g., Kutas and Federmeier, 2011, for a review). For tasks like lexical decision, reading the first word of a sentence and reading aloud, then, semantic access is easier for higher frequency words.

However, frequency does not always reliably modulate N400 amplitude. Using a regression analysis, Laszlo and Federmeier (2011) found that, in the context of a proper name detection task, lexical frequency accounted for very little of the variance in N400 amplitude for single words – despite the fact that other variables, including orthographic neighborhood size and number of lexical associates, were strong predictors, making clear that this task does elicit lexico-semantic processing. The proper name detection task in the current study replicates this finding, using a much larger difference in frequency and controlling for other lexical variables, both in terms of the lack of a frequency effect and the finding that N400 amplitude was modulated by orthographic neighborhood size. Although it is possible that participants could succeed in the proper name detection with less attention than it takes to succeed in the reading aloud task, it is unlikely that our results could be explained simply by task-related differences in attention levels. Differences in attention predict smaller N400 amplitude across the board in the proper name detection task (e.g., Ruz & Nobre, 2008), whereas we found no main effect of task on N400 amplitude. Comparisons across conditions within each frequency level revealed no significant difference in the response to low frequency words and a trend for greater facilitation (i.e., smaller N400s) for high frequency words in the reading aloud task. Thus, it is the response pattern across words -- rather than the amount of activation overall -- that changes across the tasks. Similarly, Van Petten and Kutas (1990) found that, although frequency modulated the response to sentence-initial words, there was no effect of lexical frequency on the N400 amplitude for the final words in their sentence reading task. As these examples show, in some task situations, effects of lexical frequency on N400 amplitudes are notably reduced or even eliminated.

What properties of the task determine whether a robust frequency effect on the N400 is observed? Following Becker (1976), Van Petten and Kutas (1990) explain the lexical frequency by sentence position interaction observed in their sentence reading experiment by

relying on a general principle that semantic access is easier for words that are more expected. In the absence of any contextual information, the frequency with which a word has been encountered is a good predictor of how likely it will be seen again; i.e., without knowing anything else, previous experience tells that it is more likely that we will read DOG than PINT. However, contextual information provided by the other words in a sentence allows participants to make more informed predictions about what words are likely to appear next, overriding the predictive effects of lexical frequency. In a sentence like "Let's go to the bar and have a ___", PINT is more expected than DOG, and this difference should be reflected in the N400 amplitude. Consistent with this account, the amplitude of the N400 to the last word in a sentence has long been known to be highly correlated with contextual predictability, as measured by cloze probability (Kutas & Hillyard, 1984). In sentence reading, there is no context to predict the first word, and the ease of semantic access is a function of lexical frequency. By the end of the sentence, context has refined what words are predicted, and these effects of lexical frequency disappear.

Along the same lines, certain types of semantic categorization tasks provide different predictive constraints for single words than do tasks like lexical decision or reading aloud. The lexical decision task and the reading aloud task provide no context for what words participants expect to encounter, making lexical frequency the best predictor of what word will be observed next. However, in semantic categorization tasks, particularly those with small categories, participants know that they will see a number of words from the target category and will be faster and less error prone in their responses if they are prepared for those words when they arrive. Thus, participants' expectations are likely to be shaped more by the semantic constraints of the task and less by lexical frequency when they are making category decisions than when they are making lexical decisions or reading aloud. This prediction-based account of semantic-categorization can also explain the reaction time effects reported by Balota and Chumbley (1984, 1990). Greater effects of lexical frequency on reaction time were found in lexical decision and reading aloud when compared to their semantic categorization task, but greater effects of category prototypicality were found for their semantic categorization task – suggesting that participants expect words of a particular category and are thus sensitive to category structure (c.f., Federmeier & Kutas, 1999, for a similar account of patterns in sentence reading).

This approach to the role of word frequency in visual word processing aligns well with the assumptions of the Bayesian Reader account proposed by Norris (2006). Under this framework, readers are assumed to approximate optimal Bayesian descision-makers, and, consequently, influences of lexical variables like word frequency on word recognition depend on how useful that information is for carrying out the task at hand. In contrast, explanations of frequency effects that assume that sensitivity to frequency is an intrinsic part of the processes involved in gleaning information from words would have a difficult time accounting for our results. These include accounts of frequency effects that assume that high frequency words are always more perceptually salient (e.g., Solomon & Postman, 1952), are more likely to be identified by default (e.g. Morton, 1969), are uniformly searched for more quickly (e.g. Forster, 1976) or are inherently represented more strongly because of learning (e.g. Monsell, 1991).

This approach to frequency effects can also account for seemingly contradictory results from reaction time experiments. Balota & Chumbley (1984) used fairly small semantic categories in their semantic categorization tasks (e.g. BIRDS), and their finding – reduced or non-existent frequency effects in semantic categorization – has been replicated when other small categories are used (e.g. MONTHS; Forster, (2004). When a category is small, it is easy to prepare the system to recognize words from the target category, and the predictive role of frequency in word recognition disappears. However, reaction time effects of lexical frequency are observed in some semantic categorization tasks, particularly when the categories are much larger (e.g. ANIMALS, THINGS; Forster, 2004; Forster & Shen, 1996; Monsell, Doyle & Haggard, 1989). Within a large semantic category like animals, it is difficult to predict the set of target items that are likely to be observed and retaining some frequency-based expectation can still aid performance on the task (e.g., one might expect DOG more than AARDVARK).

