A Thousand Words Are Worth a Picture: Snapshots of Printed-Word Processing in an Event-Related Potential Megastudy

Stéphane Dufau, Jonathan Grainger, Katherine J. Midgley, Phillip J. Holcomb

Psychological Science
2015, Vol. 26(12) 1887–1897
DOI: 10.1177/0956797615603934

A Thousand Words Are Worth a Picture: Snapshots of Printed-Word Processing in an Event-Related Potential Megastudy

Stéphane Dufau[1,2,3], Jonathan Grainger[1,2,3], Katherine J. Midgley[4], and Phillip J. Holcomb[4,5]

[1]Laboratoire de Psychologie Cognitive, Aix-Marseille University; [2]Centre National de la Recherche Scientifique (CNRS), Marseille, France; [3]Brain and Language Research Institute, Aix-Marseille University; [4]Department of Psychology, San Diego State University; and [5]Department of Psychology, Tufts University

Abstract

In the experiment reported here, approximately 1,000 words were presented to 75 participants in a go/no-go lexical decision task while event-related potentials (ERPs) were recorded. Partial correlations were computed for variables selected to reflect orthographic, lexical, and semantic processing, as well as for a novel measure of the visual complexity of written words. Correlations were based on the item-level ERPs at each electrode site and time slice while a false-discovery-rate correction was applied. Early effects of visual complexity were seen around 50 ms after word onset, followed by the earliest sustained orthographic effects around 100 to 150 ms, with the bulk of orthographic and lexical influences arising after 200 ms. Effects of a semantic variable (concreteness) emerged later, at around 300 ms. The overall time course of these ERP effects is in line with hierarchical, cascaded, interactive accounts of word recognition, in which fast feed-forward influences are consolidated by top-down feedback via recurrent processing loops.

Keywords

ERPs, megastudies, visual word recognition, item-level ERPs, visual complexity, open data, open materials

Received 9/12/14; Revision accepted 8/12/15

The last decade has been witness to an increasing number of large-scale behavioral investigations of visual word recognition. In so-called "megastudies," a large number of responses are collected for large samples of words, which enables item-level analyses to be performed on the data set (e.g., Balota et al., 2007; Dufau et al., 2011; Ferrand et al., 2010; Keuleers, Lacey, Rastle, & Brysbaert, 2012). However, behavior is behavior, and as such can be measured only at the very end point of processing. Therefore, given the importance of specifying the relative timing of component processes in reading (e.g., Grainger & Holcomb, 2009), one might be well-advised to look beyond behavioral results for appropriate data, and there is one measurement technique that is particularly well suited for such an endeavor. This technique involves the millisecond-by-millisecond recording of the brain's electrical activity and time locking this activity (the electroencephalogram, or EEG, signal) to the onset of a given stimulus to measure changes in electrical activity that are provoked by a given stimulus or category of stimuli—the event-related potential (ERP).

By generating item-level data for a large set of items, megastudies provide the opportunity to explore effects of different variables in a parametric, continuous manner (see Balota, Yap, Hutchison, & Cortese, 2012; Brysbaert, Keuleers, & Mandera, 2014, for a review of the advantages of the megastudy approach). For the purposes of such explorations, megastudies apply correlational approaches to data analysis.[1] Highly relevant for the present work, therefore, are prior studies that have applied regression analyses on item-level ERP data to examine the timing of component processes in visual word recognition (see Dien, Frishkoff, Cerbone, & Tucker, 2003, for an early application of this general approach to the study of word comprehension in sentence contexts,

Corresponding Author:
Jonathan Grainger, Laboratoire de Psychologie Cognitive, Aix-Marseille University and CNRS, 3 Place Victor Hugo, 13331 Marseille, France
E-mail: jonathan.grainger@univ-amu.fr

and see Rey, Dufau, Massol, & Grainger, 2009, and Madec, Rey, Dufau, Klein, & Grainger, 2012, for item-level ERP analyses with letter stimuli).

In one large-scale study, Laszlo and Federmeier (2011) tested 120 participants with 75 words and various kinds of nonword stimuli. Regression analyses were performed on item-level ERPs obtained from six electrode sites selected to best capture the N400 ERP component, a negative-going waveform that peaks around 400 ms after stimulus onset. Orthographic neighborhood and number of semantic associates were both found to significantly influence N400 amplitude. In a follow-up study, Laszlo and Federmeier (2014) performed further regression analyses of the same ERP data set using variables designed to cover orthographic, lexical, and semantic effects, correcting for the multiple comparisons that such analyses involve. The earliest reliable effects were seen between 130 and 150 ms in the form of effects of a composite "orthographic" variable combining bigram frequency, orthographic neighborhood size, and orthographic neighborhood frequency.

Other studies have revealed even earlier effects of orthographic and lexical variables on ERPs. Hauk, Davis, Ford, Pulvermüller, and Marslen-Wilson (2006) tested 20 participants with 300 words presented intermixed with an equal number of nonword stimuli in a lexical decision task. A principal component analysis was used to construct a small number of composite variables. The results revealed an early orthographic effect (combining word length and *n*-gram frequency) at around 90 ms after stimulus onset and a slightly later effect of lexical frequency at 110 ms. These findings were confirmed in a follow-up study (Hauk, Pulvermüller, Ford, Marslen-Wilson, & Davis, 2009) that revealed an effect of word length and orthographic neighborhood starting around 100 ms after word onset. Amsel (2011) tested 28 participants with 207 words and analyzed effects of word length, word frequency, and a host of semantic variables using linear mixed-effects models applied to single-trial ERPs. Like Hauk, Davis, et al. (2006; Hauk et al., 2009), Amsel found an early effect of word length starting around 110 ms and peaking at around 250 ms, but found effects of word frequency arising much later than in the Hauk, Davis, et al. (2006) study.

