



Published in final edited form as:

Lang Cogn Process. 2014 ; 29(5): 642–661. doi:10.1080/01690965.2013.866259.

Never Seem to Find the Time: Evaluating the Physiological Time Course of Visual Word Recognition with Regression Analysis of Single Item ERPs

Sarah Laszlo^{(1),(2)} and Kara D. Federmeier^{(3),(4),(5)}

⁽¹⁾Department of Psychology, State University of New York, Binghamton

⁽²⁾Program in Linguistics, State University of New York, Binghamton

⁽³⁾Department of Psychology, University of Illinois, Urbana-Champaign

⁽⁴⁾Program in Neuroscience, University of Illinois, Urbana-Champaign

⁽⁵⁾Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign

Abstract

Visual word recognition is a process that, both hierarchically and in parallel, draws on different types of information ranging from perceptual to orthographic to semantic. A central question concerns when and how these different types of information come online and interact after a word form is initially perceived. Numerous studies addressing aspects of this question have been conducted with a variety of techniques (e.g., behavior, eye-tracking, ERPs), and divergent theoretical models, suggesting different overall speeds of word processing, have coalesced around clusters of mostly method-specific results. Here, we examine the time course of influence of variables ranging from relatively perceptual (e.g., bigram frequency) to relatively semantic (e.g., number of lexical associates) on ERP responses, analyzed at the single item level. Our results, in combination with a critical review of the literature, suggest methodological, analytic, and theoretical factors that may have led to inconsistency in results of past studies; we will argue that consideration of these factors may lead to a reconciliation between divergent views of the speed of word recognition.

Keywords

Visual Word Recognition; ERPs; Multiple Regression

Introduction

Extracting meaning from text is a subjectively effortless and instantaneous process for adult, literate readers. However, despite the rapidity and ease with which it is accomplished, visual

Address correspondence to: Sarah Laszlo, Department of Psychology, 4400 Vestal Parkway East, Binghamton, New York, 13903, Phone: 607.777.3380, slaszlo@binghamton.edu.
Reprint requests should be addressed to S.L.: cogneuro@alum.mit.edu.

word recognition is actually a temporally extended process comprised of accessing information at multiple levels of representation both hierarchically and in parallel. Because reading is so massively prevalent in advanced society, and because it requires the access of so many types of information, understanding its time course of processing with a high degree of functional specificity has been of special interest to a broad cross section of researchers employing a variety of empirical methods.

Event-related potential studies of word recognition

Event-related potentials (ERPs) have been an especially useful method for studying visual word recognition as it unfolds, as they enable the collection of data continuously, and with millisecond-level resolution, throughout the recognition process (as opposed to supplying only end-state data, as with button press or naming measures). ERP research on visual word recognition has thus taken advantage of this temporal sensitivity to identify at what time points after an orthographic stimulus has been presented various types of information (e.g., orthographic, semantic) come online and are processed.

In Tables 1, 2, and 3, we present a summary of ERP studies that have looked for effects of orthographic (Table 1) and semantic (Table 3) variables in the response to visual word forms, as well as studies that have looked for effects of written frequency (Table 2; written frequency itself being neither clearly orthographic or semantic; see Simon, Lewis, & Marantz, 2012). Effects of orthographic variables are well-attested within ~150 ms of word apprehension and continue to manifest throughout the first half second of processing. Reports of semantic effects, in contrast, are sparse and variable in form until closer to 300 ms. Frequency effects have been reported early in a couple of studies but begin to become better attested in time windows in between orthographic and semantic effects.

This general pattern seen across studies has been formalized in the functional architecture of the Bimodal Interactive Activation Model (BIAM; Grainger & Holcomb, 2009). In the BIAM, the visual presentation of a word is immediately followed by the apprehension of visual features, which are combined into orthographic features that are initially location specific, then spatially invariant. Around 150 ms into processing (in the epoch of what Grainger and Holcomb call the N/P150 component), ERPs are sensitive to the repetition of single letters and words, in a manner that is independent of size but sensitive to font (Chauncey et al., 2008), letter case (Petit et al., 2006), and position (Dufau et al., 2008), suggesting that at this point in processing representations are still perceptual and retinotopic. Spatially invariant representations of orthographic features are then formed and combined into more complex orthographic representations during the epoch of the N250 component, which has a right posterior distribution and has been observed in masked priming contexts involving both repetition priming (e.g., Holcomb, & Grainger, 2006; Kiyonaga, Midgley, Holcomb, & Grainger, 2007) and semantic priming (Carreiras, Andoni Dunabeitia, & Molinario, 2009). N250 priming effects are not observed with cross-modal presentation (e.g., Kiyonaga et al., 2007), nor for orthographically unrelated masked primes, and are thus thought to be selective to orthographic processing at the form-meaning interface. Because pseudowords show similar N250 repetition priming effects to those elicited by words, it

seems unlikely that orthographic inputs have been uniquely identified at this stage of processing.

The N250 is followed, between about 250 and 500 ms, by the N400 component, which has a central posterior distribution and is thought to have diverse neural sources, including particularly strong sources in the left anterior temporal lobe (e.g., Halgren, Dhond, Christensen, Van Petten, Marinkovic, & Lewine, 2002). In the BIAM, as in a multitude of other frameworks (for review, see Kutas & Federmeier, 2011), the N400 is thought to reflect semantic processing -- for both verbal and nonverbal stimuli (e.g., Barrett & Rugg, 1990; Ganis & Kutas, 2003). Importantly, N400 amplitude reductions elicited by repetition and predictability in higher-level (e.g., sentence) contexts are observed not only for words (which clearly have semantics) but also for pseudowords (e.g., GORK) and even illegal consonant strings (e.g., XFQ; Laszlo & Federmeier 2008; 2009; 2011). In fact, even without the benefit of highly constraining sentence contexts, N400 effects can still be observed to meaningless, illegal strings of letters (Laszlo, Stites, & Federmeier, 2012), suggesting that, although semantic access is being attempted during the N400 epoch, neural activity in this time window is not limited to one uniquely identified word form.

A strength of the BIAM is that it is consistent with what is known about the anatomical and physiological bases of the elicited brain activity in early time windows (e.g., Di Russo, Martinez, Sereno, Pitzalis, & Hillyard, 2001) and accords well with the timecourse of visual processing and recognition described in the ERP literature on face and object processing. As reviewed in Federmeier, Kutas, & Dickson (in press), it takes about 150–200 ms for the brain to extract enough visual information to allow stimuli to be categorized at the basic level (as faces, objects, or strings: e.g., Schendan, Ganis, Kutas, 1998, Rossion & Jacques, 2011). Between about 200 and 300 ms, ERP responses to objects and faces reflect grouping (e.g., Schendan & Kutas, 2007) and “structural encoding” processes (Schweinberger et al., 2002) that create a more complex perceptual representation of the stimulus. Across these literatures, semantic access (e.g., sensitivity to whether or not a face or object is familiar or contextually congruent) is associated with N400-like potentials beginning around 250–300 ms (e.g., Ganis et al., 1996; Paller et al., 2000). The ERP literatures on face, object, and word processing thus converge in providing strong general support for a timecourse of processing that culminates in a common semantic access process reflected in the N400 component.

However, despite its strong points, further validation of the BIAM is needed. First, the timecourse of orthographic effects described by the BIAM has been largely derived from a single paradigm: masked priming. For the BIAM to serve as a viable model for general word recognition processes, this timecourse needs to be replicated outside of the confines of a single (and arguably not particularly ecologically valid) paradigm. Replicating the timecourse of processing described by the BIAM is also important because, as can be seen in Table 3, there have been occasional reports (none of which came from masked priming paradigms) of early (i.e., pre N400) semantic effects -- or frequency effects that are interpreted as reflecting later-stage, whole word processing -- that are at odds with the timecourse the BIAM describes. These early effects have not yet been very consistent or well-replicated, thus, one way or the other, both the BIAM and accounts of earlier semantic

processing could benefit from replication. Interestingly, in another parallel across literatures, early semantic effects have sometimes been reported in the ERP literature on object and face processing as well; however, further investigation has pointed to alternative explanations for these effects, in terms of perceptual properties (spatial frequency, luminance) that correlate with semantic variables (see, e.g., VanRullen & Thorpe, 2001; Roission & Carahel, 2011; Rossion & Jacques, 2008). It is important, therefore, to examine the time course of availability of semantic information in a design that can account for variability due to lower-level stimulus properties.

Eyetracking studies of word recognition

Although reports of early semantic effects in the ERP literature are not well-attested, they gain some weight -- and thus especially merit further examination -- because they have been taken to be more consistent with the timecourse of word processing as described in the literature derived from studies using eyetracking methods. The fixations and saccades that allow the intake of information from text during natural reading can be monitored continuously with high temporal resolution, making eyetracking methods, like ERPs, particularly useful for studying the timecourse of word recognition. Strikingly, the functional time course described in the BIAM has often been taken to be at odds with the time course of word processing revealed by eye-tracking data (see Sereno & Rayner, 2003; Rayner & Clifton, 2009), which places the temporal locus of unique identification of a wordform as being between 100 and 200 ms post-stimulus onset—that is, within the epoch of the evoked P1/N1 complex (e.g., Sereno, Rayner, & Posner, 1998; Sereno & Rayner, 2003).

Just as the functional architecture of the BIAM theoretically consolidates the results of many ERP studies of visual word recognition, an examination of the E-Z Reader model (Pollatsek, Reichle, & Rayner, 2006), a prominent model of eye control during reading lines of connected text, can present a good theoretical summary of the eye-tracking research on the same topic. In the E-Z Reader model, early visual processing derives information (for example, where spaces are between words) that is relevant to the programming of the next eye movement (see Reichle, Rayner, & Pollatsek, 2012). In the model, the visual processing stage takes a fixed 50 ms to run to completion. This figure is based on the empirically-derived “eye-mind lag”; that is, the amount of time it takes for information to get from the retina to the striate cortices (e.g., Clark, Fan, & Hillyard, 1995). The visual processing stage is then followed by two substages of lexical processing: the familiarity check and lexical access.

In the model, the time it takes to complete the familiarity check is a function of the frequency of a word and its predictability in context. Inclusion of the familiarity check as a separate stage from lexical access allows the model to explain spillover effects (e.g., Rayner & Duffy, 1986), that is, the finding that the difficulty of identifying a word can increase the time it takes to identify a subsequent word. The time it takes to complete lexical access is then simply a fixed proportion of the time it takes to complete the familiarity check, which has the result that, the faster the familiarity check, the faster lexical access unfolds (see Pollatsek & Rayner, 1990). The familiarity check provides information about how difficult

the currently attended word is likely to be to process, and, when completed, signals the programming of an eye movement to the subsequent word (e.g., Reichle, Tokowicz, Liu, & Perfetti, 2011). Completion of lexical access, in contrast, provides unique recognition of the currently attended word and signals a shift in attention to the subsequent word; in the E-Z Reader model, only a single word can be attended at any particular time (e.g., Reichle, Pollatsek, & Rayner, 2012b). Importantly, the completion of lexical access for a fixated word typically occurs *before* the subsequent word is fixated; thus, gaze duration for a particular word is often taken to be an *upper bound* on the lexical processing and unique identification of that word (e.g., Pollatsek et al., 2006).

With this assumption in mind, it becomes evident that the E-Z Reader model estimates a time course of unique visual word recognition that is faster than that proposed in the BIAM¹. The mean duration of fixation on a single word during silent reading is often given as ~200–250 ms (e.g., Rayner, 2009; Dambacher & Kliegl, 2007). The proposition that 250 ms is an upper bound on the amount of time it takes to uniquely identify a word, is, then, in conflict with the BIAM's proposition that high level orthographic processing (which precedes word identification and semantic access) is still occurring as late as the N250 component.

Much has been made of the apparent discrepancy in processing times suggested by eye movements and ERPs (e.g., Clifton & Rayner, 2009), and so some studies have attempted to minimize differences between experiments using both ERP and eyetracking in the service of trying to identify a systematic relationship between data obtained with the two methods. For example, one study (Dambacher & Kliegl, 2007) presented a common set of materials to two participant groups, with one group undergoing EEG as they read the materials serially at fixation and the other having their eye movements tracked as they read the same texts presented all at once. There was a strong correlation between fixation durations and N400 mean amplitudes, leading the authors to conclude that the two measures must index processes sharing at least one “common stage”.