This category size account is unlikely to explain the lack of a frequency effect in the proper name detection task; proper names is a pretty open category, probably more similar in size to the category ANIMALS then the category BIRDS. But the fact that category size can influence frequency demonstrates that the role of frequency in word recognition depends, at least in part, on task demands. Category size is just one factor that can determine the utility of frequency. For the proper name categorization task used in the present study, none of the proper names will be high frequency words. Because of this, frequency is not useful information for carrying out this task. Indeed, in the name detection task, we found evidence suggesting that frequency may actually be a distraction. Frequency effects were observed in the proper name detection test in the form of a (post-N400) P300 effect observed for high frequency words, which closely resembled the much larger correct response to targets. This pattern suggests that participants sometimes experienced the high frequency words as lures for the task (even though they were eventually able to correctly suppress their overt responses to these items), attesting to the utility of using non-frequency-based expectation strategies. Our data thus suggest that the effects of lexical frequency on semantic access, as indexed by the N400, depend on the utility of frequency, as opposed to other task- or context-based factors, for tuning expectations about what word is likely to be encountered next.

It is critical to note that this view does not imply that frequency effects will *never* be observed in the context of a categorization task. Rather, it predicts reduced effects of frequency in some tasks compared to others, evinced by the frequency by task interaction observed on N400 amplitude in our experiment. For example, Vergara-Martinez & Swaab, 2012) did find significant frequency effects on N400 amplitude in a semantic categorization task, wherein participants were instructed to press a button when an animal name was presented. The proposed account suggests that if the words used in Vergara-Martinez & Swaab, 2012) were also administered in a reading aloud task, larger N400 frequency effects would be observed. Moreover, there may also be differences in the influence of frequency on N400 amplitude depending on the type of categorization task. For example, for the same set of non-exemplars (items not responded to in a semantic categorization task), we predict differential effects of frequency on N400 amplitude, with small or non-existent frequency

effects with semantic categorization tasks that rely on small categories (e.g. BIRDS, MONTHS) and larger effects of frequency with semantic categorization tasks in which exemplars are drawn from larger categories (e.g. ANIMALS, THINGS); similarly, we predict differential effects of frequency on N400 amplitude, when the task requires detecting members of a category whose members are all infrequent (e.g. PROPER NAMES) compared to detecting members of a category that includes some high frequency words (e.g. ANIMALS). Our data support a growing move toward thinking of frequency as one of many sources of information that can be used to estimate the probability of upcoming words – which may be superseded when other, more task-relevant sources of probability, are available (e.g., Smith and Levy, 2013).

Although additional work is required to determine precisely which factors determine the size of the N400 frequency effect in a given task, our data show – for the first time – that task demands can significantly modulate frequency effects on the N400. The N400 has been argued to index the initial linking between a word form and meaning information stored in long-term memory (Kutas & Federmeier, 2011). N400 amplitude modulations are observed even under conditions of reduced awareness, such as with masking (Stenberg et al. 2000), during the attentional blink (Vogel et al. 1998), and during sleep (for auditory words; see review in Ibanez et al. 2008); thus, the N400 reflects aspects of processing well before those associated with the use of a decision criteria. Our N400 findings, therefore, clearly support the hypothesis that task demands can have early effects, modulating core aspects of word processing (e.g. Norris, 2006).

In contrast to the N400 effect of frequency and its interaction with task, we found no evidence for effects of spelling-to-sound regularity on the N400 in either task. As discussed in the introduction, there has been little previous ERP research that looks at how spelling-to-sound regularity modulates reading-related ERP components. The N400 has been shown to be sensitive to some manipulations of phonology (e.g., Briesemeister et al., 2009; Newman & Connolly, 2004), suggesting that phonological representations are (or at least can be) activated in some form prior to semantic access from visual words (see Grainger & Holcomb, 2009 for a detailed description of the time-course of single word processing as revealed by ERPs). However, spelling-to-sound regularity effects have been argued to arise when mismatching phonological information has been elicited by a particular word form. Effects in behavioral paradigms have been attributed to processes involved in resolving the resulting conflict. The present data suggest that these effects take place after semantic access, manifesting in components that follow the N400, such as the Late Positive Complex (LPC).

LPC, resolution of conflicting pronunciations and the frequency by regularity interaction

We first observed effects of regularity in the time window of the LPC, in the form of a frequency by regularity interaction, which, in turn, was task sensitive – emerging only in the reading aloud task. If frequency-by-regularity interactions reflect the relative ease of resolving conflicts between competing phonological representations, the failure to observe this interaction on the LPC in the proper name detection may be due to the fact that accurate performance does not require settling on a single pronunciation. Alternatively, it is possible

that during the name detection task some types of phonological information are not generated (e.g., on a dual-route account, that the print-to-sound pathway is not deployed).