One key comparison point is missing in all these prior studies: the influence of purely visual factors[2] that can be used as a benchmark against which the timing of downstream orthographic and lexical influences can be evaluated. As noted by Laszlo and Federmeier (2014), effects of word frequency found prior to 100 ms after word onset (e.g., Sereno & Rayner, 2003) merit a certain amount of suspicion given current knowledge of the timing of visual object-identification processes, plus the difficulties associated with controlling for the very large number of statistical comparisons that ERP time-course analyses can involve.

In the present experiment, 75 participants were tested with 960 words in a go/no-go lexical decision task, in which "go" responses were made to nonword stimuli presented approximately every 7 trials. The main aim of the analyses was to compare the timing of purely visual effects with orthographic, lexical, and semantic influences and to examine the evolution of these effects over time. To do so, we selected seven variables hypothesized to be sensitive to various combinations of visual, orthographic, lexical, and semantic processing: (a) visual complexity, (b) proportion of consonants versus vowels, (c) mean positional bigram frequency, (d) word length in letters, (e) orthographic similarity with other words, (f) word frequency, and (g) concreteness. We expected the sensitivity of ERP recordings to millisecond-by-millisecond changes in brain activity, combined with the high power of the experiment (75 participants tested with 960 words), to reveal the earliest influences of visual processing followed by subsequent orthographic, lexical, and semantic influences.

Method

Participants

Seventy-five healthy individuals (36 male, 39 female; average age = 20.4 years, range = 18–25 years) from Tufts University took part in the experiment as paid volunteers. All participants were right-handed native speakers of English and reported having normal or corrected-to-normal vision. The number of participants was predetermined as being sufficient to obtain at least 40 data points per word, after data loss, for item-based correlation analyses (see the ERP recording and analysis section for further details). No participants were excluded prior to analysis.

Procedure and stimuli

After completing informed consent and handedness forms, participants were seated in a comfortable chair in a sound-attenuated darkened room. Stimuli consisted of 960 nouns between four and eight letters in length (the critical stimuli) and 140 nonword items, which served as probes. Both types of stimuli were used in a go/no-go lexical decision task in which participants were instructed to respond with their right index finger as fast as possible whenever they detected a nonword and to withhold button presses for the remaining critical real-word items. Nonwords appeared once approximately every seven trials. Pronounceable, orthographically legal nonwords were formed by replacing one or two letters in internal positions of real words that were not in the list of critical items.

Stimuli were presented as white letters on a black background on a 19-in. CRT monitor in lowercase Arial font. Viewing distance was 120 cm, and all words subtended between 1° (four letters) and 2° (eight letters) of horizontal visual angle. Each trial began with a 400-ms presentation of a letter string followed by a 600-ms black screen. Participants were instructed to minimize blinking during the task and were given short 1-min rest breaks every 55 trials and a longer 4-min break between each block. Participants completed four blocks of 275 trials each.

EEG recording and analysis

An electrode cap (Electro-Cap, Eaton, OH) with tin electrodes was used to record continuous EEG from 29 sites on the scalp: left and right frontopolar (FP1, FP2), frontal (F3, F4, F7, F8), frontocentral (FC1, FC2, FC5, FC6), central (C3, C4), temporal (T7, T8), central-parietal (CP1, CP2, CP5, CP6), parietal (P3, P4, P7, P8), and occipital (O1, O2) areas and five midline sites over the frontal pole (FPz), frontal (Fz), central (Cz), parietal (Pz), and occipital (Oz) areas. In addition, four electrodes were attached to the face and neck: one below the left eye (to monitor for vertical eye movement and blinks), one to the right of the right eye (to monitor for horizontal eye movements), one over the left mastoid (reference), and one over the right mastoid (recorded actively to monitor for differential mastoid activity). All EEG electrode impedances were maintained below 5 kΩ (impedance for eye electrodes was less than 10 kΩ). The EEG was amplified by an SA Instrumentation (San Diego, CA) bioamplifier with a band-pass filter from 0.01 to 40 Hz, and the EEG was continuously sampled at a rate of 250 Hz.

The ERP data were time locked to word presentation and were recorded for 920 ms after target onset as well as for 100 ms before target onset to establish a baseline. A semiautomated method (automatic threshold-based detection and manual confirmation) was used to reject epochs with eye movements, blinks, or muscle artifacts. Each grand-average word ERP was calculated using the unique waveform from each participant generated by a given word. The minimal number of artifact-free trials per word was 43 (*M* = 60.01, *SD* = 3.76, range = 43–71), which gave an acceptable signal-to-noise ratio for the entire set of stimuli. On average, the pool of participants used to form each of the grand-average word ERPs overlapped by 65.57% (*SD* = 6%, range = 33%–89%).

Statistical analysis

For each time sample (4 ms) and each of the 29 scalp electrodes, a vector of 960 ERP values (corresponding to the 960 different words) was first extracted from the data set. For each ERP vector, outliers (more than 2 standard deviations from the mean) were removed, and two-tailed linear partial correlations were computed for each of the seven key variables. The concreteness (vs. abstractness) of the critical stimuli (*M* = 4.37, *SD* = 1.14, range = 1.65–6.90) was rated in a separate experiment.[3] Word frequency was the log frequency from the CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995; *M* = 2.44, *SD* = 0.81, range = 0.30–4.10). Orthographic distance was defined as the average Levenshtein distance of the 20 most orthographically similar words (Yarkoni, Balota, & Yap, 2008; *M* = 2.15, *SD* = 0.70, range = 1.00–8.00), where Levenshtein distance is the minimum number of letter substitutions, deletions, additions, or transpositions required to transform one word into another. The mean number of letters was 6.00 (*SD* = 1.41, range = 4.00–8.00), and the mean positional bigram frequency (Balota et al., 2007) was 532.34 (*SD* = 219.27, range = 81.25–1,369.30). Consonant-vowel proportion was calculated by dividing the number of consonants in a word by the number of letters (*M* = .61, *SD* = .10, range = .25–.90).