A small number of additional studies have attempted to reconcile ERP and eye-movement data more directly, by recording the two measures simultaneously, despite the methodological difficulties (Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, 2011; Dimigen, Kliegl, & Sommer, 2012; for review see Kliegl, Dambacher, Dimigen, Jacobs, & Sommer, 2012). Two important conclusions of this work are that, first, N400 priming effects can be demonstrated to onset much earlier when parafoveal preview is available than when it is not (Dimigen, Kliegl, & Sommer, 2012) and, second, that when ERPs are collected during natural reading, N400 effects in response to a large proportion of items can be observed to onset while those items are still being fixated (Dimigen et al., 2011). Thus, it would seem that estimates of lexical processing time may be less discrepant across methodologies than has sometimes been suggested. Nevertheless, the longstanding controversy between the timecourses suggested across methods, combined with occasional contradictory reports in

¹In this paper, we focus on one prominent model of eye control during reading, which has been central to arguments about timing discrepancies between ERPs and eye movements. There are, of course, other models of eye control such as the SWIFT model (Engbert, Nuthmann, Richter, &

the ERP literature, compels further research examining the timecourse of word processing with a design that can address some of the weakness of prior research, and thus possibly reconcile apparently discrepant findings.

The Present Study

In the present study, we measure ERPs in order to examine when various sources of information (orthographic, frequency-based, semantic) affect word processing, strengthening the extant literature by using (1) a different task (not masked priming), (2) high levels of power (helpful for looking for infrequently replicated, and therefore, possibly relatively small, effects), and (3) novel analytical techniques that allow us to look at the influence of multiple sources of variability in parallel.

One common feature of almost all of the studies reviewed in Tables 1–3 is that they used factorial designs and analyses in order to determine points in time at which particular lexical variables began to reliably affect the waveform—for example, measuring when average waveforms elicited by high and low frequency words began to differ in order to infer the point at which frequency affects the waveform. As others have pointed out (e.g., Hauk et al., 2006; Hauk, Pulvermuller, Ford, Marslen-Wilson, & Davis, 2009), this type of analysis, although common, is not ideal, largely due to the problem of the pervasive intercorrelation of lexical variables. Hauk et al., (2012), for example, have demonstrated that even when exceptional care is taken to match item groups on important lexical variables, it is pragmatically impossible to do so for even as few as 4 variables. Intercorrelations between lexical variables make inferences about when a particular lexical property begins to have an effect on the waveform problematic, as it is difficult if not impossible to know whether the source of an effect is, for example, orthographic, semantic, or a mixture of both.

A solution that has been put forward for this problem (e.g., Hauk et al., 2006; Hauk, Pulvermuller, Ford, Marslen-Wilson, & Davis, 2009; Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004) is a move away from factorial designs and towards the use of items regression. Items multiple regression, in particular, has the attractive property of enabling exploration of the simultaneous effects of an arbitrary number of intercorrelated variables. Despite its many advantages as a statistical technique, however, items multiple regression is used relatively rarely in the ERP literature. The reason for this is straightforward: with the number of participants typically included in an ERP study, there is usually not enough power to successfully perform items analysis (see Dambacher, Kliegl, Hofman, & Jacobs, 2006, for an example of this problem) without the use of statistical tests that do not correct for the number of comparisons that must be done in order to identify a unique starting time for the influence of lexical variables (i.e., separate tests every 10–50 ms; Sereno et al., 1998; Hauk et al., 2006, 2009).

In the present study, therefore, we aimed to make use of our large corpus of visual word recognition ERP data, which includes stable single-item ERPs (the “single item ERP corpus,” Laszlo & Federmeier, 2011) to perform critical analyses of the time course of visual word recognition out of context (and without masking) that have previously not been possible. The corpus consists of ERPs from 6 channels, offering coverage of the frontal, parietal, and occipital scalp, representing responses averaged over participants—but not over

items—for single presentations of 75 each words (e.g., DOG), familiar acronyms (e.g., DVD), pseudowords (e.g., DAL), and illegal strings of letters (e.g., DSN), presented in a relatively passive list-reading paradigm (participants monitored an unconnected stream of text for proper English first names; items were presented at a comfortable pace with no visual interference). The single item ERP corpus is particularly well suited for analysis by items multiple regression for two reasons. First, the unusually large number of participants ($N=120^2$) included in the corpus allows for the creation of stable single-item ERPs, providing sufficient power for successful items multiple regression, as we have reported in previous work (Laszlo & Federmeier, 2011), and the ability to find even small effects if they are truly present. Second, by design, the items included in the corpus vary broadly in their lexical characteristics, making them well-suited for analysis with continuous measures; the items essentially encompass the full range in English of the majority of the variables we will analyze (e.g., written frequency, orthographic neighborhood size, number of lexical associates).

In what follows, we take advantage of the unique nature of the single item ERP corpus to conduct a novel set of analyses which hopefully provide a novel view of the time course of visual word recognition out of context. Through consecutive simultaneous items multiple regressions (that is, items multiple regressions performed on consecutive 10 ms bins), we track the time course of influence of orthographic variables (orthographic neighborhood size, log bigram frequency, and log summed neighbor frequency) on the one hand and semantic variables (concreteness, imageability, number of senses, number of lexical associates, and noun-verb ambiguity) on the other. This approach differs from the more typical, component-centered analysis often utilized in ERP studies by being more data-driven: time windows of interest are not determined in advance nor constrained by waveform morphology (e.g., component peaks) but are revealed by the data. Such an approach is especially useful when the issue at hand is one of *when* an effect begins to take place (see Amsel, 2011).

We conduct the semantic analysis both with and without a prior step of removing variance explained by orthographic variables. This approach is preferable to a factorial approach of (for example), simply comparing waveforms elicited by words and nonwords to see when an influence of the (presumably greater) semantics of the words begins to be evident, in that words and nonwords can differ on any number of other lexical properties besides lexicality *per se* (e.g., bigram frequency, orthographic neighborhood size, neighbor consistency, frequency of neighbors, to name only a few). This approach also allows relative certainty as to the identity of the variables that actually produce a given effect; for example, when semantic effects are observed after variance due to orthographic variables has already been accounted for, it is clear that those effects are truly semantic, and not, instead, effects of correlated orthographic variables.

²In order to collect high-quality data from such a large number of participants in a reasonable period of time, a reduced recording montage was used with only 6 electrodes at Middle Prefrontal, Middle Central, Left, Right, and Central Parietal, and Middle Occipital sites. As is demonstrated in the meta-reviews of Tables 1, 2, and 3, this montage covers the vast majority of topographic regions where effects of lexical variables have been observed in the literature.

We compare the time course revealed in these analyses with those predicted by both the BIAM and frameworks that point to an earlier locus of lexical processing, such as E-Z Reader. Under the BIAM, we would expect orthographic effects in the temporal regions of the N/P150 and N250, with semantic effects emerging in the region of the N400. Under an earlier lexical processing framework, in contrast, we would expect to see a much faster time course, with similar early effects of orthographic variables, but with semantic effects emerging prior to 200 ms post stimulus onset. The primary question of interest is whether the BIAM will be replicated here with a new task, or if, rather, our powerful analytic methods will enable us to replicate early semantic effects that would raise challenges for its assumptions.

Methods

Event-Related Potentials

The methods pertaining to collection of the single-item ERP corpus have been described in detail elsewhere (Laszlo & Federmeier, 2011). However, we present salient details here for clarity, as well as additional details pertaining to the analyses conducted here that were not conducted in Laszlo & Federmeier, 2011. EEG was recorded continuously from 120 participants (58 female, age range 18–24, mean age 19.1), who monitored an unconnected list of words (75), acronyms (75), pseudowords (75), and illegal strings (75) for proper names (150). Participants were required to press a button in response to names; no response was required for the other four, critical, item types. Words, pseudowords, acronyms, and illegal strings all repeated once at an inter-item lag of 0, 2, or 3 intervening items. Acronym familiarity was assessed by a paper and pencil post-test (identical to that described in Laszlo & Federmeier, 2007), and only EEG responses to acronyms correctly identified by a given participant were included in the averaged ERPs computed for that participant.

EEG was recorded from 6 Ag/AgCl electrodes embedded in an electrocap. We sampled from middle prefrontal, middle parietal, middle central, left middle central, right middle central, and middle occipital electrode sites. This reduced electrode montage was necessary in order to enable the collection of 120 participants in a reasonable period of time. In order to choose an appropriate reduced montage, we first conducted a pilot experiment identical in all ways to that reported here (and in Laszlo & Federmeier, 2011), with the exception that EEG was digitized at 64 scalp electrodes. The results of that study suggested that effects of the variable types of interest (e.g., lexical frequency, orthographic neighborhood size) are maximal in both early and late time windows over midline parietal channels (see Figure 1). This result conforms to the meta-review of the literature presented in Tables 1, 2, and 3, which makes clear that the vast majority of effects of lexical variables on any epoch of the ERP are visible over this selection of sites³. Figure 1 displays the topography of the lexical frequency effect around the peak of the P2 and N400 components. These results, as well as our extensive review of the literature, are the basis for our decision to use the specific 6 electrodes mentioned above.

³An exception are effects reported on the frontal N200 component in go/nogo tasks; however, we did not use a go/nogo task and thus would not expect N200 effects.

The electrooculogram (EOG) was also recorded using a bipolar montage of electrodes placed at the outer canthi of the left and right eyes in order to monitor eye movements; blinks were monitored with an electrode at the suborbital ridge. Single-item ERPs were formed time-locked to the onset of each item, averaged over participants, but not over items or over multiple presentations of items. More traditional, item-aggregated ERPs (representing, for example, the response to all words), were also formed by averaging over both items and participants.

By design, the lexical characteristics of the items varied widely, in anticipation of submitting them to regression analyses. Table 4 displays examples of each item type, along with means for each item type for all the lexical variables considered in the present analyses: log bigram frequency, log whole item frequency, orthographic neighborhood size, log summed frequency of orthographic neighbors, concreteness, imageability, number of senses, number of lexical associates, and noun verb ambiguity. Bigram frequency was retrieved from the Medical College of Wisconsin Orthographic Wordform Database (MCWord, Medler & Binder, 2005). Written frequency was obtained from the Wall Street Journal Corpus (Marcus, Santorini, & Marcinkiewicz, 1993). Orthographic neighborhood size was computed as the total number of words that could be formed by replacing one letter of a target item (i.e., Coltheart's N; Coltheart, Davelaar, Jonasson, & Besner, 1977), as indicated by MCWord. Neighbor frequency was, in turn, computed as the logarithm of the summed frequency of all of an item's orthographic neighbors, with frequency estimates drawn from the Wall Street Journal corpus (Marcus, et al., 1993). Concreteness and imageability were both retrieved from the MRC Psycholinguistic Database (Coltheart, 1981). Number of senses was retrieved from WordNet (Princeton University, 2010). Number of lexical associates was retrieved from the South Florida Norms (Nelson, McEvoy, & Schreiber, 1998). Noun verb ambiguity was computed by first retrieving the frequency of occurrence of each item as either a noun or a verb from WordNet. Then an ambiguity metric was computed as the frequency of the dominant usage divided by the summed frequency of both usages. Thus, higher numbers on the ambiguity metric indicate less ambiguous items.

Statistical Methods

In preparation for items multiple regression analysis, single-item ERPs were bandpass filtered between .2 and 20 Hz and were then measured for mean amplitude in consecutive 10ms bins from 0 ms to 920 ms post stimulus onset—the full duration of each EEG sweep. This resulted in 600 time course vectors with 92 entries at each EEG channel—that is, one vector for each channel for each of 2 presentations of each item, with mean amplitude entries every 10 ms.

The selection of a regression model proceeded as follows. We began by considering a large model simultaneously including a total of 11 predictors of the semantic, orthographic, and lexical types. However, when we computed the intercorrelations of the predictor variables in this model, we observed that four of the 11 variables had Variance Inflation Factors (VIFs) with values > 4 . For this reason, we discarded the large model.