Although the current data do not clearly adjudicate between these alternative hypotheses, we favor the former account. Pseudohomophone effects, seen both in behavior and on the N400, indicate that sound is activated from print even in tasks for which activating those sounds would be disadvantageous, as in a lexical decision task. In proper name detection, generating a pronunciation from a letter string could even be beneficial, as the same proper name can often be spelled in multiple ways (e.g., Catherine, Kathryn, Katherine). Therefore, it seems unlikely that phonological representations are not elicited in this task. Furthermore, the time window of the LPC is quite close to the average reaction time in speeded reading aloud. Participants of a similar age and education as our own set, reading similar sets of words, begin to initiate speech approximately 550–650 ms after the word was presented (Paap & Noel, 1991; Seidenberg et al., 1984). Our analysis of the LPC focused on a later time window (620–960ms) to isolate it from the N400; however, as can be seen in Figure 1, LPC effects begin to emerge around 450–500ms post stimulus onset. These effects, therefore, reside in a time window between semantic access (the N400) and the initiation of speech – i.e., when resolution of conflicting representations would be expected to occur.

Indeed, a number of previous studies have suggested that late positivities, similar to the LPC, are sensitive to conflict resolution processes (e.g., Coderre, Conklin & Van Heuven, 2011; West, 2003). In language, late positivities (including the P600 and the semantic P600) are often observed in response to words that make sentences ungrammatical (e.g., Coulson, King & Kutas, 1998) or semantically anomalous (e.g., Kuperberg, Sitnikova, Caplan & Holcomb, 2003; Kos, van den Brink, & Hagoort, 2012). These positivities have been interpreted by some as reflecting conflict that arises from a mismatch in sentence interpretation via semantic and syntactic (Kim & Osterhout, 2005; Kim & Sikos, 2011) or combinatorial and noncombinatorial (Kuperberg, 2007) routes to sentence comprehension. Although more work is needed to identify the precise functional sensitivity of the LPC, the idea that it reflects processes important for conflict resolution is broadly consistent with the current proposal. Our claim is that the effect of regularity on the LPC reflects the resolution of conflicting phonological representations engendered by words that violate regular spelling-to-sound patterns, especially when those words are low frequency (and thus less practiced). Such processes, however, are only necessary in tasks that involve actual production of a single form – and, our data show, seem only to be invoked under those task circumstances.

Conclusions

Most theories of visual word processing assume that meaning and sound are obligatorily processed whenever a written word is presented, regardless of the task being carried out; however, a number of results have suggested that the structure of the task can dictate how much meaning and sound are processed from the written word. The present findings speak to the time course and nature of the effects of task-context in a way that the previous findings could not. Specifically, we showed task-related differences in how basic lexical variables – lexical frequency and spelling-to-sound regularity – modulate ERP components

related to initial stages of semantic access (N400) and phonological conflict resolution (LPC). Although meaning and sound may be extracted from written words whenever they are presented, the nature of how these processes operate, as indexed by how much they are influenced by basic lexical variables, is a function of the goals of the task.

Acknowledgments

References

Balota DA, Chumbley JI. Are lexical decisions a good measure of lexical access? The role of the neglected decision stage. Journal of Experimental Psychology: Human Perception and Performance. 1984; 3:340–357. [PubMed: 6242411]

Balota DA, Chumbley JI. Monsell S, Doyle MC, Haggard PN. Where are the effects of frequency in visual word recognition tasks? Right where we said they were. Journal of Experimental Psychology: General. 1990; 119:231–237. 1989. [PubMed: 2141355]

Balota DA, Yap MJ, Cortese MJ, Hutchison KI, Kessler B, Loftis B, Neely JH, Nelson DL, Simpson GB, Treiman R. The English Lexicon Project. Behavior Research Methods. 2007; 39:445–459. [PubMed: 17958156]

Barber HA, Vergara M, Carreiras M. Syllable-frequency effects in visual word recognition: Evidence from ERPs. NeuroReport. 2004; 15:545–548. [PubMed: 15094520]

Becker CA. Allocation of attention during visual word recognition. Journal of Experimental Psychology: Human Perception and Performance. 1976; 2:556–566. [PubMed: 1011005]

Bentin S, Mouchetant-Rostaing Y, Giard MH, Echallier JF, Pernier J. ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. Journal of Cognitive Neuroscience. 1999; 11:235–260. [PubMed: 10402254]

Berndt RS, Reggia JA, Mitchum CC. Empirically derived probabilities for grapheme-to-phoneme correspondences in English. Behavior Research Methods, Instruments, & Computers. 1987; 19:1–9.

Briesemeister BB, Hofmann MJ, Tamm S, Kuchinke L, Braun M, Jacobs AM. The pseudohomophone effect: Evidence for an orthography-phonology-conflict. Neuroscience Letters. 2009; 455:124–128. [PubMed: 19368860]

Coderre E, Conklin K, van Heuven WJ. Electrophysiological measures of conflict detection and resolution in the Stroop task. Brain Research. 2011; 1413:51–59. [PubMed: 21840503]

Coltheart M, Rastle K, Perry C, Langdon R, Ziegler J. DRC: a dual route cascaded model of visual word recognition and reading aloud. Psychological Review. 2001; 108:204–256. [PubMed: 11212628]

Coulson S, King JW, Kutas M. Expect the unexpected: Event-related brain response to morphosyntactic violations. Language and Cognitive Processes. 1998; 13:21–58.