Finally, visual complexity (*M* = 65.37, *SD* = 6.35, range = 44.84–82.20) was measured as the mean perimetric complexity of the component letters, with perimetric complexity defined as the square of the length of the perimeter of the letter divided by the total ink area (Pelli, Burns, Farell, & Moore-Page, 2006; however, Pelli et al. used the sum of complexity of individual letters rather than the average, which confounds complexity with orthographic length). The word "lull" had the lowest complexity (44.8) in our stimuli and "poem" the highest (82.2; see Fig. 1). A correction for multiple comparisons using the false-discovery-rate (FDR) method was applied (Benjamini & Hochberg, 1995; Benjamini & Yekutieli, 2001; Groppe, Urbach, & Kutas, 2011).

Results

The partial correlations for each time slice and electrode are shown in Figure 2. The most significant correlations ($p < .01$, FDR corrected) are color coded to show the direction and the strength of the correlation. The remaining significant correlations ($p < .05$, FDR corrected) are shown in gray simply to indicate that a significant correlation was present between a given variable and the voltage values obtained at a given electrode site at a given time. Plotting the results in this way enables an immediate appreciation of the timing and spatial distribution of the most robust effects among the variables we chose to analyze.

Figure 3 shows example waveforms obtained by averaging voltages separately for the bottom 25% and the top 25% of the 960 values for each of the seven

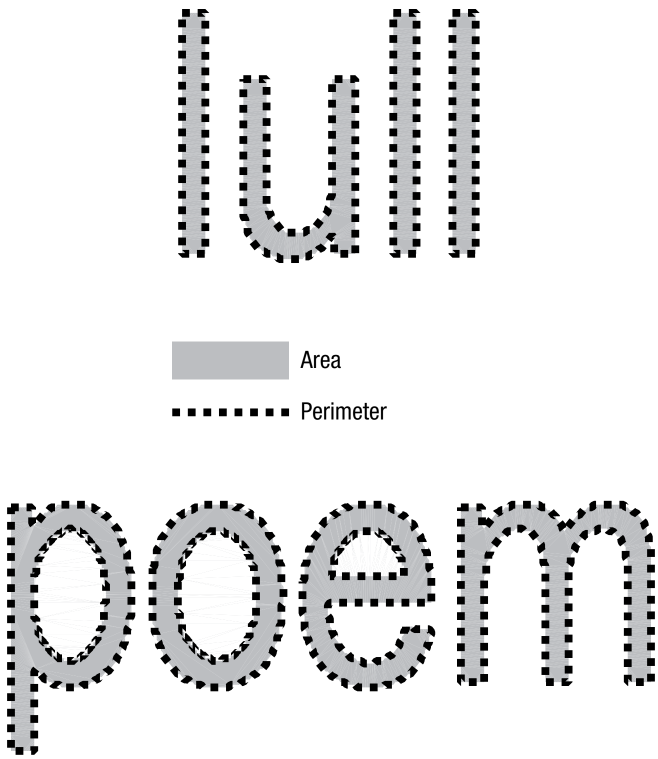

Fig. 1. Procedure for determining visual complexity of printed word stimuli (measured using the perimetric-complexity method of Pelli, Burns, Farell, & Moore-Page, 2006). Visual complexity was first computed for each letter by squaring the sum of its inside and outside perimeters (dashed lines) and dividing by the area between the lines. The word's visual complexity was the average complexity of its component letters.

variables at the electrode sites selected to best illustrate their effects. This information is presented for illustrative purposes only, since the effects of each individual variable are contaminated by the influence of uncontrolled variables in this figure. Figure 2 provides the key results of the present experiment, the overall pattern of which can be approximately divided into five time windows.

In Time Window 1 (32–52 ms after stimulus onset), there was an initial burst in the effects of visual complexity and some early effects of consonant-vowel proportion, as well as more isolated effects of concreteness and word frequency. In Time Window 2 (100–152 ms), there was an initial burst in the effects of word length (number of letters), accompanied by effects of word frequency in posterior electrode sites, as well as effects of consonant-vowel proportion. Effects of visual complexity became stronger and more widely distributed in this time window. In Time Window 3 (180–280 ms), there were widespread effects of word length and, lagging behind these, an increase in the effects of word frequency in posterior electrode sites and the emergence of effects of bigram frequency in frontal and central electrodes. Visual complexity continued to have a widespread influence on ERPs in this time window.

In Time Window 4 (280–380 ms), there were widespread effects of concreteness accompanied by effects of word frequency in frontal electrode sites that were opposite in polarity to the earlier posterior effects. There were also relatively widespread influences of orthographic distance and consonant-vowel proportion in this time window and a continuing effect of visual complexity. Finally, in Time Window 5 (380–500 ms), there were widespread effects of word frequency accompanied by effects of word length in posterior electrode sites, as well as a continuing but diminishing influence of concreteness, orthographic distance, bigram frequency, and consonant-vowel proportion. In order to provide a more detailed appreciation of the scalp distribution of the different effects, we also provide the topographic distribution of the partial correlations of each variable of interest in these five time windows (see Fig. 4).

Finally, we also examined ERP activities prior to word onset in the −100 ms to 0 ms baseline time window and from 500 ms to 900 ms after word onset. Prior to word onset, there were only two significant ERP activities, which were driven by differences in number of letters ($p < .05$) at two different time points and two different electrodes. Most of the effects seen from 500 ms to 900 ms after word onset were driven by word frequency, which continued to have a strong impact on ERPs. Consonant-vowel proportion also continued to have an influence up to about 600 ms after word onset, and there were some smaller more isolated effects of orthographic distance and number of letters.