The matrix of intercorrelations for the large model had three eigenvalues > 1 . Therefore, we performed a Factor Analysis (FA) of the predictors in the large model to identify three

underlying factors. The factor analysis revealed that the first latent factor in these predictor variables loaded most strongly on frequency (highest loading: log Wall Street Journal frequency, 2nd loading: log trigram frequency; 3rd loading, COALS vector length. COALS is a frequency-sensitive measure of co-occurrence in text; Rohde, Gonnerman, & Plaut, Under Review). The second factor loaded most strongly on orthography (all three top loadings were measures of orthographic neighborhood), and the third factor loaded most strongly on semantics (all three top loadings were measures of lexical-semantic association). For analysis we therefore divided the original, large, regression model into three independent models including either orthographic, semantic, or frequency variables. Using the FA to guide the selection of regression models in this manner ensured that no variable in any of the models used has a VIF > 4. Table 5 displays the intercorrelations of variables in each of the final regression models, along with VIF for each variable.

Two predictor matrices were formed: one for what we will refer to as the *orthographic* multiple regression, and one for what we will refer to as the *semantic* multiple regression. The orthographic multiple regression included as predictors: log bigram frequency, orthographic neighborhood size (N), and log summed frequency of orthographic neighbors. The semantic multiple regression included as predictors: concreteness, imageability, number of senses, number of lexical associates, and noun verb ambiguity. All orthographic information was available for all 300 items, but semantic information was available for only subsets of the word items. For this reason, only words were submitted to semantic regressions. Words with a missing value on a particular semantic predictor (e.g., a word not present in the MRC and thus without a concreteness value) were assigned the mean value for that predictor.

Significance levels for all statistical tests were obtained using the permutation test technique. To estimate the overall significance of regression models with permutation tests, the unpermuted R-squared is compared with a distribution of values of R-squared obtained by recomputing the statistic for every permutation of the response vector (Anderson, 2001). The p-value for R-squared is then simply the proportion of permutations that produce a more extreme statistic than the observed statistic, such that if a very extreme statistic is observed, there will be very few permutations that produce a more extreme statistic.

To estimate significance levels for individual predictors, a slightly different approach is used. First, the t-statistic for each individual predictor in a given model is computed. The distribution of permutation t-statistics is then generated based on permuting residuals of a reduced regression model (i.e., a model not including the predictor of interest) rather than on permuting the response vector directly (for a full description of this procedure, see Freedman & Lane, 1983).

Permutation tests produce exact p values when all permutations of the response vector are considered (Blair & Karniski, 1993); however even with only 75 items, this would have required 3.78×10^{22} multiple regressions to be computed at each time point, which was computationally excessive. Fortunately, permutation tests still provide precise p value estimates when as few as 10,000 permutations are considered (Blair & Karniski, 1993). We

therefore estimated multiple regression R^2 and t p-values from distributions of R^2 and t formed over 20,000 permutations at each of the 50 time points under consideration.

Results

We began by conducting the orthographic multiple regression every 10 ms over the middle occipital channel and the semantic multiple regression every 10 ms over the middle parietal channel. The middle occipital channel was selected for the orthographic regression because, from the reduced electrode array available, it was the channel closest to those at which early effects of orthographic variables have most often been observed (see Table 1); some early frequency and semantic effects have been reported in this region as well (Tables 2 and 3). The middle parietal channel was selected for the semantic regression because it is where semantic effects on the N400 are typically largest, as well as being one channel where early effects of frequency and semantic variables have been observed in the past (Tables 2 and 3). In these and all subsequent regressions, a clustering criterion is applied such that a predictor is not considered reliable unless its t -statistic has a permutation p -value of $< .01$ for at least two consecutive 10 ms bins. To limit the number of necessary statistical tests, and because we were concerned mainly here with when information first becomes available, we will consider only the first 500 ms of processing in the results that follow. R -squared and t -statistics for *all* reliable tests, along with associated permutation test p -values are available for review in Appendix 1.

The orthographic multiple regression was first reliable from 130–150 ms post stimulus onset. The model was then again reliable from 180–210 ms. During this second phase, bigram and neighbor frequency were both continuously reliable. The orthographic regression was again reliable for a third phase, from 230–470 ms. During this third phase, bigram frequency was a reliable predictor from 240–440 ms, and orthographic neighborhood from 310 to 380 ms. Note that these three temporal windows correspond roughly to the windows of the N/P150, N250, and N400 components. Figure 2 displays the time course of R -squared for the full orthographic model, and Figure 3 displays the time course of t -statistics for each individual predictor in the orthographic model. Figure 5 overlays the windows of reliability for each of the regression models on grand averaged ERPs elicited by words, pseudowords, acronyms, and illegal strings at the middle parietal and middle occipital electrode site.

The semantic multiple regression, in contrast, was not first reliable until 300 ms post stimulus onset, and stayed reliable until 340 ms post stimulus onset. During the 300–340 ms post stimulus onset epoch, by far the strongest predictor of waveform amplitude among those included in the semantic regression was number of lexical associates. The individual t -statistic for number of lexical associates was reliable throughout this window. The other individual t -statistics were not reliable in this window. Figure 2 displays the time course of R -squared for the full semantic model, while Figure 4 displays the time course of t -statistics for individual predictors in the semantic model. Figure 5 overlays the windows of reliability for each of the regression models on grand averaged ERPs elicited by words, pseudowords, acronyms, and illegal strings at the middle parietal and middle occipital electrode site. One important point that can be understood on the basis of Figure 5 is that although the N400

peaks around 400 ms post stimulus onset, effects of semantics are clearly visible before then—in this case by 300 ms.

Because the orthographic regression and semantic regression were reliable in overlapping epochs, we were interested in determining whether the results of the semantic regression would be different if we employed a hierarchical regression procedure in which the orthographic regression was first performed on data from words only, the residuals computed, and the semantic regression run on those residuals only. That is, we wondered if the results of the semantic regression would change if variance explained by orthographic factors were already accounted for. With this approach, the semantic regression was reliable from 310–350 ms, and, again, number of lexical associates was the only individually reliable predictor of waveform amplitude in this epoch. That is, results of the semantic regression were essentially unchanged in the hierarchical procedure, suggesting that the semantic regression is truly revealing effects of semantics on the waveform, not correlated effects of orthography.⁴

Effects of lexical frequency have played a particularly important role in the literature examining the timecourse of word processing. However, as discussed, the nature of lexical frequency effects, when observed, can be controversial. In particular, it is not clear that it would be appropriate to classify lexical frequency as either a semantic or an orthographic variable (see especially Simon, Lewis, & Marantz, 2012). Because of this, and the fact that lexical frequency was suggested as a separate factor in the Factor Analysis, we conducted a focused analysis of this variable alone. First, we conducted a regression of lexical frequency alone over the middle occipital channel—the channel in our data set most similar to where early lexical frequency effects have been reported in the past, and the channel for which our pilot study with 64 electrodes revealed especially strong effects of lexical frequency in this task and item set. Second, we conducted a regression of bigram frequency over the middle occipital channel, and then conducted a regression of lexical frequency on the residuals from the bigram frequency analysis. The single-step analysis revealed that lexical frequency was a reliable predictor of waveform amplitude from 270–360 ms post stimulus onset. When bigram frequency was first accounted for in the hierarchical procedure, lexical frequency was a reliable predictor of waveform amplitude from 280–350, suggesting that a small portion of the lexical-frequency-only effect was driven by bigram frequency.

Finally, we were interested in illustrating what happens when less conservative statistical tests than those employed thus far are used. We therefore conducted further analyses using approaches similar to those used in past work advocating for early lexical access. That is, we conducted the semantic multiple regression over all 6 scalp electrodes, and looked for the earliest effect of semantics, at any electrode site, with no clustering criterion and an uncorrected alpha of .05. This analysis indicated that the first epoch in which the semantic regression was reliable was 120–160 ms post stimulus onset, over the middle occipital channel—quite near to the window and scalp site in which early effects of semantic variables have been reported in past studies (160 ms, Hauk et al., 2006).

⁴Similarly, the time course of the semantic model, and the prominence of the lexical association effect, are both retained if lexical frequency is simultaneously entered in to the semantic model.

Second, we conducted t tests on waveform mean amplitude for lexical items (words and acronyms) and nonlexical items (pseudowords and illegal strings) every 10 ms over the middle parietal channel, where early effects of lexicality have been reported in the past on the basis of t-maps (Sereno et al., 1998). Again, we did not adjust for multiple comparisons. In this analysis, the first reliable effect of lexicality was observed at 50 ms post stimulus onset, even earlier than that observed by Sereno et al. (1998) and at a time that all models would consider to be prelexical (visual) in nature.

Discussion

We set out to examine the functional time course of out-of-context visual word recognition, as evidenced in electrophysiology, with a novel, data-driven, items multiple-regression analysis approach. In order to characterize the time windows during which orthographic and semantic processing were evidenced in the ERPs, we conducted multiple regressions with both orthographic and semantic predictor variables in consecutive 10 ms bins from 0 to 500 ms post stimulus onset. In doing so, we aimed to discover whether the BIAM would be supported with a paradigm that did not use masked priming and whether we would be able to replicate reports of early effects of lexical frequency and/or semantics with a combination of a very large data set and powerful analytic techniques.

The orthographic model first became reliable from 130–150 ms, in the latency of the N/P150 component. This is consistent with the BIAM in suggesting that during this time window, processing reflects the formation of orthographic features (Grainger & Holcomb, 2009). The orthographic model was again reliable from 180 to 210 ms. Bigram frequency and neighbor frequency were both reliable predictors throughout this epoch. To our knowledge, these effects are novel in that bigram frequency effects *per se* (i.e., as opposed to effects of lexical or syllabic frequency) have not been reported in this time window of the ERP, and neighbor frequency effects on the ERP have only been reported once previously, by us (Laszlo & Federmeier, 2011), in a study focusing on the N400 component and therefore not examining potential early effects. However, lexical frequency effects have been observed throughout this time window in numerous studies (e.g., Hauk & Pulvermuller, 2004; Hauk et al., 2006; Rugg, 1990). Our interpretation of these early effects of bigram and neighbor frequency is essentially the same as that typically offered for lexical frequency effects: namely, that bigram frequency and neighbor frequency each are measures of how often a particular (sub-lexical) series of letters have been previously encountered and therefore have early and long-lasting effects on the processing of visual input.

The orthographic regression became reliable for a third and final time between 230 and 470 ms. During this time, bigram frequency was a reliable predictor from 240 to 440 ms and neighborhood size was a reliable predictor from 310 to 380 ms. This result is consistent with our prior finding that orthographic neighborhood size is an especially strong predictor of N400 amplitude (Laszlo & Federmeier, 2011), which we interpret as resulting from the fact that items with more orthographic neighbors initially activate not only their own semantics, but also the semantics of their neighbors. Supporting this, Laszlo and Plaut (2012) have shown in an implemented computational model of the N400 that items with more

orthographic neighbors produce more activation at the semantic level of representation than their counterparts with fewer neighbors.

We observed effects of whole item frequency from 270–360 ms post stimulus onset. Whole item frequency has been important in the visual word recognition literature, as effects of frequency have often been taken as evidence that lexical access has occurred (Serenio et al., 1998; Serenio & Rayner, 2003). However, this assumption is controversial, given that frequency can have effects at multiple processing levels, from perception to semantics (see Embick, Hackl, Schaeffer, Kelepir, & Marantz, 2001; Simon et al., 2012). Because whole item frequency effects cannot be neatly categorized as “orthographic” or “semantic” in nature, we assessed frequency separately from our purely orthographic or purely semantic regression models; this decision was supported by a factor analysis of our predictor variables. We found that the earliest effects of whole item frequency (beginning at 270 ms) lagged the earliest effects of bigram frequency (beginning at 180 ms) by 90 ms. Within another 30 ms, then, beginning at 300 ms, we began to observe effects of whole-item semantic variables, especially number of lexical associates. This pattern of effects suggests a transition from processing items at the sub-item level (as revealed by bigram and neighbor frequency effects), to processing “lower” level characteristics of whole items (lexical frequency), to processing “higher” level characteristics (e.g., lexical associations). Additionally, this timecourse was essentially consistent for tests performed simultaneously and hierarchically, with a frequency model applied to the residuals of a bigram frequency model indicating effects of whole item frequency independent of effects of bigram frequency occurring between 280 and 350 ms.