Dale, AM. Source Localization and Spatial Discriminant Analysis of Event-Related Potentials: Linear Approaches. Department of Cognitive Science. San Diego, La Jolla, CA: University of California; 1994.

Davelaar E, Coltheart M, Besner D, Jonasson JT. Phonological recoding and lexical access. Memory & Cognition. 1978; 6:391–402.

Donchin E, Coles MGH. Is the P300 component a manifestation of context updating. The Behavioral and Brain Sciences. 1988; 11:355–425.

Fabiani, M.; Gratton, G.; Federmeier, KD. Event-related brain potentials: Methods, theory, and application. In: Cacioppo, JT.; Tassinary, L.; Berntson, G., editors. Handbook of Psychophysiology. 3rd edition. Cambridge: Cambridge University Press; 2007. p. 85-119.

Federmeier KD, Kutas M. A rose by any other name: Long-term memory structure and sentence processing. Journal of Memory and Language. 1999; 41:469–495.

Forster, KI. Accessing the mental lexicon. In: Wales, RJ.; Walker, ECT., editors. New approaches to language mechanisms. North Holland: Amsterdam; 1976.

Forster KI. Category size effects revisited: Frequency and masked priming effects in semantic categorization. Brain and Language. 2004; 90:276–286. [PubMed: 15172545]

Forster KI, Shen D. No enemies in the neighborhood: absence of inhibitory neighborhood effects in lexical decision and semantic categorization. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1996; 22:696–713.

Grainger J, Holcomb PJ. Watching the word go by: On the time-course of component processes in visual word recognition. Language and Linguistics Compass. 2009; 3:128–156. [PubMed: 19750025]

Harm MW, Seidenberg MS. Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. Psychological Review. 2004; 111:662–720. [PubMed: 15250780]

Ibáñez AM, San Martín R, Hurtado E, López V. Methodological considerations related to sleep paradigm using event related potentials. Biological Research. 2008; 41:271–275. [PubMed: 19399340]

Jared D. Spelling-sound consistency affects the naming of high-frequency words. Journal of Memory and Language. 1997; 36:505–529.

Kim A, Osterhout L. The independence of combinatory semantic processing: Evidence from event-related potentials. Journal of Memory and Language. 2005; 52:205–225.

Kim A, Sikos L. Conflict and surrender during sentence processing: An ERP study of syntax-semantics interaction. Brain and Language. 2011; 118:15–22. [PubMed: 21481924]

Kos M, Van den Brink D, Hagoort P. Individual variation in the late positive complex to semantic anomalies. Frontiers in Psychology. 2012:3. [PubMed: 22291676]

Kuperberg GR, Sitnikova T, Caplan D, Holcomb PJ. Electrophysiological distinctions in processing conceptual relationships within simple sentences. Cognitive Brain Research. 2003; 17:117–129. [PubMed: 12763198]

Kuperberg GR. Neural mechanisms of language comprehension: Challenges to syntax. Brain Research. 2007; 1146:23–49. [PubMed: 17400197]

Kutas M, Federmeier KD. Electrophysiology reveals semantic memory use in language comprehension. Trends in Cognitive Science. 2000; 4:463–470.

Kutas M, Federmeier KD. Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). Annual Review of Psychology. 2011; 62:621–647.

Kutas M, Hillyard SA. Brain potentials during reading reflect word expectancy and semantic association. Nature. 1984; 307:161–163. [PubMed: 6690995]

Kutas M, McCarthy G, Donchin E. Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. Science. 1977; 197:792–795. [PubMed: 887923]

Laszlo S, Federmeier KD. Better the DVL You Know Acronyms Reveal the Contribution of Familiarity to Single-Word Reading. Psychological Science. 2007; 18:122–126. [PubMed: 17425530]

Laszlo S, Federmeier KD. The N400 as a snapshot of interactive processing: Evidence from regression analyses of orthographic neighbor and lexical associate effects. Psychophysiology. 2011; 48:176–186.

Laszlo S, Federmeier KD. Never seem to find the time: evaluating the physiological time course of visual word recognition with regression analysis of single-item event-related potentials. Language, Cognition and Neuroscience. 2014:1–20. ahead-of-print.

Lund K, Burgess C. Producing high-dimensional semantic spaces from lexical co-occurrence. Behavior Research Methods, Instruments, & Computers. 1996; 28:203–208.

MacLeod CM. Half a century of research on the Stroop effect: an integrative review. Psychological Bulletin. 1991; 109:163–203. [PubMed: 2034749]

Monsell, S. The nature and locus of the word frequency effect in reading. In: Besner, D.; Humphreys, G., editors. Basic processes in reading: Visual word recognition. Hillsdale, N.J.: Erlbaum; 1991.

Monsell S, Doyle MC, Haggard PN. Effects of frequency on visual word recognition tasks: Where are they. Journal of Experimental Psychology: General. 1989; 118:43–71. [PubMed: 2522506]

Morton J. The interaction of information in word recognition. Psychological Review. 1969; 76:165–178.

Newman RL, Connolly JF. Determining the role of phonology in silent reading using event-related brain potentials, Cognitive Brain. Research. 2004; 21:94–105.