Table 1 provides the partial-correlation matrix for the seven variables examined in the present experiment, and Table 2 provides the partial correlations between each of these variables and mean lexical decision response time (RT) extracted from the English Lexicon Project database (Balota et al., 2007). In Table 1, note the correlations between the new variables that we introduced in the present experiment (visual complexity, consonant-vowel proportion) and the other variables. These correlations might help explain some of the divergences in the pattern of ERP effects reported here with respect to prior studies.

Lexical decision RT had the strongest correlation by far with word frequency ($pr = -.59$), followed by number of letters ($pr = .09$). What is more interesting is the relatively strong correlation between consonant-vowel proportion and mean RT ($pr = -.08$), plus the significant positive correlation between bigram frequency and consonant-vowel proportion ($pr = .12$). The latter correlation reflects that fact that the most frequent bigrams in English are composed of consonants. Thus, within the list of the 100 most-frequent English bigrams found in Google books, 6 are composed of two vowels, and 19 are

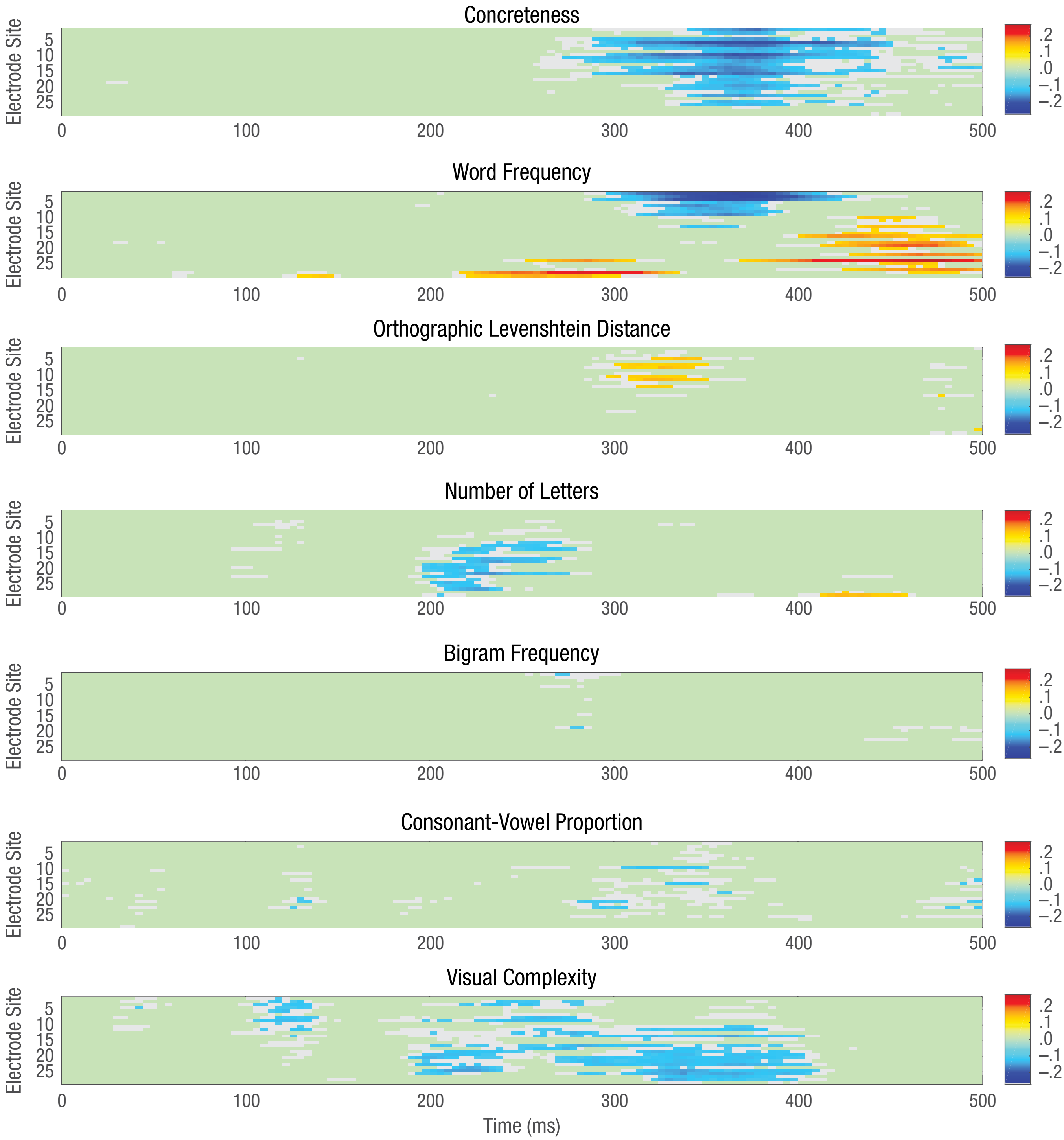

Fig. 2. Partial correlations between event-related potential voltage and each variable of interest at each electrode and time slice. Color coding is shown for correlations that are significant at less than .01 (false-discovery-rate corrected), and other significant correlations ($p < .05$) are shown in gray. On the y-axes, frontal electrode sites are at the top, and occipital sites are at the bottom.

composed of two consonants (Norvig, 2013). These two correlations point to a possible explanation for the relative fragility of effects of bigram frequency, as confirmed in the present ERP data (see Discussion). Finally, it should be noted that visual complexity did not have a significant influence on RT. Future work will examine how the visual complexity of a word's component letters can be combined with factors such as letter visibility and the information carried by each letter (Stevens & Grainger, 2003) to predict word-identification times.