The semantic model was first reliable from 300–340 ms, which falls squarely within the traditional range of the N400 component (e.g., 250–450 ms). This period is slightly abbreviated when compared to the traditional N400 window, although few prior studies have explicitly examined the fine-grained time course of semantic effects within the larger window in a data-driven (as opposed to component-centered) manner. In the present study, target detection responses (P3b component; see Polich, 2007) to the proper names peaked around 400 ms; it may be that once participants were able to classify the words as non-targets, they did not engage in much additional semantic processing. During the range of significance of the semantic model, only number of lexical associates was a reliable individual predictor. The primacy of number of lexical associates as a predictor of N400 amplitude is consistent with our prior work, in which we have demonstrated that, at least for words out of context, variables that locate an item within the larger lexico-semantic network tend to be stronger predictors of N400 amplitude than variables that pertain only to that item (Laszlo & Federmeier, 2011). In fact, we have shown that number of lexical associates is a variable that predicts N400 amplitude even in the face of item repetition, a manipulation that, for example, reduces or eliminates effects of concreteness and imageability (e.g., Kounios & Holcomb, 1994). Although we did not find reliable effects of all of the individual semantic predictors (i.e., imageability and concreteness) that have previously been shown to affect the N400, it is important to note that our study is the first to consider all of these variables simultaneously. Thus, it could certainly be the case that inter-correlations between variables in past studies have inflated (or deflated) the apparent importance of any of the individual predictors. Moreover, these variables have been shown to be sensitive to task

demands (e.g., imageability effects are larger in tasks that encourage imagery; West and Holcomb, 2000), which may have reduced their contribution in the current experiment.

Overall, our time course results are remarkably similar to the predictions of the BIAM. That is, we observed sub-lexical effects in the range of the N/P150, relatively “low” level effects of properties of the full lexical item in the range of the N250, and semantic effects in the range of the N400. This correspondence is remarkable for at least two reasons. First, we obtained these results in a “data-driven” manner. That is, we did not pick out the N/P150, N250, and N400 range based on, as is often done, visual characteristics of the waveform, and then look for each effect in each *a priori* window; instead, we let the data guide us to ranges of significance. Second, we used an experimental paradigm, consisting of the presentation of single words in the clear, that differs in important ways from the masked priming studies that have provided core data for the BIAM (for review, see Grainger & Holcomb, 2009). Thus, our results support a generalization of the theoretical framework of the BIAM to a novel analytic and empirical context.

Our results do not support the hypothesis, as suggested by models such as EZ Reader, that lexical access is completed within the first 200 ms of stimulus processing. First, we did not find effects of semantic variables prior to 200 ms, despite a larger data set and more powerful analytic methods than have been used in most prior studies. Second, we observed ongoing processing of nonwords subsequent to 200 ms, which would seem inconsistent with the proposal that unique lexical identification has occurred by that time. Thus, both our positive (ongoing processing of nonwords) and negative (no early effects of semantics) results seem just as inconsistent with models postulating rapid lexical access as they are consistent with the BIAM.

Early semantic ERP effects?

A question raised by our result pattern is why prior studies have reported early effects of semantics that did not manifest here. We suggest that two factors may be important for understanding this discrepancy: (1) what statistical methods were employed and (2) what task participants performed.

It is notable that many previous ERP studies that report early lexico-semantic effects do not describe any correction for multiple comparisons (although, of course, it is possible that they were applied but not described; Hauk & Pulvermuller, 2004; Hauk et al., 2006; Hauk et al., 2009; Sereno et al., 1998). This is an important omission, as the multiple consecutive tests that are required to identify when a variable of interest is first reliable can result in considerable inflation of false positive rate. For example, even with only 8 consecutive tests (as in Sereno et al., 1998), the chances of a false positive at an alpha of .05 is 40%. This problem is exacerbated in these studies by the fact that electrode arrays of greater than 30 scalp channels were used, with testing done on multiple electrode sites/groups; other authors have pointed out that analysis of this type results in an effective false positive rate of 1 (Amsel, 2011). Especially when the influence of a given lexical variable is then found at one electrode site, determined post hoc rather than based on motivated, *a priori* regions of interest, the risk of a false positive is even further increased. For example, no motivation is given for selection of the left occipital electrode used in Sereno et al. (1998) to suggest that

lexical frequency effects are observed prior to 200 ms post stimulus onset. Each of these issues decreases the statistical robustness of the data taken in support of early lexical access, and, as these data are less well-replicated than reports of later effects of lexical semantics, they must be sought with especial statistical care. Here, we avoided these statistical issues in part by limiting our analysis to a small number of channels determined *a priori* by a pilot study with the same materials and in part by application of 1) a clustering criterion and 2) a reduced alpha value.

We investigated the importance of correcting for multiple comparisons in our own data set and observed that if multiple regressions are conducted every 10 ms, the first observable effect of semantic variables on the waveforms at an alpha of .05 can indeed be found considerably sooner than the 300 ms post stimulus onset suggested by more conservative analyses. In fact, with more liberal statistical tests, we were able to observe “reliable” effects of semantic variables as early as 120 ms post stimulus onset. One clear message of this investigation is thus that differences in estimated time courses of visual word recognition advocated by past ERP work may in part be due to differences in the statistical techniques employed, with uncorrected tests tending to indicate an earlier influence of semantics than that indicated by more conservative tests such as those employed here. This conclusion is further supported by previous work also using the permutation technique that similarly observed pre-N400 effects of lexical-semantic variables, but only prior to correction for multiple comparisons (Dimigen et al., 2011).

A more convincing case for early effects of semantics comes from a pair of studies (Hauk et al., 2012; Amsel et al., 2013) using the Go/Nogo task paradigm and showing differentiation of words depicting living and nonliving things on the N200, a component that constitutes part of the normal ERP response in the context of Go/Nogo tasks and that has been linked to response inhibition and conflict monitoring. The N200 is not part of the normal response to words during other types of language tasks, and, in the context of Go/Nogo tasks, is not specific to word processing; indeed, early responses to semantic properties of scenes (VanRullen and Thorpe, 2001) have also been seen on N200s in Go/Nogo paradigms. The Go/Nogo paradigm brings with it particular task demands and constraints: participants are explicitly directed to pay attention to only certain aspects of the stimuli, they can develop strong expectancies about what type of stimuli will be presented, and they need to respond quickly but based only on whatever type and degree of information is sufficient to sort the items into binary categories. It is therefore possible that fast -- but coarse -- feed forward information about perceptual and/or more abstract features of stimuli may be made available to the motor system and harnessed in these tasks. The extent to which similar processes are elicited in the context of other tasks and, if so, whether they have functional consequences for normal language processing remains unclear. However, such data point to the possibility that the effective timecourse of word processing may importantly vary with factors such as task -- or, as we discuss next, the availability of context.

Early effects in eyetracking studies

Although procedural differences (either in task or data analysis) may explain discrepancies in findings within the ERP literature, they do not explain the apparent differences in the

timecourse of word processing suggested by the BIAM and the current ERP study on the one hand and by studies using eye tracking during natural reading on the other. This disconnect has been even more troubling to the literature than that amongst ERP studies, because the eye-tracking studies that support (for example) E-Z Reader are extremely well-replicated—much more so than ERP studies arguing for early semantics. To reconcile this disconnect, we suggest that it is critical to consider the role of context (including parafoveal preview) in word processing.

Theories of language comprehension vary considerably in the degree and nature of the influence that they propose accrues from context onto the identification of individual words, with important concomitant consequences for understanding the effect of context on the time course of visual word recognition. In one longstanding tradition within psycholinguistics, instantiated in explicit models of visual word recognition that are mostly or entirely bottom-up (e.g., Forster, 1989; Gaskell & Marslen-Wilson, 1997), context has been viewed as having little or no influence on basic word identification processes. Under such a view, the timecourse and nature of early stages of processing for words should be essentially the same whether those words are encountered out of context or in the context of a sentence or larger text, and thus there must be a methodological difference between the ERP and the eyetracking approaches that is responsible for their divergent timecourse estimates. Alternatively, the time course of subprocesses of word recognition could actually differ in and out of context, contrary to the traditional view that word recognition is an entirely bottom-up process. We propose that both of these factors may contribute to the discrepancy between time courses estimated from eye tracking and ERP studies.

One especially important aspect of eye tracking methodology that could contribute to the emergence of what have been taken to be earlier effects of lexico-semantic processing is that participants read relatively naturally, with full preview of the text to come. ERPs, in contrast, are typically collected while participants read single words; this is done because eye movements produce large electrical deviations that can obscure the brain signals of interest. Thus, participants reading in an eye tracking study have significantly more information available to them about upcoming words than that available to participants in a typical ERP reading study. Indeed, when ERPs have been recorded in paradigms where in parafoveal preview is available, it has been demonstrated that fairly high level information—such as semantic congruency with sentence context—can be extracted from previewed words (Barber, et al., 2010; Barber, Ben-Zvi, Bentin, & Kutas, 2011). In terms of timing in particular, some have suggested that processing of a previewed word can be sped by as much as 50 ms (Pollatsek et al., 2006); if this is added to the ~225–250 ms first fixation duration suggested as an upper bound on lexical processing time by E-Z Reader, an estimate of how long lexical processing might take when there is no preview is now approximately 300 milliseconds—precisely when we first observed semantic effects in the present study and well within the range of the N400, which tends to peak ~400 ms but is evident 100 to 150 ms prior to its peak (See Figure 5).

However, the effect of context information on processing extends beyond just the possibility of preview; context can provide information about likely upcoming words prior to their apprehension. ERP studies, for example, have been instrumental in providing evidence for

context-based predictive processing in language comprehension, and, in particular, in showing that such prediction facilitates the processing of incoming words at semantic (Federmeier & Kutas, 1999), morphological (DeLong, Urbach, & Kutas, 2005), and even orthographic (Laszlo & Federmeier, 2009) levels of representation. Correspondingly, ERP effects of expectancy for words in context can be seen earlier in time. For example, P2 responses (beginning around 200 ms) to expected words in constraining sentence frames differ from those to words that are less expected because they occur in less constraining sentence frames (e.g., Federmeier, Mai, & Kutas, 2005). Importantly, this type of effect is clearly not bottom-up: unpredictable words in highly constraining contexts show the same enhanced P2 as expected words in those contexts (Wlotko & Federmeier, 2007). Thus, the P2 enhancement effect, while elicited in response to target words, reflects processing differences that arise as a function of the predictive strength of the context in which those words are embedded⁵.

Similarly, early effects (peaking as early as 100 ms post stimulus onset, often referred to as the “ELAN”; Gunter, Friederici, & Hahne, 2002) have been reported for words whose physical properties (e.g., presence or absence of morphological markings) either match or mismatch strong expectations established by the prior context. Although the nature of these effects remains controversial, growing evidence suggests that they arise at perceptual levels of processing. For example, in the MEG literature, what has been taken to be the magnetic equivalent of the ELAN, the M100 component, is known to be generated in visual cortex (Dikker, Rabagliati, & Pykkänen, 2009). Thus, when context affords predictions, features of words that are not themselves “semantic” in nature can nevertheless provide information relevant to determining the (likely) semantics of a word and making appropriate behavioral responses (such as moving one’s eyes). In other words, the brain’s response to a word is flexible, multi-faceted, and context-driven, with information being continuously derived at multiple levels of representation over time, and with the utility of that information varying depending on the comprehender’s goals and the circumstances under which a word is encountered.

Summary and Conclusions

There has been a sense in the literature that “word recognition” is a unitary process that unfolds with the same timecourse in all possible contexts and eliciting conditions, and that it is possible (even desirable) to identify that invariant time at which a lexical item can be uniquely specified (the ‘magic moment’ as it is referred to by Balota & Yap, 2006). However, as we have reviewed, there are contextual, methodological, and theoretical reasons that the observed time course of word recognition might vary from one situation to another, that, when taken together, seem to diffuse the discrepancies in time course suggested across task and eliciting method.