Norris D. The Bayesian Reader: Explaining word recognition as an optimal Bayesian decision process. Psychological Review. 2006; 113:327–357. [PubMed: 16637764]

Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia. 1971; 9:97–113. [PubMed: 5146491]

Paap KR, Noel RW. Dual-route models of print to sound: Still a good horse race. Psychological Research. 1991; 53:13–24.

Plaut DC, McClelland JL, Seidenberg MS, Patterson K. Understanding normal and impaired word reading: computational principles in quasi-regular domains. Psychological Review. 1996; 103:56–115. [PubMed: 8650300]

Polich J. Updating P300: an integrative theory of P3a and P3b. Clinical Neurophysiology. 2007; 118:2128–2148. [PubMed: 17573239]

Polich J, Donchin E. P300 and the word frequency effect. Electroencephalography and Clinical Neurophysiology. 1988; 70:33–45. [PubMed: 2455628]

Rugg MD. Event-related brain potentials dissociate repetition effects of high-and low-frequency words. Memory & Cognition. 1990; 18:367–379. [PubMed: 2381316]

Rugg MD, Doyle MC. Event-related potentials and recognition memory for low and high-frequency words. Journal of Cognitive Neuroscience. 1992; 4:69–79. [PubMed: 23967858]

Ruz M, Nobre AC. Dissociable top-down anticipatory neural states for different linguistic dimensions. Neuropsychologia. 2008; 46:1151–1160. [PubMed: 18083202]

Seidenberg MS, Waters GS, Barnes MA, Tanenhaus MK. When does irregular spelling or pronunciation influence word recognition. Journal of Verbal Learning and Verbal Behavior. 1984; 23:383–404.

Sereno SC, Rayner K, Posner MI. Establishing a time-line of word recognition: evidence from eye movements and event-related potentials. NeuroReport. 1998; 9:2195–2200. [PubMed: 9694199]

Smith NJ, Levy R. The effect of word predictability on reading time is logarithmic. Cognition. 2013; 128:302–319. [PubMed: 23747651]

Solomon RL, Postman L. Frequency of usage as a determinant of recognition thresholds for words. Journal of Experimental Psychology. 1952; 43:195–201. [PubMed: 14927822]

Spironelli C, Angrilli A. Influence of phonological, semantic and orthographic tasks on the early linguistic components N150 and N350. International Journal of Psychophysiology. 2007; 64:190–198. [PubMed: 17363097]

Stenberg G, Lindgren M, Johansson M, Olsson A, Rosen I. Semantic processing without conscious identification: evidence from event-related potentials. Journal of Experimental Psychology: Learning, Memory & Cognition. 2000; 26:973–1004.

Van Petten C, Kutas M. Interactions between sentence context and word frequency in event-related brain potentials. Memory & Cognition. 1990; 18:380–393. [PubMed: 2381317]

Van Petten C&, Luka BJ. Prediction during language comprehension: Benefits, costs, and ERP components. International Journal of Psychophysiology. 2012; 83:176–190. [PubMed: 22019481]

Vergara-Martinez M, Swaab T. Orthographic neighborhood effects as a function of word frequency: An event-related potential study. Psychophysiology. 2012; 49:1277–1289. [PubMed: 22803612]

Vogel EK, Luck SJ, Shapiro KL. Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. Journal of Experimental Psychology: Human Perception & Performance. 1998; 24:1656–1674. [PubMed: 9861716]

West R. Neural correlates of cognitive control and conflict detection in the Stroop and digit-location tasks. Neuropsychologia. 2003; 41:1122–1135. [PubMed: 12667546]