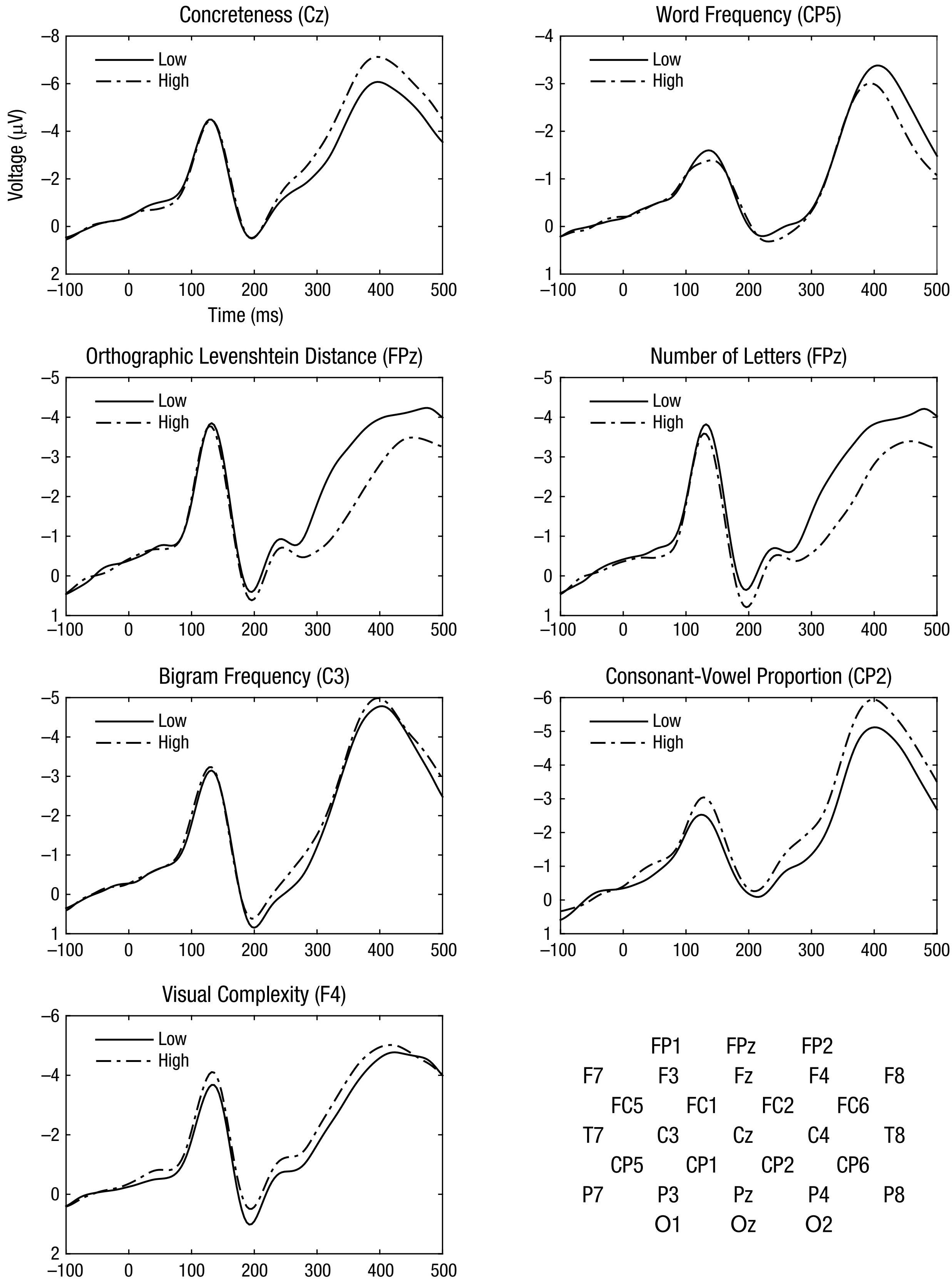

Fig. 3. Grand-average waveforms in response to word stimuli in the lexical decision task. For each variable, the bottom 25% (low) and top 25% (high) values are shown for a representative electrode site. The positions of the selected electrodes are shown at the bottom right (F = frontal, C = central, T = temporal, P = parietal, O = occipital, z = midline; odd numbers are in the left hemisphere, and even numbers are in the right hemisphere).