In some sense, this should not be surprising, because the brain’s goal when encountering a word is generally not to invariantly link it to specific lexical, syntactic, or meaning-related

⁵Eyetracking studies showing early effects of expectancy are generally comparing words in highly versus more weakly constraining contexts. Thus, in these studies, it is not possible to determine if effects on gaze are driven by the fit of the target word itself or, as in the ERP data, by the constraint of the prior context.

information. Instead, the optimal strategy may be better conceptualized as flexibly adopting a range of approaches in order to understand the message level information being conveyed (Balota & Yap, 2006; Christianson, Williams, Zacks, & Ferreira, 2006). The use of a given word in text may range from skipping it completely (as occurs for predictable words in natural reading), to checking merely that low-level properties of the input are consistent with expectations that have been built from various sources of context, to richly processing its features and using that information to create, update, or revise the ongoing message-level representation. These considerations apply even to the processing of words *out* of context. As our own work has demonstrated, the extent of high level (e.g., semantic) processing attempted for a string of letters even in an unconnected stream of text is strongly dependent on the nature of the substantive task being performed (Laszlo, Stites, & Federmeier, 2012). This is a concept that has been embraced in the behavioral literature (e.g., Norris, 2006; Balota & Yap, 2006) but has been missing from the discourse attempting to reconcile divergent time courses suggested from the many ERP and eye tracking studies of visual word recognition.

Thus, asking “what is **THE** time course of visual word recognition?” might not be a well-formed question. Instead, we submit that to understand the cognitive and neural mechanisms involved in comprehension, it is important to assess and compare what the brain can do, and how fast it can do it, when it encounters letter strings under a range of different circumstances. In doing so, we may discover that some aspects of processing indeed have important temporal constraints. Our data, obtained with a novel analytical approach, suggest a protracted time course for the extraction of increasingly high level information from out-of context words. The data accord with other many other ERP studies of word recognition, as summarized in the BIAM model, as well as with data about the timecourse of recognition for other types of visual objects. Notably, the timing of the peak of the N400 component, which has been postulated to reflect a binding process that links the current input to associated information in long-term memory (Federmeier & Laszlo, 2009), is remarkably stable across task circumstances, including word processing in and out of context, as well as face and visual object processing. This suggests that, although the information available to be bound may differ across contexts and tasks, time may play an important role in the binding process itself. In contrast, other brain responses linked to language comprehension, such as the P600, do not seem subject to the same kind of temporal constraints (e.g., Gouvea, Phillips, Kazanina, & Poeppel, 2010). Exploring these kind of differences promises to help us build a fuller understanding of how the brain processes language inputs under different circumstances in order to achieve the flexible, situation dependent goals it must achieve across the range of comprehension scenarios it encounters.

Acknowledgments

The authors acknowledge B. C. Armstrong, E. Wlotko, D. Groppe, and D.C. Plaut for their insightful discussion of the statistical issues involved in analyzing the single-item ERP corpus. Numerous research assistants are to be thanked for their part in the collection of the single-item data, especially P. Anaya, H. Buller, and C. Laguna. This research was supported by NIMH T32 MH019983, NICHD F32HD062043, NIA 5R01 AG026308, the James S. McDonnell Foundation 21st Century Science Initiative Scholar Award in Understanding Human Cognition, NSF-CAREER BCS-1252975, and the Research Foundation of the State University of New York.

References

- Amsel BD. Tracking real-time neural activation of conceptual knowledge using single-trial event-related potentials. *Neuropsychologia*. 2011; 49:970–983. [PubMed: 21219919]
- Amsel BD, Urbach TP, Kutas M. Alive and grasping: Stable and rapid semantic access to an object category but not object graspability. *NeuroImage*. 2013
- Altmann GTM. Language can mediate eye movement control within 100 milliseconds, regardless of whether there is anything to move the eyes to. *Acta Psychologica*. 2010; 137:190–200.
- Anderson MJ. Permutation tests for univariate or multivariate analysis of variance and regression. *Canadian Journal of Fisheries and Aquatic Sciences*. 2001; 58:626–639.
- Balota, DA. The role of meaning in word recognition. In: Balota, DA.; Flores D'Arcais, GB.; Rayner, K., editors. *Comprehension processes in reading*. Erlbaum; 1990.
- Balota DA, Cortese MJ, Sergent-Marshall SD, Spieler DH, Yap MJ. Visual Word Recognition of Single-Syllable Words. *Journal of Experimental Psychology: General*. 2004; 133:283–316. [PubMed: 15149254]
- Balota DA, Yap MJ. 10 Attentional control and the flexible lexical processor: Explorations of the magic moment of word recognition. *From inkmarks to ideas: Current issues in lexical processing*. 2006:229.
- Barber HA, Doñamayor N, Kutas M, Münte T. Parafoveal N400 effect during sentence reading. *Neuroscience letters*. 2010; 479(2):152–156. [PubMed: 20580772]
- Barber HA, Ben-Zvi S, Bentin S, Kutas M. Parafoveal perception during sentence reading? An ERP paradigm using rapid serial visual presentation (RSVP) with flankers. *Psychophysiology*. 2011; 48(4):523–531. [PubMed: 21361965]
- Barrett SE, Rugg MD. Event-related potentials and the semantic matching of pictures. *Brain and Cognition*. 1990; 14:201–212. [PubMed: 2285513]
- Blair RC, Karniski W. An alternative method for significance testing of waveform difference potentials. *Psychophysiology*. 1993; 30:518–524. [PubMed: 8416078]
- Carreiras M, Andoni, Dunabeitia J, Molinaro N. Consonants and Vowels Contribute Differently to Visual Word Recognition: ERPs of Relative Position Priming. *Cerebral Cortex*. 2009; 19:2659–2670. [PubMed: 19273459]
- Chauncey K, Holcomb PJ, Grainger J. Effects of stimulus font and size on masked repetition priming: An ERP investigation. *Language and Cognitive Processes*. 2008; 23:183–200. [PubMed: 19590754]
- Christianson K, Williams CC, Zacks RT, Ferreira F. Younger and older adults' "good-enough" interpretations of garden-path sentences. *Discourse Processes*. 2006; 42:205–238. [PubMed: 17203135]
- Clark VP, Fan S, Hillyard SA. Identification of Early Visual Evoked Potential Generators by Retinotopic and Topographic Analyses. *Human Brain Mapping*. 1995; 2:170–187.
- Coltheart, M.; Davelaar, E.; Jonasson, J.; Besner, D. Access to the internal lexicon. In: Dornic, S., editor. *Attention & performance IV*. Hillsdale, NJ: Erlbaum; 1977. p. 535-555.
- Coltheart M. The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*. 1981; 33A:497–505.
- Dambacher M, Kliegl R, Hofman M, Jacobs AM. Frequency and predictability effects on event-related potentials during reading. *Brain Research*. 2006; 1084:89–103. [PubMed: 16545344]
- Dambacher M, Kliegl R. Synchronizing timelines: Relations between fixation durations and N400 amplitudes during sentence reading. *Brain Research*. 2007; 1155:147–162. [PubMed: 17499223]
- Dambacher M, Dİmigen O, Braun M, Wille K, Jacobs AM, Kliegl R. Stimulus onset asynchrony and the timeline of word recognition: Event-related potentials during sentence reading. *Neuropsychologia*. 2012; 50:1852–1870. [PubMed: 22564485]
- Dehaene S, Cohen L, Sigman M, Vinckier F. The neural code for written words: a proposal. *TRENDS in Cognitive Sciences*. 2005; 9:335–341. [PubMed: 15951224]
- Delong KA, Urbach TP, Kutas M. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*. 2005; 8:1117–1121.

- Dikker S, Rabagliati H, Pylkkanen L. Sensitivity to syntax in visual cortex. *Cognition*. 2009; 110:293–321. [PubMed: 19121826]
- Dimigen O, Sommer W, Hohlfield A, Jacobs AM, Kliegl R. Coregistration of Eye Movements and EEG in Natural Reading: Analyses and Review. *Journal of Experimental Psychology: General*. 2011; 140:552–572. [PubMed: 21744985]
- Dimigen O, Kliegl R, Sommer W. Trans-saccadic parafoveal preview benefits in fluent reading: A study with fixation-related brain potentials. *NeuroImage*. 2012; 62:381–393. [PubMed: 22521255]
- Di Russo F, Martínez A, Sereno MI, Pitzalis S, Hillyard SA. Cortical Sources of the Early Components of the Visual Evoked Potential. *Human Brain Mapping*. 2001; 15:95–111. [PubMed: 11835601]
- Dufau S, Grainger J, Holcomb PJ. An ERP investigation of location invariance in masked repetition priming. *Cognitive, Affective, and Behavioral Neuroscience*. 2008; 8:222–228.
- Embick D, Hackl M, Schaeffer J, Kelepir M, Marantz A. A magnetoencephalo-graphic component whose latency reflects lexical frequency. *Cognitive Brain Research*. 2001; 10:345–348. [PubMed: 11167059]
- Engbert R, Longtin A, Kliegl R. A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*. 2002; 42:621–636. [PubMed: 11853779]
- Fabre-Thorpe M, Delorme A, Marlot C, Thorpe S. A Limit to the Speed of Processing in Ultra-Rapid Visual Categorization of Novel Natural Scenes. *Journal of Cognitive Neuroscience*. 2001; 13:171–180. [PubMed: 11244543]
- 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.
- Federmeier KD, Mai H, Kutas M. Both sides get the point: Bihemispheric sensitivity to sentential constraint. *Memory and Cognition*. 2005; 33:871–886. [PubMed: 16383175]
- Federmeier, KD.; Kutas, M.; Dickson, DS. A common neural progression to meaning in about a third of a second. In: Hickok, GS.; Small, SL., editors. *Neurobiology of Language*. Holland: Elsevier; in press
- Forster KI, Davis C. Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1984; 10:680–698.
- Forster, KI. Basic issues in lexical processing. In: Marslen-Wilson, WD., editor. *Lexical representation and process*. Cambridge, MA: MIT Press; 1989. p. 75-107.
- Freedman D, Lane D. A Nonstochastic Interpretation of Reported Significance Levels. *Journal of Business & Economic Statistics*. 1983; 1:292–298.
- Ganis G, Kutas M, Sereno MI. The search for “common sense”: An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience*. 1996; 8(2): 89–106. [PubMed: 23971417]
- Ganis G, Kutas M. An electrophysiological study of scene effects on object identification. *Cognitive Brain Research*. 2003; 16:123–144. [PubMed: 12668221]
- Gaskell MG, Marslen-Wilson WD. Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*. 1997; 12:613–656.
- Gouvea AC, Phillips C, Kazanina N, Poeppel D. The linguistic processes underlying the P600. *Language and Cognitive Processes*. 2010; 25:149–188.
- 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]
- Hagoort, P.; Baggio, G.; Willems, RM. Semantic Unification. In: Gazzaniga, M., editor. *The Cognitive Neurosciences*. 4. Boston: MIT Press; 2009. p. 819-836.
- Hauk O, Pulvermüller F. Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology*. 2004; 115:1090–1103. [PubMed: 15066535]
- Hauk O, Davis MH, Ford M, Pulvermüller F, Marslen-Wilson WD. The time course of visual word recognition as revealed by linear regression analysis of ERP data. *NeuroImage*. 2006; 30:1383–1400. [PubMed: 16460964]