Yarkoni T, Balota DA, Yap MJ. Beyond Coltheart's N: A new measure of orthographic similarity. Psychonomic Bulletin & Review. 2008; 15:971–979. [PubMed: 18926991]
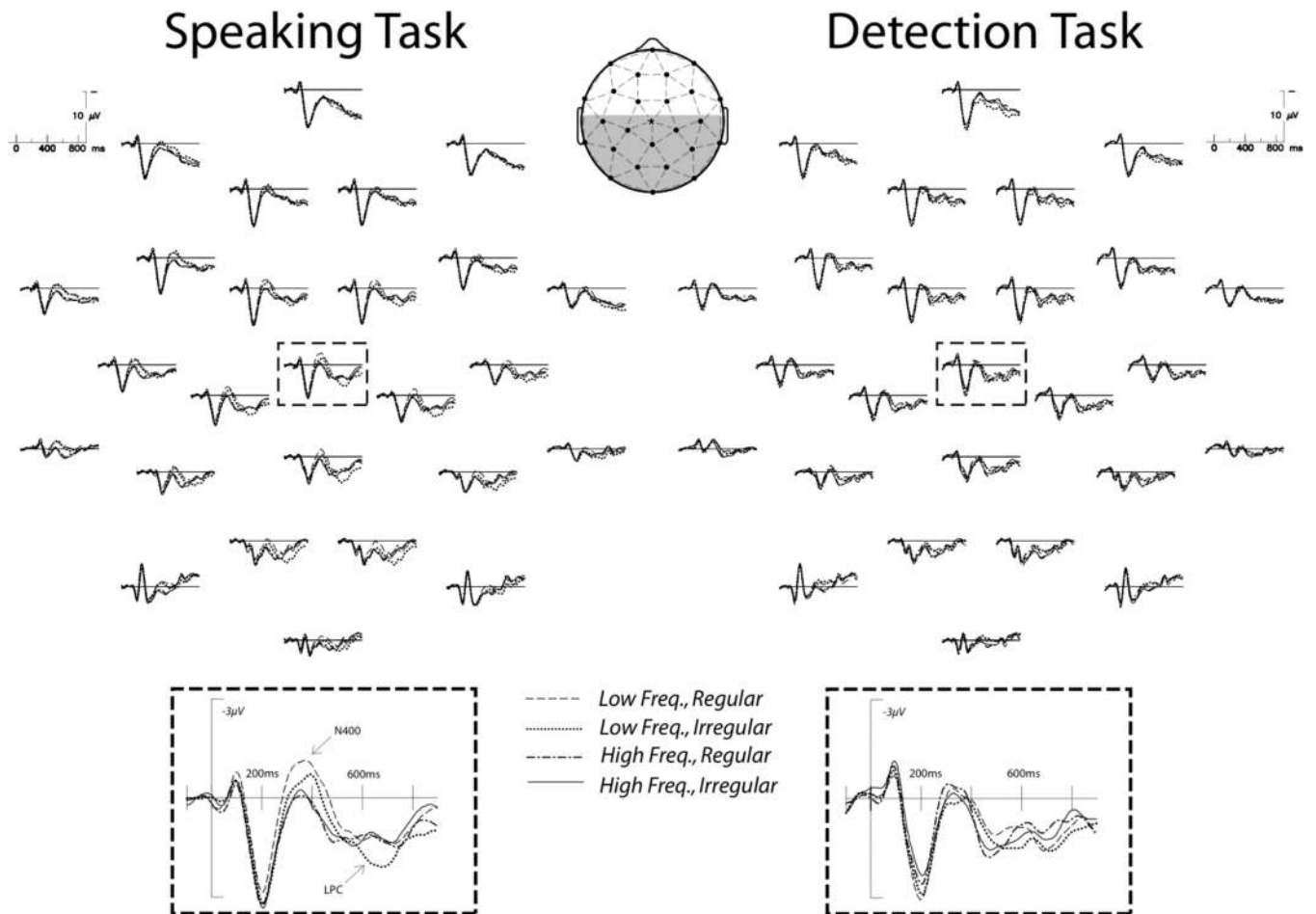
Figure 1.
Grand average ERPs for each task, plotted at all channels for each of the four item types (see legend) distributed in accordance with the head schematic (channels in grey were used in the LPC statistical analysis). Pull out plots show a larger view of a representative middle central channel.
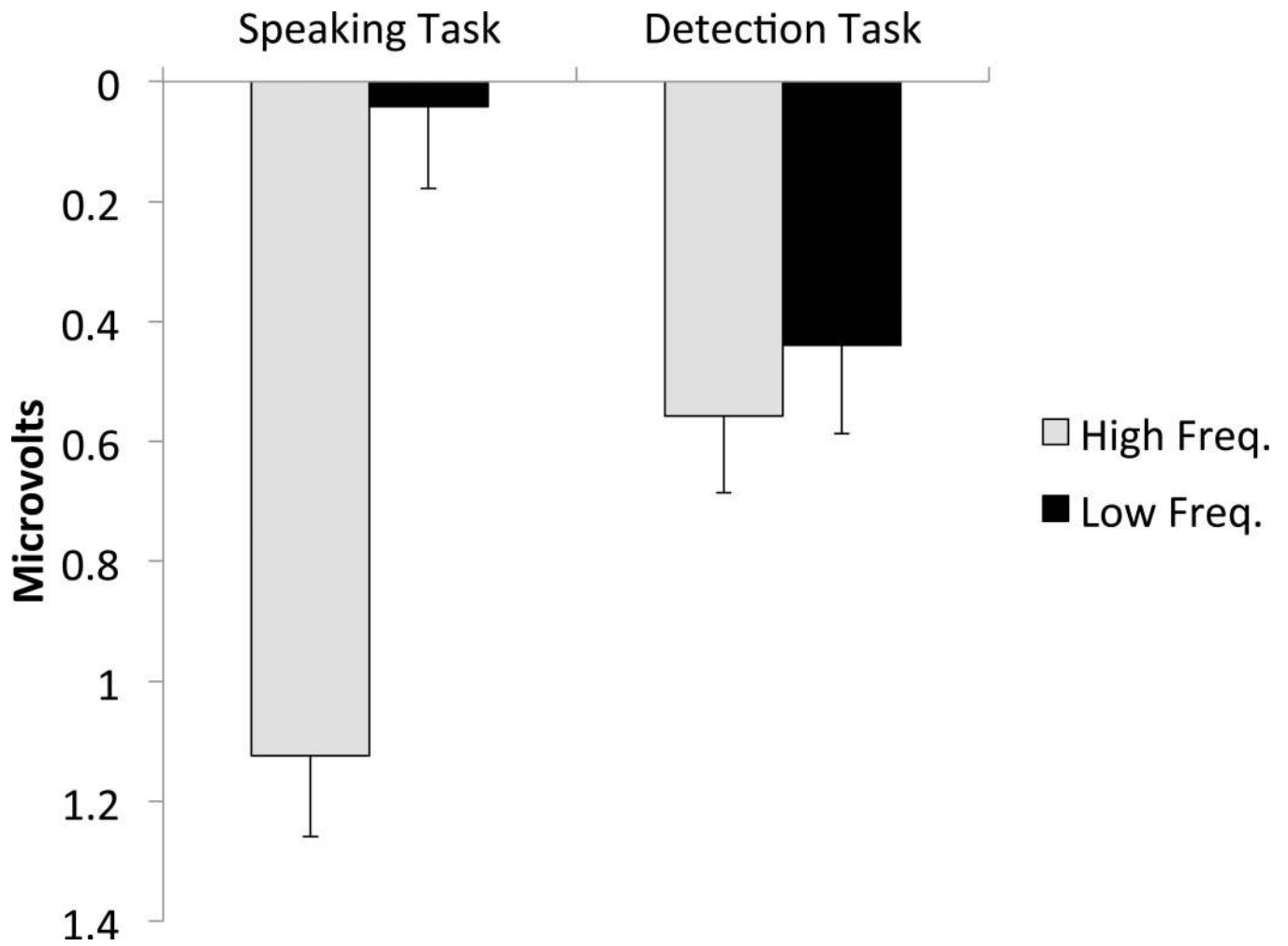
Figure 2.