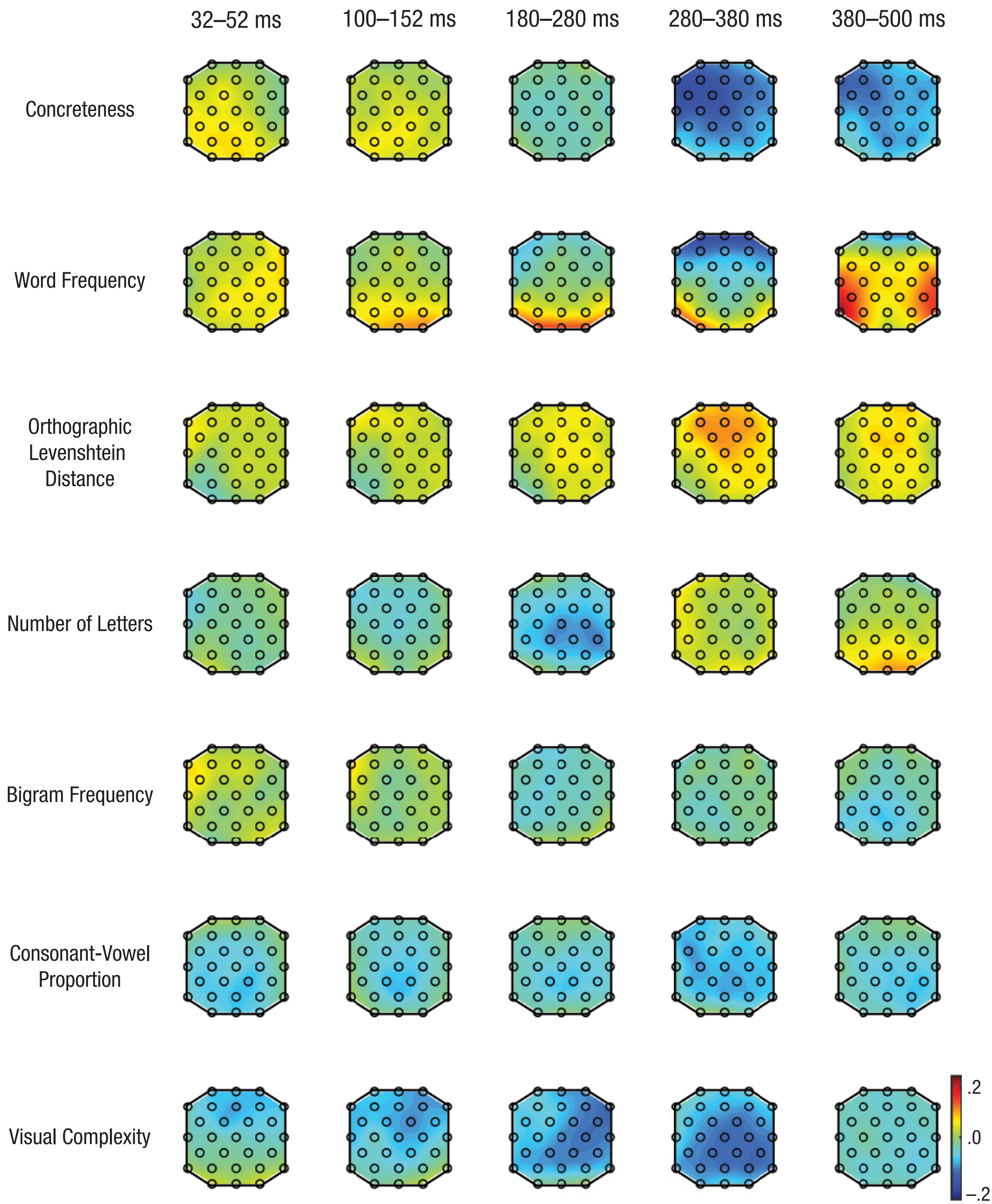

Fig. 4. Topographic maps showing partial correlations between each variable and the average event-related potential voltage values at each electrode site, separately for each time window. Values were smoothed across the scalp (see Fig. 3 for the electrode montage).

Table 1. Partial Correlations Between the Variables of Interest

Variable	1	2	3	4	5	6
1. Concreteness	—					
2. Frequency	−.17***	—				
3. Orthographic Levenshtein distance	−.02	−.24***	—			
4. Number of letters	−.11***	.11***	.75***	—		
5. Mean bigram frequency	.04*	.08**	−.32***	.35***	—	
6. Consonant-vowel proportion	.09**	−.03*	−.06*	−.06*	.12***	—
7. Visual complexity	−.02	.04*	.06*	−.12***	−.03*	−.10**

*$p < .05$. **$p < .01$. ***$p < .001$.

Discussion

Item-level data obtained in a large-scale ERP lexical decision experiment were analyzed in order to reveal the time course of orthographic, lexical, and semantic influences during the processing of printed word stimuli. Crucially, effects of linguistic variables were compared with the effects of a visual variable, visual complexity, analyzed for the first time in an ERP study of visual word recognition. Although temporally and spatially isolated effects of several variables arose prior to 100 ms after word onset, only two variables had more widespread effects—visual complexity and consonant-vowel proportion. The timing of the early effects of visual complexity provides a baseline against which the effects of linguistic variables can be better evaluated. It suggests that caution should be exercised when interpreting the isolated effects of concreteness and word frequency seen in roughly the same time window. It further suggests that the more widespread effects of consonant-vowel proportion seen in the same epoch could well be driven by visual differences between consonants and vowels that were not captured by our measure of visual complexity.

Effects of word length (number of letters) emerged around 100 ms after word onset, in line with prior reports of length effects emerging in a similar time window (e.g., Amsel, 2011; Hauk, Davis, et al., 2006; Hauk et al., 2009). Rapidly after the onset of effects of word length, we saw an influence of word frequency in posterior electrode sites. Word-frequency effects gradually became stronger and extended to parietal electrode sites between 200 and 300 ms, and they became even more widespread between 400 and 500 ms. The effect of word frequency seen between 120 and 160 ms after stimulus onset is in line with the estimated onset of frequency effects reported in prior studies (e.g., Chen, Davis, Pulvermüller, & Hauk, 2015; Hauk, Davis, et al., 2006; Strijkers, Bertrand, & Grainger, 2015).

Effects of orthographic distance started to emerge around 280 ms. The observed pattern is in line with prior investigations manipulating the number of single-letter-substitution neighbors (e.g., Holcomb, Grainger, & O'Rourke, 2002; Laszlo & Federmeier, 2011). In previous research (Holcomb et al., 2002), we found that orthographic overlap with other words results in increased negativity in the ERP waveforms because of the greater overall activity in lexical representations. The fact that words with a greater proportion of consonants generated more negative ERPs suggests that this effect might be akin to effects of orthographic neighborhood, with more consonants leading to more activity in whole-word representations (Carreiras, Duñabeitia, & Molinaro, 2009).[4] In a similar vein, the very limited effects of bigram frequency took the form of negative correlations with ERP voltage, such that the greater the bigram frequency of a word, the more negative the voltage. The fact that we found no evidence for an early effect of bigram frequency, somewhat in contradiction with prior studies (Hauk, Davis, et al., 2006; Hauk, Patterson, et al., 2006; Laszlo & Federmeier, 2014), could be due to the *n*-gram frequency effects reported in prior research being mainly driven by effects of consonant-vowel proportion.