- Hauk O, Pulvermuller F, Ford M, Marslen-Wilson WD, Davis MH. Can I have a quick word? Early electrophysiological manifestations of psycholinguistic processes revealed by event-related regression analysis of the EEG. *Biological Psychology*. 2009; 80:64–74. [PubMed: 18565639]
- Hauk O, Coutout C, Holden A, Chen Y. The time-course of single-word reading: Evidence from fast behavioral and brain responses. *NeuroImage*. 2012; 60:1462–1477. [PubMed: 22281671]
- Holcomb PJ, Kounios J, Anderson JE, West WC. Dual-coding, context-availability, and concreteness effects in sentence comprehension: an electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1999; 25:721–742.
- Holcomb PJ, Grainger J. On the Time Course of Visual Word Recognition: An Event-Related Potential Investigation using Masked Repetition Priming. *Journal of Cognitive Neuroscience*. 2006; 18:1631–1643. [PubMed: 17014368]
- Kliegl R, Dimigen O, Jacobs AM, Sommer. Eye movements and brain electric potentials during reading. *Psychological Research*. 2012; 76:145–158. [PubMed: 21915693]
- Kiyonaga K, Midgley KJ, Holcomb PJ, Grainger J. Masked cross-modal repetition priming: An ERP investigation. *Language and Cognitive Processes*. 2007; 22:337–376. [PubMed: 18163153]
- Kounios J, Holcomb PJ. Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1994; 20(4):804.
- 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.
- Laszlo S, Federmeier KD. Minding the PS, queues, and PXQs: Uniformity of semantic processing across multiple stimulus types. *Psychophysiology*. 2008; 45:458–466. [PubMed: 18221447]
- Laszlo S, Federmeier KD. A beautiful day in the neighborhood: An event-related potential study of lexical relationships and prediction in context. *Journal of Memory and Language*. 2009; 61:326–338. [PubMed: 20161064]
- Laszlo S. What goes up, must come down: Multiple regression analysis of single-item ERPs reveals yoked oscillation of orthographic and semantic processing. *Psychophysiology*. 2011; 48(Supplement 1):S43.
- 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.; Plaut, DC. Simulating Event-Related Potential Reading Data in a Neurally Plausible Parallel Distributed Processing Model. *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, Mahwah, NJ: Lawrence Erlbaum Associates; 2011.
- Laszlo S, Plaut DC. A neurally plausible parallel distributed processing model of event-related potential word reading data. *Brain and Language*. 2012; 120:271–281. [PubMed: 21945392]
- Laszlo S, Stites M, Federmeier KD. Won't Get Fooled Again: An Event-Related Potential Study of Task and Repetition Effects on the Semantic Processing of Items without Semantics. *Language and Cognitive Processes*. 2012; 27:257–274. [PubMed: 22518068]
- Luck SJ, Chelazzi L, Hillyard SA, Desimone R. Neural Mechanisms of Spatial Selective Attention in Areas V1, V2, and V4 of Macaque Visual Cortex. *Journal of Neurophysiology*. 1997; 77:24–42. [PubMed: 9120566]
- Marcus M, Santorini B, Marcinkiewicz M. Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*. 1993; 19:313–330.
- Medler, DA.; Binder, JR. MCWord: An on-line orthographic database of the English language. 2005. Retrieved from <http://www.neuro.mcw.edu/mcword/>
- Nelson, DL.; McEvoy, CL.; Schreiber, TA. The University of South Florida word association, rhyme, and word fragment norms. 1998. Retrieved from: <http://usf.edu/FreeAssociation/>
- Nobre AC, Allison T, McCarthy G. Word recognition in the human inferior temporal lobe. *Nature*. 1994; 372:260–263. [PubMed: 7969469]
- Norris D. The Bayesian Reader: Explaining Word Recognition as an Optimal Bayesian Decision Process. *Psychological Review*. 2006; 113:327–357. [PubMed: 16637764]

- Paller KA, Gonsalves B, Grabowecky M, Bozic VS, Yamada S. Electrophysiological correlates of recollecting faces of known and unknown individuals. *Neuroimage*. 2000; 11(2):98–110. [PubMed: 10679183]
- Polich J. Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology*. 2007; 118(10):2128. [PubMed: 17573239]
- Pollatsek, A.; Rayner, K. Eye movements and lexical access in reading. In: Balota, DA.; Flores d'Arcais, GB.; Rayner, K., editors. *Comprehension processes in reading*. Hillsdale, NJ: Erlbaum; 1990. p. 143-163.
- Pollatsek A, Reichle ED, Rayner K. Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*. 2006; 52:1–56. [PubMed: 16289074]
- Princeton University. About WordNet. 2010. Retrieved from <http://wordnet.princeton.edu>
- Pulvermuller F. Words in the brain's language. *Behavioral and Brain Sciences*. 1999; 22:253–336. [PubMed: 11301524]
- Pylkkanen L, Marantz A. Tracking the time course of word recognition with MEG. *Trends in Cognitive Sciences*. 2003; 7:187–189. [PubMed: 12757816]
- Rayner K, Duffy SA. Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*. 1986; 14:191–201. [PubMed: 3736392]
- Rayner K. Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*. 2009; 62:1457–1506. [PubMed: 19449261]
- Rayner K, Clifton C Jr. Language Processing in Reading and Speech Perception is Fast and Incremental: Implications for Event-Related Potential Research. *Biological Psychology*. 2009; 80:4–9. [PubMed: 18565638]
- Reichle E, Pollatsek A, Fisher DL, Rayner K. Toward a model of eye movement control in reading. *Psychological Review*. 1998; 105:125–157. [PubMed: 9450374]
- Reichle E, Tokowicz N, Liu Y, Perfetti CA. Testing an assumption of the E-Z Reader model of eye-movement control during reading: Using event-related potentials to examine the familiarity check. *Psychophysiology*. 2011; 48:993–1003. [PubMed: 21261631]
- Reichle E, Rayner K, Pollatsek A. Eye movements in reading versus nonreading tasks: Using E-Z Reader to understand the role of word/stimulus familiarity. *Visual Cognition*. 2012; 20:360–390. [PubMed: 22707910]
- Reichle E, Rayner K, Pollatsek A. Using E-Z Reader to Simulate Eye Movements in Nonreading Tasks: A Unified Framework for Understanding the Eye-Mind Link. *Psychological Review*. 2012b; 119:155–185. [PubMed: 22229492]
- Rohde DLT, Gonnerman L, Plaut DC. An improved model of semantic similarity based on lexical co-occurrence. *Cognitive Science*. Under Review.
- Rossion B, Jacques C. Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *NeuroImage*. 2008; 39:1959–1979. [PubMed: 18055223]
- Rossion B, Caharel S. ERP evidence for the speed of face categorization in the human brain: Disentangling the contribution of low-level visual cues from face perception. *Vision Research*. 2011; 51:1297–1311. [PubMed: 21549144]
- Rossion B, Jacques C. The N170: understanding the time-course of face perception in the human brain. *The Oxford handbook of ERP components*. 2011:115–142.
- Rugg MD. Event-related brain potentials dissociate repetition effects of high- and low-frequency words. *Memory and Cognition*. 1990; 18:367–379. [PubMed: 2381316]
- Schendan HE, Ganis G, Kutas M. Neurophysiological evidence for visual perceptual categorization of words and faces within 150 ms. *Psychophysiology*. 1998; 35:240–251. [PubMed: 9564744]
- Schendan HE, Kutas M. Neurophysiological evidence for transfer appropriate processing of memory: processing versus feature similarity. *Psychonomic bulletin & review*. 2007; 14(4):612–619. [PubMed: 17972722]

- Schweinberger SR, Pickering EC, Jentzsch I, Burton AM, Kaufmann JM. Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Cognitive Brain Research*. 2002; 14(3):398–409. [PubMed: 12421663]
- 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]
- Sereno SC, Rayner K. Measuring word recognition in reading: eye movements and event-related potentials. *TRENDS in Cognitive Sciences*. 2003; 7:489–493. [PubMed: 14585445]
- Simon DA, Lewis G, Marantz A. Disambiguating form and lexical frequency effects in MEG responses using homonyms. *Language and Cognitive Processes*. 2012; 27:275–287.
- Thomas G, Friderici A, Hahne A. Brain responses during sentence reading: visual input affects central processes. *NeuroReport*. 1999; 10:3175–3178. [PubMed: 10574555]
- VanRullen R, Thorpe SJ. The Time Course of Visual Processing: From Early Perception to Decision Making. *Journal of Cognitive Neuroscience*. 2001; 13:454–461. [PubMed: 11388919]
- West WC, Holcomb PJ. Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. *Journal of Cognitive Neuroscience*. 2000; 12(6): 1024–1037. [PubMed: 11177422]
- Wlotko E, Federmeier KD. Finding the right word: Hemispheric asymmetries in the use of sentence context information. *Neuropsychologia*. 2007; 45:3001–3014. [PubMed: 17659309]

APPENDIX

A.1 Statistics for the Orthographic Regression Model

R-squared statistics, as well as t-statistics for each predictor included in the model. The associated permutation test p-value for each statistic is given in brackets. R-squared p-values were estimated by permuting observations; t p-values were estimated by permuting residuals. The orthographic regression was computed over the middle occipital electrode site. Reliable statistics are indicated by highlighting. In this table, as in all others in the Appendix, **only** temporal epochs in which the full model was reliable are included. See Figures 2 and 3 for full time courses of the overall model and the individual predictors, respectively.

Time (ms)	Orth Full Model (R-Squared)	Bigram (t)	Neighborhood (t)	Neighbor Frequency (t)
130	0.037 [0.011]			
140	0.035 [0.013]			
180	0.043 [0.004]	2.537 [0.011]		-2.713 [0.007]
190	0.046 [0.003]	2.518 [0.012]		-2.928 [0.004]
200	0.038 [0.008]	2.416 [0.017]		-2.734 [0.008]
230	0.036 [0.012]			
240	0.056 [0.001]	2.469 [0.014]		
250	0.090 [0.000]	3.043 [0.002]		
260	0.136 [0.000]	3.907 [0.000]		
270	0.194 [0.000]	5.052 [0.000]		
280	0.248 [0.000]	6.060 [0.000]		
290	0.288 [0.000]	6.752 [0.000]		
300	0.312 [0.000]	7.079 [0.000]		

Time (ms)	Orth Full Model (R-Squared)	Bigram (t)	Neighborhood (t)	Neighbor Frequency (t)
310	0.321 [0.000]	7.109 [0.000]	2.569 [0.012]	
320	0.323 [0.000]	6.879 [0.000]	2.751 [0.006]	
330	0.327 [0.000]	6.516 [0.000]	2.918 [0.004]	
340	0.330 [0.000]	6.064 [0.000]	3.033 [0.003]	
350	0.328 [0.000]	5.666 [0.000]	3.040 [0.002]	
360	0.315 [0.000]	5.309 [0.000]	2.864 [0.005]	
370	0.291 [0.000]	5.095 [0.000]	2.586 [0.010]	
380	0.258 [0.000]	4.962 [0.000]		
390	0.228 [0.000]	4.926 [0.000]		
400	0.204 [0.000]	4.786 [0.000]		
410	0.182 [0.000]	4.424 [0.000]		
420	0.153 [0.000]	3.717 [0.000]		
430	0.121 [0.000]	2.731 [0.005]		
440	0.091 [0.000]			
450	0.064 [0.000]			
460	0.042 [0.005]			

A.2 Statistics for the Whole-Item Frequency Regressions

R-squared statistics for the lexical frequency model, applied to raw amplitude data (first column) and residual data after a preliminary regression on bigram frequency. The associated permutation test p-value for each statistic is given in brackets. R-squared p-values were estimated by permuting observations. The frequency regression was computed over the middle occipital electrode site. Reliable statistics are indicated by highlighting. See Figure 2 for the full time course of the model.

Time (ms)	Frequency Model (R- Squared)	Frequency Model, Bigram Residuals (R-Squared)
270	0.086 [0.011]	
280	0.128 [0.001]	0.101 [0.005]
290	0.162 [0.000]	0.127 [0.002]
300	0.182 [0.000]	0.142 [0.001]
310	0.186 [0.000]	0.144 [0.001]
320	0.171 [0.000]	0.131 [0.001]
330	0.142 [0.001]	0.112 [0.003]
340	0.107 [0.004]	0.089 [0.010]
350	0.084 [0.012]	

A.3 Statistics for the Full Semantic Regression

R-squared statistics, as well as t-statistics for each predictor included in the model. The associated permutation test p-value for each statistic is given in brackets. R-squared p-values were estimated by permuting observations; t p-values were estimated by permuting residuals. The semantic regression was computed over the middle parietal electrode site. Reliable statistics are indicated by highlighting. See Figures 2 and 4 for full time courses of the over all model and the individual predictors, respectively.

Time (ms)	Sem Full Model (R-Squared)	Concreteness (t)	Imageability (t)	Number of Senses (t)	Number of Lexical Associates (t)	Noun/Verb Ambiguity (t)
300	0.183 [0.016]				2.775 [0.008]	
310	0.209 [0.006]				3.088 [0.002]	
320	0.207 [0.007]				3.128 [0.003]	
330	0.182 [0.016]				2.895 [0.005]	

A.4 Statistics for the Hierarchical Semantic Regression

R-squared statistics, as well as t-statistics for each predictor included in the model. The associated permutation test p-value for each statistic is given in brackets. R-squared p-values were estimated by permuting observations; t p-values were estimated by permuting residuals. The semantic regression was computed over the middle parietal electrode site. Reliable statistics are indicated by highlighting.