Mean amplitude of high and low frequency items in each task collapsed across spelling-to-sound regularity, measured between 300 and 450ms and averaged over all 26 channels. Positive is plotted down, in keeping with ERP figure conventions.
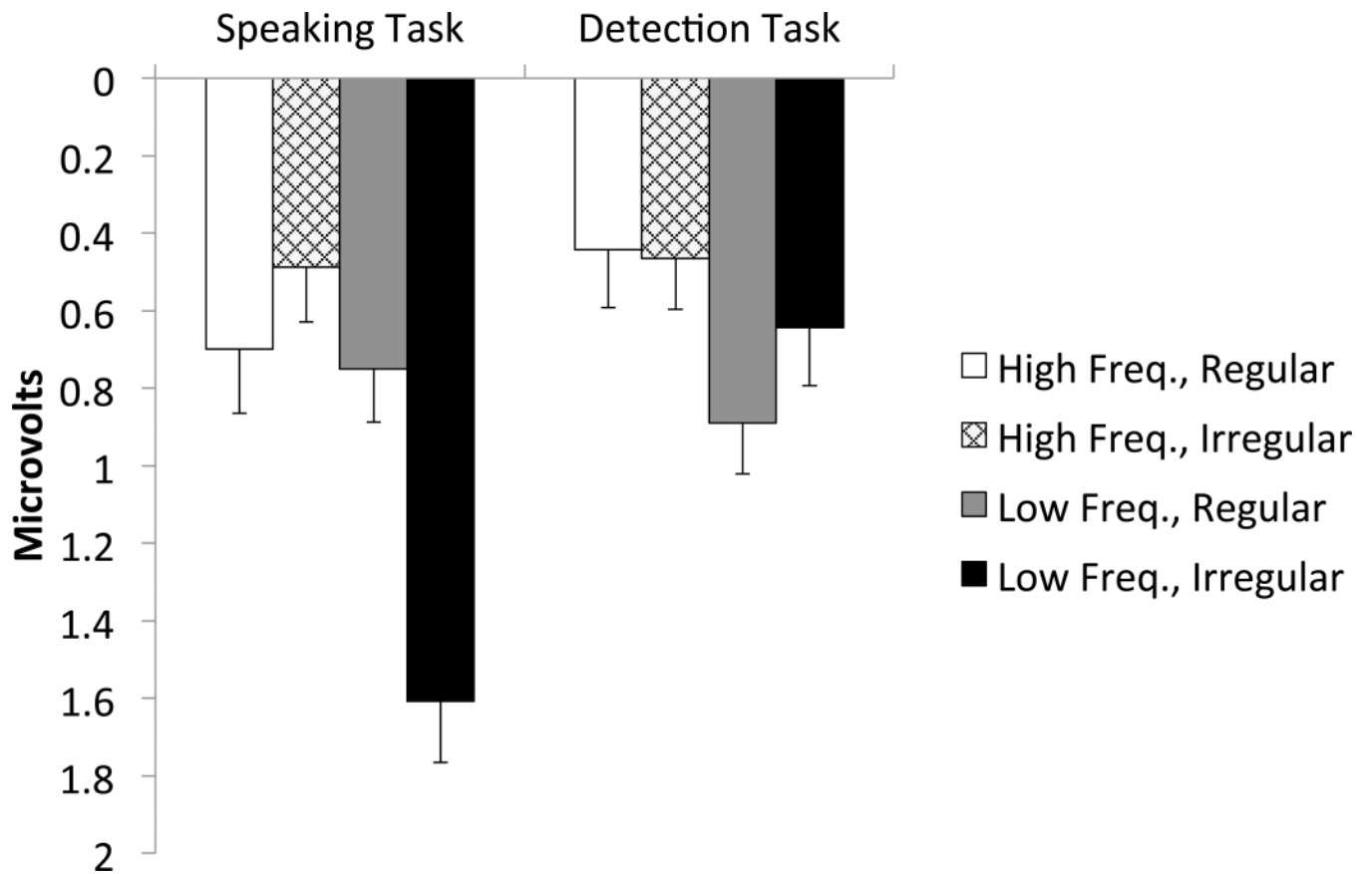
Figure 3.
Mean amplitude of the four critical item types in each task, measured between 620 and 920ms and averaged over the 15 posterior channels. Positive is plotted down, in keeping with ERP figure conventions.
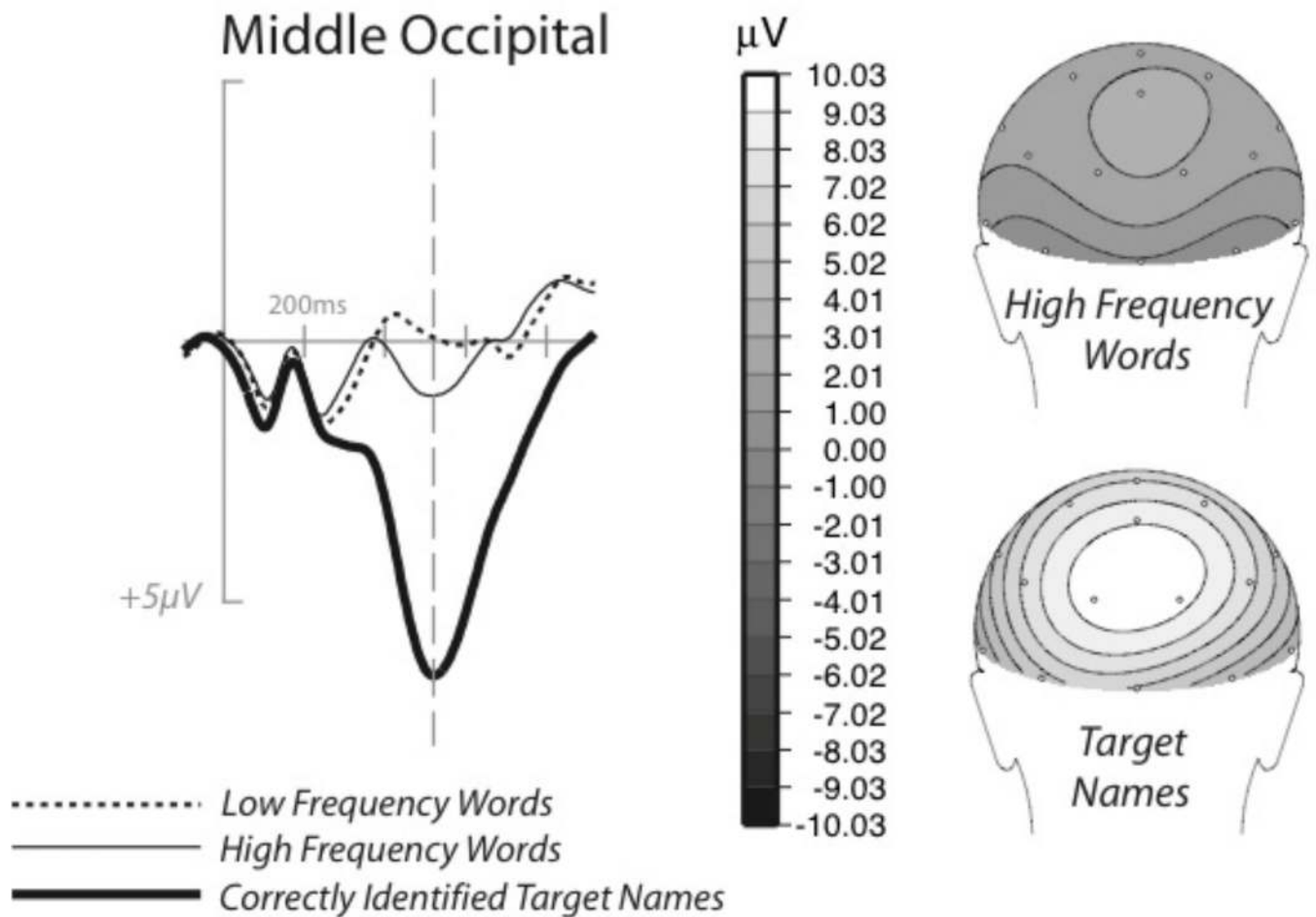
Figure 4.
Grand average ERPs and from the detection task in response to correctly identified targets and high and low frequency items, collapsed across spelling-to-sound regularity, as well as the scalp map (at 515 ms) for the high frequency word items and the correctly identified targets. The dashed line highlights the similarity between the P300 peak to targets and the response to high frequency word items at 515 ms, as does the similar distribution of the effect shown in the scalp maps.

Table 1

Lexical and orthographic properties of the stimuli

		Log Freq.	Regularity	Coltheart's N	OLD20	Length	Bigram Mean
HF/Reg	TURN	11.4	.94	6.5	1.6	4.3	1474
HF/Irr	MOVE	11.4	.73	5.5	1.7	4.4	1336
LF/Reg	LUMP	7.8	.97	5.9	1.6	4.3	1152
LF/Irr	CHUTE	7.5	.67	5.8	1.6	4.3	1294