Contrary to a number of prior studies, the present results showed no evidence for early effects of a semantic variable (concreteness). Three studies in particular (Amsel, Urbach, & Kutas, 2013; Chen et al., 2015; Hauk,

Table 2. Partial Correlations Between Mean Lexical Decision Response Time (RT) and the Variables of Interest

Variable	*pr*
Concreteness	−.07 (p = .03)
Frequency	−.59 (p < .01)
Orthographic Levenshtein distance	.08 (p = .02)
Number of letters	.09 (p < .01)
Mean bigram frequency	.06 (p = .09)
Consonant-vowel proportion	−.08 (p = .01)
Visual complexity	−.06 (p = .10)

Note: RTs were extracted from the English Lexicon Project database (Balota et al., 2007) and log-10 transformed; RTs exceeding 2 standard deviations from the mean were removed, which left 916 data points for the analysis.

Coutout, Holden, & Chen, 2012) converged on an estimate of 160 ms for the emergence of semantic influences in the EEG and magnetoencephalogram (MEG) signal. This estimate was obtained from ERP effects in a living/nonliving categorization task (Amsel et al., 2013; Hauk et al., 2012) and from regression analyses of the effects of imageability and concreteness on EEG and MEG responses (Chen et al., 2015). Chen et al. also reported that the effects of imageability and concreteness were stronger in a silent reading task than in the lexical decision task. More generally, the task modulation of early orthographic, lexical, and semantic effects reported in the Chen et al. (2015) study points to a key role for proactive top-down mechanisms that modify stimulus-driven processing (see also Strijkers et al., 2015, for task modulation of word-frequency effects). In a go/no-go living/nonliving categorization task, for example, preactivation of the semantic features associated with the target category would enable these same features to reach criterion levels of activation faster on stimulus presentation, compared with presentation of the same word in the lexical decision task (see Laszlo & Federmeier, 2014, for a similar proposal). This, however, cannot account for the early effect of imageability (concreteness) found in a silent reading task by Chen et al. (2015). In this respect, the timing of effects of concreteness in the present experiment is more in line with the effects of semantic variables (including concreteness) emerging around 300 ms after stimulus onset in Laszlo and Federmeier's (2014) regression analysis and in line with the results of factorial manipulations of abstract versus concrete words showing effects on the N400 component (e.g., Barber, Otten, Kousta, & Vigliocco, 2013; West & Holcomb, 2000).

Overall, the timing of the effects of the different variables examined in this experiment suggests a fast initial feed-forward sweep of neural activity cascading through visual, orthographic, and lexical representations. This feed-forward activity would represent a fragile initial state of the network prior to stabilization through feedback (Grainger & Holcomb, 2009). As pointed out by Strijkers et al. (2015), this reactive feedback needs to be complemented with proactive mechanisms that enable preparatory activity prior to stimulus presentation. The combination of reactive and proactive top-down influences is likely to be at least partly responsible for the discrepancies in the timing estimates of component processes in visual word recognition obtained from ERP data, with some effects being visible only following feedback consolidation and some effects being particularly sensitive to task-related preparatory mechanisms. Finally, the results of the present experiment strongly suggest a need for the inclusion of measures of visual influences on the ERP signal when effects of linguistic variables are evaluated. It will also be important to examine the extent to which visual influences, as reflected in effects of variables such as visual complexity, are sensitive to the nature of the task being performed. Another useful manipulation for future research would be to present stimuli in both lowercase and uppercase formats in order to further separate the visual from the linguistic in printed word perception.

Author Contributions

S. Dufau analyzed the data, prepared all graphics, and participated in the writing. J. Grainger participated in the design of the experiment and wrote the first draft of the manuscript. K. J. Midgley participated in setting up the experiment and supervised the data collection. P. J. Holcomb designed the experiment and participated in the data analyses and writing of the manuscript.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This experiment was funded by the European Research Council (Grant No. 230313) and the National Institutes of Health (Grant No. HD25889).

Open Practices

All data and materials have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/72b89/. The complete Open Practices Disclosure for this article can be found at http://pss.sagepub.com/content/by/supplemental-data. This article has received badges for Open Data and Open Materials. More information about the Open Practices badges can be found at https://osf.io/tvyxz/wiki/1.%20View%20the%20Badges/ and http://pss.sagepub.com/content/25/1/3.full.

Notes

1. Megastudies also provide databases that can be used to perform "virtual" factorial experiments (e.g., Kuperman, 2015), and clearly the more items there are in the database, the more possibilities there are to perform such virtual experiments.
2. Word length could be considered a visual variable, but the visual component of word length (i.e., physical length) is completely confounded with its orthographic component (i.e., length in letters) in all the studies cited here.
3. Twenty different participants were asked to rate the 960 words tested in the present experiment on a 7-point Likert scale (7 = *very concrete*, 1 = *very abstract*). Participants were encouraged to use the full range of values. The obtained ratings were found to correlate highly with the concreteness ratings published by Brysbaert, Warriner, and Kuperman (2014) for the 931 words common to both studies ($r = .90$, $p < .001$).
4. Within words, consonant-vowel proportion correlates with number of syllables, a factor known to influence visual word

recognition (e.g., Chétail, 2014). However, entering number of syllables into the partial correlation analyses revealed a much reduced impact on ERPs compared with consonant-vowel proportion.

References

Amsel, B. D. (2011). Tracking real-time neural activation of conceptual knowledge using single-trial event-related potentials. *Neuropsychologia*, *49*, 970–983.