Time (ms)	Sem Hierarchical Model (R-Squared)	Concreteness (t)	Imageability(t)	Number of Senses (t)	Number of Lexical Associates (t)	Noun/Verb Ambiguity (t)
310	0.199 [0.008]				2.901 [0.005]	
320	0.213 [0.005]				3.113 [0.003]	
330	0.205 [0.006]				3.133 [0.002]	
340	0.187 [0.012]				2.956 [0.004]	

END OF APPENDIX

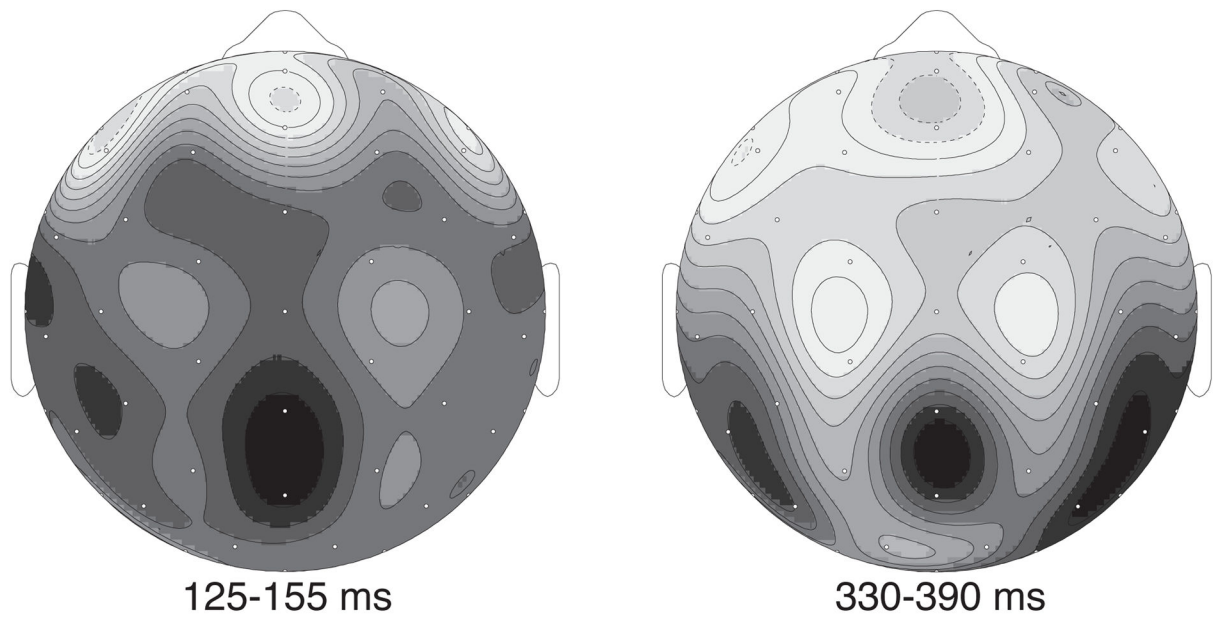


Figure 1. Distribution of the word frequency effect. The map on the left depicts the distribution of the word frequency effect around the peak of the P2 component. The map on the right depicts the distribution of the same effect around the peak of the N400 component. In both time windows the effect is especially large over midline parietal electrode sites. This observation motivated the selection of electrode sites for the present study.

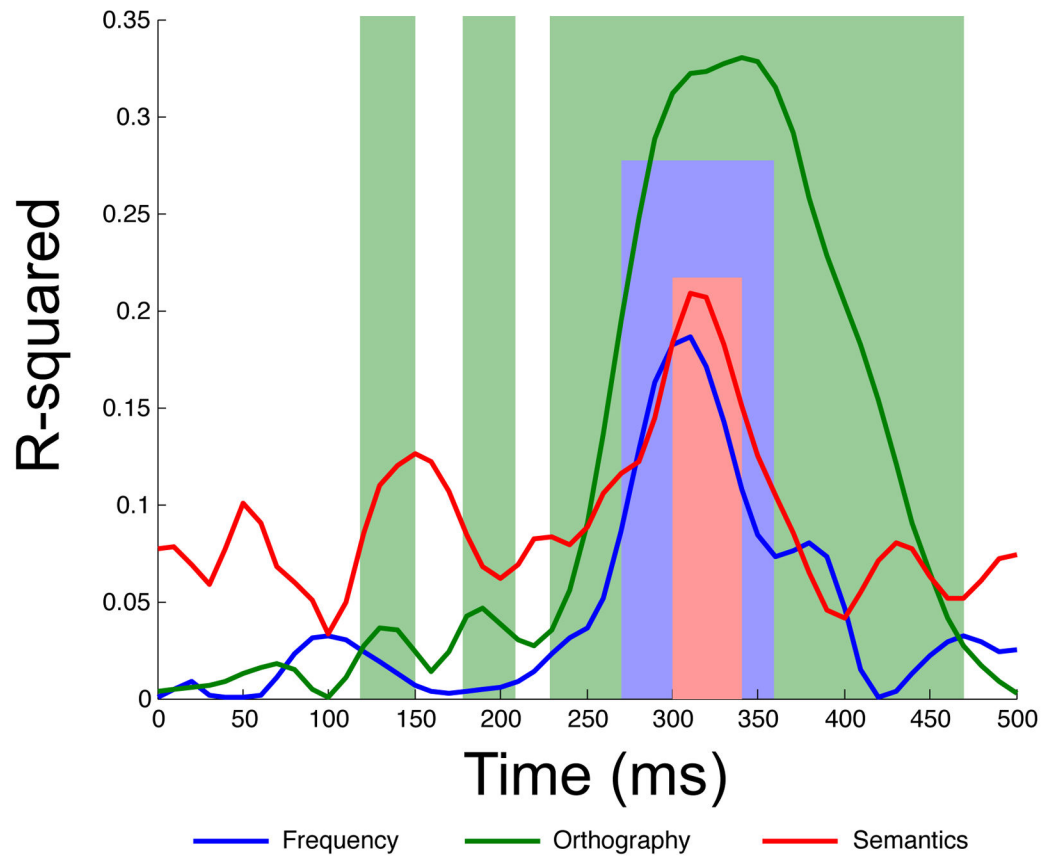


Figure 2. Time courses of R-squared for the lexical frequency, orthography, and semantic regression models. Epochs significant at $\alpha = .01$ are highlighted; blue highlighting for frequency, green for orthography, and red for semantics.

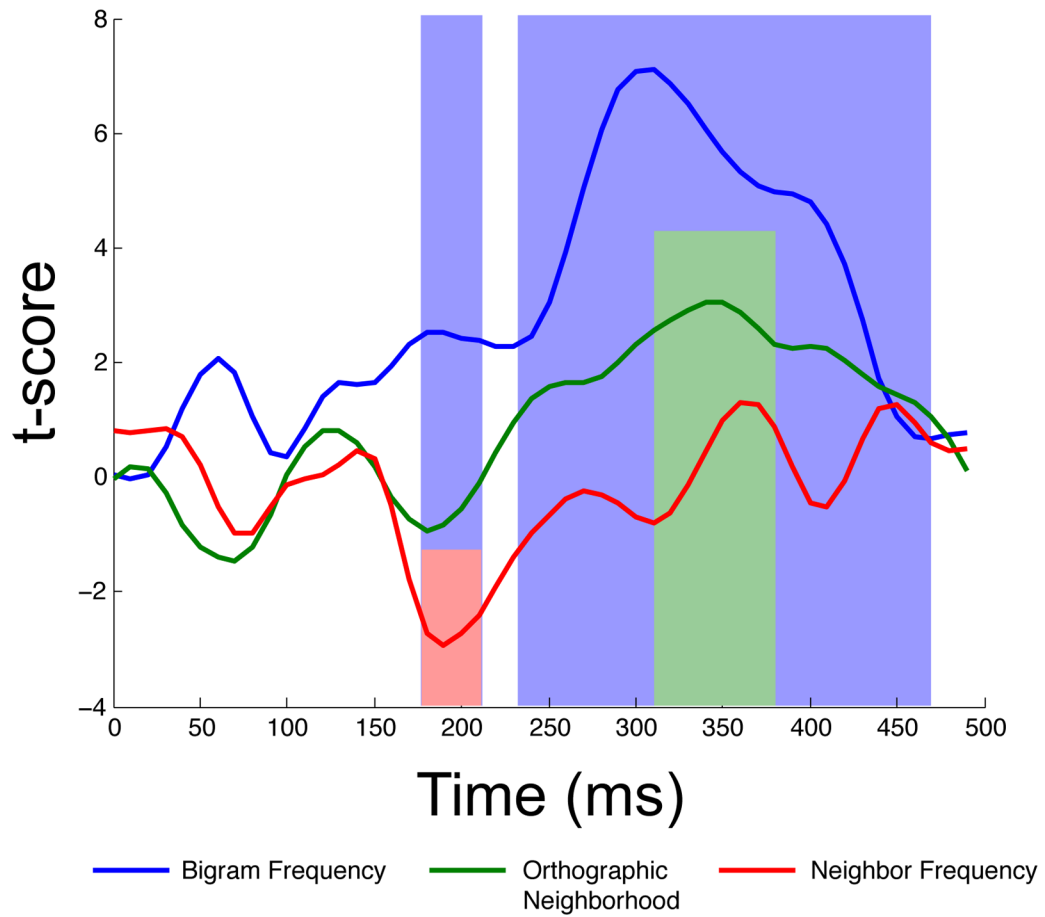


Figure 3. Time courses of t-scores for the individual predictors in the orthographic regression model. Epochs significant at $\alpha = .01$ are highlighted; blue highlighting for bigram frequency, green for orthographic neighborhood, and red for neighbor frequency.

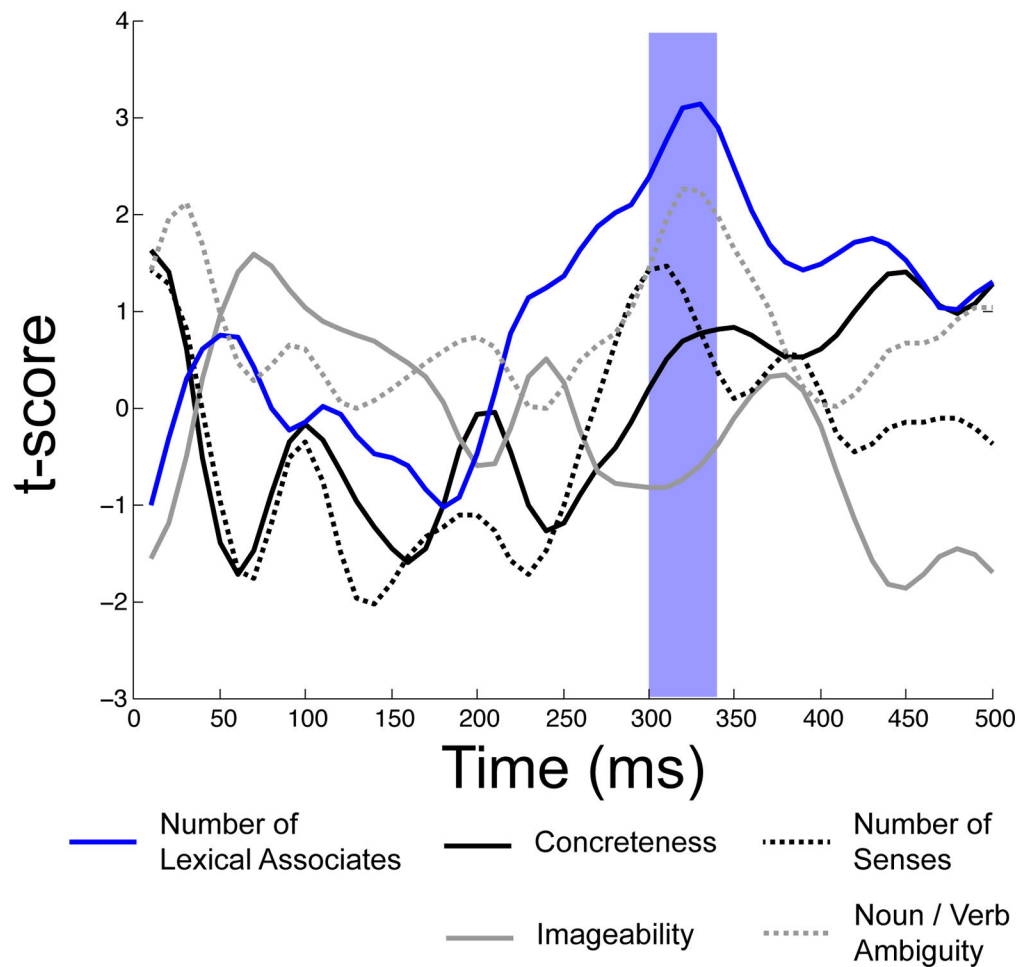


Figure 4. Time courses of t-scores for the individual predictors in the semantic regression model. Epochs significant at $\alpha = .01$ are highlighted. Only the Number of Lexical Associates predictor is reliable in an epoch in which the full model is also reliable.

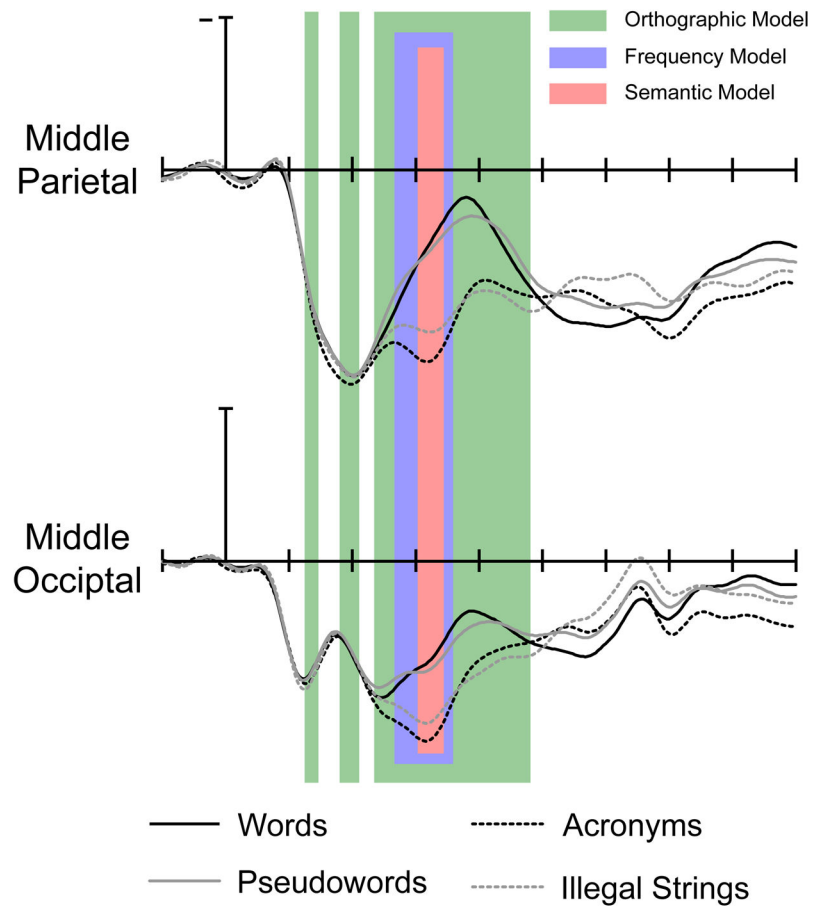


Figure 5. Grand average waveforms depicting the response to words, pseudowords, acronyms, and illegal strings over the middle parietal and middle occipital electrode sites. The semantic regression was computed over the middle parietal channel, while the orthographic and lexical frequency regressions were computed over the middle occipital channel.

Table 1

Orthographic Effects

Latency	Location	Citations
N/P150*: ~100–250 ms	Occipital [MiOc]	Carreiras, Dunabeitia, & Molinaro (2009); Chauncey, Holcomb, & Grainger (2008); Dufau, Grainger, & Holcomb (2008); Hauk & Pulvermuller (2004); Hauk, Davis, Ford, Pulvermuller, & Marslen-Wilson (2006); Hauk, Pulvermuller, Ford, Marslen-Wilson, & Davis (2009); Holcomb & Grainger (2006); Grainger, Kiyonaga, & Holcomb (2006); Petit, Midgley, Holcomb, & Grainger (2006); Massol, Grainger, Midgley, & Holcomb (2012); Carreiras, Perea, Vergara, & Pollatsek (2009); Chauncey, Holcomb, & Grainger, 2008
P200* ~150–250 ms	“Posterior”, “Anterior” [MiPf, MiOc]	Petit, Midgley, Holcomb, & Grainger (2006); Massol, Midgley, Holcomb, & Grainger (2011)
N250*: ~200–300 ms	Occipital [MiOc]	Carreiras, Dunabeitia, & Molinaro (2009); Hauk, Pulvermuller, Ford, Marslen-Wilson, & Davis (2009); Holcomb & Grainger (2006); Massol, Midgley, Holcomb, & Grainger (2011); Massol, Grainger, Midgley, & Holcomb (2012); Chauncey, Holcomb, & Grainger (2008)
P325* ~300–350 ms	Left Posterior [LmPa]	Holcomb & Grainger (2006); Chauncey, Holcomb, & Grainger (2008)
N400*: ~300–500 ms	Central Parietal [MiPa, LmPa, RmPa]	Carreiras, Dunabeitia, & Molinaro (2009); Holcomb, Grainger, & O’Rourke (2002); Holcomb & Grainger (325); Laszlo & Federmeier (2008); Laszlo & Federmeier (2009); Laszlo & Federmeier (2011); Massol, Grainger, Dufau, & Holcomb (2010); Massol, Midgley, Holcomb, & Grainger (2011); Carreiras, Vergara, Perea (2008); Vergara-Martinez, Perea, Marin, & Carreiras (2010); Chauncey, Holcomb, & Grainger (2008)

Meta-review of studies that have reported effects of orthographic manipulations. Epochs in the Latency column with a * are those where orthographic effects were found in the present study. In the Location column, topographic regions are given indicating where effects have been found for each latency; channels given in brackets underneath are those covering the relevant areas in the present study. MiPf: Middle Prefrontal, MiPa: Middle Parietal, LmPa: Left Middle Parietal, RmPa: Right Middle Parietal, MiOc: Middle Occipital.

Table 2

Lexical Frequency Effects

Latency	Location	Citations
N1: ~100– 150 ms	“Posterior” [MiPa, LmPa, RmPa, MiOc]	Dambacher, Dimigen, Braun, Wille, Jacobs, & Kliegl (2012); Hauk, Davis, Ford, Pulvermuller, & Marslen-Wilson (2006)
P2*: ~150– 250 ms	Fronto-Central, Occipital [MiPf, MiOc]	Dambacher, Kliegl, Hofmann, & Jacobs (2006); Hauk & Pulvermuller (2004); Sereno, Rayner, & Posner (1998); Dambacher, Dimigen, Braun, Wille, Jacobs, & Kliegl (2012); Hauk, Davis, Ford, Pulvermuller, & Marslen- Wilson (2006)
FSN*: 280– 335 ms	Left Anterior	King & Kutas (1998)
N400*: ~300– 500 ms	Central, Parietal, Occipital [MiCe, MiPa, LmPa, RmPa, MiOc]	Rugg (1990); Dambacher, Kliegl, Hofmann, & Jacobs (2006); Van Petten & Kutas (1990); Munte, Wieringa, Weyerts, Szentkuti, Matzke, & Johannes (2001); Hauk & Pulvermuller (2004); Osterhout, Bersick, & McKinnon, (1997); Young & Rugg (1992); Dambacher & Kliegl (2007); Hauk, Davis, Ford, Pulvermuller, & Marslen-Wilson (2006)

Meta-review of studies that have reported effects of lexical frequency. Epochs in the Latency column with a * are those where frequency effects were found in the present study. In the Location column, topographic regions are given indicating where effects have been found for each latency; channels given in brackets underneath are those covering the relevant areas in the present study. MiCe: Middle Central, MiPf: Middle Prefrontal, MiPa: Middle Parietal, LmPa: Left Middle Parietal, RmPa: Right Middle Parietal, MiOc: Middle Occipital.

Table 3

Semantic Effects

Latency	Location	Citations
80–100 ms	Frontal, Parietal [MiPf, MiPa, LmPa, RmPa]	Hauk, Pulvermuller, Ford, Marslen-Wilson, & Davis (2009)
104, 168 ms	Parietal, Occipital [MiPa, LmPa, RmPa, MiOc]	Segalowitz & Zheng, 2009
112 ms	Parietal [MiPa, LmPa, RmPa]	Sereno, Rayner, & Posner, 1998
160 ms	Frontal, Parietal, Occipital [MiPf, MiPa, RmPa, LmPa, MiOc]	Hauk, Davis, Ford, Pulvermuller, & Marslen-Wilson (2006); Hauk, Coutout, Holden, & Chen (2012)
N200: 150–200 ms	Fronto-Central [MiPf]	Amsel (2011); Amsel, Urbach, & Kutas (2013)
N400*: ~300–500 ms	Central Parietal, Occipital [MiPa, LmPa, RmPa, MiOc]	Barber, Donamayor, Kutas, & Munte (2010); Barber, Ben-zvi, Bentin, & Kutas (2010); Dambacher, Kliegl, Hofmann, & Jacobs (2006); Dambacher, Dimigen, Braun, Wille, Jacobs, & Kliegl (2012); Deacon, Dynowska, Ritter, & Grose-Fifer (2004); DeLong, Urbach, & Kutas (2005); Federmeier & Kutas (1999); Hauk, Davis, Ford, Pulvermuller, & Marslen-Wilson (2006); Hauk, Coutout, Holden, & Chen (2012); Holcomb, Grainger, & O'Rourke (2002); Kounios & Holcomb (1994); Kutas & Hillyard (1980; 1983; 1984) Laszlo & Federmeier (2007;2008;2009;2011;2012); Lee & Federmeier (2009); Rugg & Nagy (1987); Review in Kutas & Federmeier (2000; 2011); Grainger & Holcomb (2009); Lau, Phillips, & Poeppel (2008); <i>Examples from a larger set</i>

Meta-review of studies that have reported effects of semantic variables. Clearly, more studies have reported semantic effects in the N400 window than in all early windows combined (in fact, only a selection of N400 studies are included, along with several reviews). Reports of earlier effects are less consistent. Epochs in the Latency column with a * are those where semantic effects were found in the present study. In the Location column, topographic regions are given indicating where effects have been found for each latency; channels given in brackets underneath are those covering the relevant areas in the present study. MiPf: Middle Prefrontal, MiPa: Middle Parietal, LmPa: Left Middle Parietal, RmPa: Right Middle Parietal, MiOc: Middle Occipital.

Table 4

Mean lexical characteristics of each critical item type, with examples.

Item Type	Examples	Log Bigram Frequency	Log Whole Item Frequency	Orthographic Neighborhood Size	Log Frequency of Orthographic Neighbors	Concreteness	Imageability	Number of Senses	Number of Lexical Associates	Noun Verb Ambiguity
Word	DOG, CAR	6.3	2.39	13.0	4.3	374	379	5.9	10.5	0.84
Pseudoword	MIP, VAP	5.6	N/A	11.0	4.1	N/A	N/A	N/A	N/A	N/A
Acronym	BLT, DVD	-0.79	N/A	1.9	2.7	N/A	N/A	N/A	N/A	N/A
Illegal String	MBM, HCP	-0.38	N/A	2.4	3.0	N/A	N/A	N/A	N/A	N/A

Table 5

Intercorrelation matrices and Variance Inflation Factors (VIFs) for the orthographic (A) and semantic (B) regression models (these measures do not apply to the lexical frequency model, which has only 1 predictor).

(A): Orthographic Regressors	Bigram Frequency	Neighborhood	Neighbor Frequency
Bigram Frequency	1		
Neighborhood	0.65	1	
Neighbor Frequency	0.56	0.64	1
VIF	1.81	2.17	1.80

(B): Semantic Regressors	Concreteness	Imageability	Number of Senses	Number of Lexical Associates	Noun/Verb Ambiguity
Concreteness	1				
Imageability	0.82	1			
Number of senses	0.10	0.11	1		
Number of Lexical Associates	0.03	0.08	0.28	1	
Noun/Verb Ambiguity	0.09	0.17	-0.34	0.01	1
VIF	3.07	3.20	1.21	1.09	1.20