Amsel, B. D., Urbach, T. P., & Kutas, M. (2013). Alive and grasping: Stable and rapid semantic access to an object category but not object graspability. *NeuroImage*, *77*, 1–13.

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

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. I., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*, 445–459.

Balota, D. A., Yap, M. J., Hutchison, K. A., & Cortese, M. J. (2012). Megastudies: What do millions (or so) of trials tell us about lexical processing? In J. S. Adelman (Ed.), *Visual word recognition* (Vol. 1, pp. 90–115). Hove, England: Psychology Press.

Barber, H. A., Otten, L. J., Kousta, S.-T., & Vigliocco, G. (2013). Concreteness in word processing: ERP and behavioral effects in a lexical decision task. *Brain & Language*, *125*, 47–53.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*, 289–300.

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, *29*, 1165–1188.

Brysbaert, M., Keuleers, E., & Mandera, P. (2014). A plea for more interactions between psycholinguists and natural language processing research. *Computational Linguistics in the Netherlands*, *4*, 209–222.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavioral Research Methods*, *46*, 904–911. doi:10.3758/s13428-013-0403-5

Carreiras, M., Duñabeitia, J. A., & Molinaro, N. (2009). Consonants and vowels contribute differently to visual word recognition: ERPs of relative position priming. *Cerebral Cortex*, *19*, 2659–2670.

Chen, Y., Davis, M. H., Pulvermüller, F., & Hauk, O. (2015). Early visual word processing is flexible: Evidence from spatiotemporal brain dynamics. *Journal of Cognitive Neuroscience*, *27*, 1738–1751. doi:10.1162/jocn_a_00815

Chétail, F. (2014). Effect of number of syllables: New insights from the lexical decision task. *Language, Cognition and Neuroscience*, *29*, 1249–1256.

Dien, J., Frishkoff, G. A., Cerbone, A., & Tucker, D. M. (2003). Parametric analysis of event-related potentials in semantic comprehension: Evidence for parallel brain mechanisms. *Cognitive Brain Research*, *15*, 137–153.

Dufau, S., Duñabeitia, J. A., Moret-Tatay, C., McGonigal, A., Peeters, D., Alario, F.-X., . . . Grainger, J. (2011). Smart phone, smart science: How the use of smartphones can revolutionize research in cognitive science. *PLoS ONE*, *6*(9), Article e24974. Retrieved from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0024974

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., . . . Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods*, *42*, 488–496.

Grainger, J., & Holcomb, P. J. (2009). Watching the word go by: On the time-course of component processes in visual word recognition. *Language & Linguistics Compass*, *3*, 128–156.

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, *48*, 1711–1725.

Hauk, O., Coutout, C., Holden, A., & Chen, Y. (2012). The time-course of single-word reading: Evidence from fast behavioral and brain responses. *NeuroImage*, *60*, 1462–1477.

Hauk, O., Davis, M. H., Ford, M., Pulvermüller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *NeuroImage*, *30*, 1383–1400.

Hauk, O., Patterson, K., Woollams, A., Watling, L., Pulvermüller, F., & Rogers, T. T. (2006). [Q:] When would you prefer a SOSSAGE to a SAUSAGE? [A:] At about 100 msec. ERP correlates of orthographic typicality and lexicality in written word recognition. *Journal of Cognitive Neuroscience*, *18*, 818–832.

Hauk, O., Pulvermüller, F., Ford, M., Marslen-Wilson, W. D., & Davis, M. H. (2009). Can I have a quick word? Early electrophysiological manifestations of psycholinguistic processes revealed by event-related regression analysis of the EEG. *Biological Psychology*, *80*, 64–74.

Holcomb, P. J., Grainger, J., & O'Rourke, T. (2002). An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Journal of Cognitive Neuroscience*, *14*, 938–950.

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods*, *44*, 287–304.

Kuperman, V. (2015). Virtual experiments in megastudies: A case study of language and emotion. *Quarterly Journal of Experimental Psychology*, *68*, 1693–1710. doi:10.1080/17470218.2014.989865

Laszlo, S., & Federmeier, K. D. (2011). The N400 as a snapshot of interactive processing: Evidence from regression analyses of orthographic neighbor and lexical associate effects. *Psychophysiology*, *48*, 176–186.

Laszlo, S., & Federmeier, K. D. (2014). Never seem to find the time: Evaluating the physiological time course of visual word recognition with regression analyses of single-item event-related potentials. *Language, Cognition and Neuroscience*, *29*, 642–661.

Madec, S., Rey, A., Dufau, S., Klein, M., & Grainger, J. (2012). The time course of visual letter perception. *Journal of Cognitive Neuroscience*, *24*, 1645–1655.

Norvig, P. (2013). *English letter frequency counts: Mayzner revisited*. Retrieved from http://norvig.com/mayzner.html

Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, *46*, 4646–4674.

Rey, A., Dufau, S., Massol, S., & Grainger, J. (2009). Testing computational models of letter perception with item-level event-related potentials. *Cognitive Neuropsychology*, *26*, 7–22.

Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: Eye movements and event-related potentials. *Trends in Cognitive Sciences*, *7*, 489–493.

Stevens, M., & Grainger, J. (2003). Letter visibility and the viewing position effect in visual word recognition. *Perception & Psychophysics*, *65*, 133–151.

Strijkers, K., Bertrand, D., & Grainger, J. (2015). Seeing the same words differently: The time course of automaticity and top-down intention in reading. *Journal of Cognitive Neuroscience*, *27*, 1542–1551. doi:10.1162/jocn_a_00797

West, W. C., & Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: An electrophysiological investigation. *Journal of Cognitive Neuroscience*, *12*, 1024–1037.

Yarkoni, T., Balota, D. A., & Yap, M. J. (2008). Moving beyond Coltheart's *N*: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*, 971–979.