Predictive pre-activation of orthographic and lexical-semantic representations facilitates visual word recognition

Abbreviated Title: Predictive pre-activation in word recognition

Susanne Eisenhauer[1*], Benjamin Gagl[1,2*] & Christian J. Fiebach[1,2,3] (*equal contribution)

[1] Department of Psychology, Goethe University Frankfurt, Frankfurt am Main, Germany

[2] Center for Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt am Main, Germany

[3] Brain Imaging Center, Goethe University Frankfurt, Frankfurt am Main, Germany

Address for Correspondence:

Susanne Eisenhauer

eisenhauer@psych.uni-frankfurt.de

Department of Psychology

Goethe University Frankfurt

Theodor W. Adorno-Platz 6

60323 Frankfurt am Main, Germany

Abstract

To a crucial extent, the efficiency of reading results from the fact that visual word recognition is faster in predictive contexts. Predictive coding models suggest that this facilitation results from pre-activation of predictable stimulus features across multiple representational levels before stimulus onset. Still, it is not sufficiently understood which aspects of the rich set of linguistic representations that are activated during reading – visual, orthographic, phonological, and/or lexical-semantic – contribute to context-dependent facilitation. To investigate in detail which linguistic representations are pre-activated in a predictive context and how they affect subsequent stimulus processing, we combined a well-controlled repetition priming paradigm, including words and pseudowords (i.e., pronounceable nonwords), with behavioral and magnetoencephalography measurements. For statistical analysis, we used linear mixed modeling, which we found had a higher statistical power compared to conventional multivariate pattern decoding analysis. Behavioral data from 49 participants indicate that word predictability (i.e., context present vs. absent) facilitated orthographic and lexical-semantic, but not visual or phonological processes. Magnetoencephalography data from 38 participants show sustained activation of orthographic and lexical-semantic representations in the interval before processing the predicted stimulus, suggesting selective pre-activation at multiple levels of linguistic representation as proposed by predictive coding. However, we found more robust lexical-semantic representations when processing predictable in contrast to unpredictable letter strings, and pre-activation effects mainly resembled brain responses elicited when processing the expected letter string. This finding suggests that pre-activation did not result in 'explaining away' predictable stimulus features, but rather in a 'sharpening' of brain responses involved in word processing.

Impact Statement

Visual word recognition is facilitated in predictive contexts. Predictive coding postulates that context-dependent facilitation involves the pre-activation of expected stimulus features, but it is not clear on which linguistic representations this mechanism relies during word recognition. Combining magnetoencephalography with high-powered linear mixed modeling, we show that context-dependent facilitation relies on pre-activation of orthographic and lexical-semantic representations in neuronal signals before actually perceiving an expected word.

1 Introduction

Predictive contexts can facilitate word processing, in the sense of increasing reading speed (e.g., Rayner et al., 2011) and decreasing neuronal activation elicited during recognition of expected words (as reflected, for example, in the reduction of the N400 component of the event-related brain potential; e.g., Kutas and Hillyard, 1984; Kutas and Federmeier, 2011). These often-replicated findings have been linked to domain-general theories that postulate active, top-down (i.e., from hierarchically higher to lower processing levels) prediction of expected stimulus characteristics before actually perceiving the stimulus (predictive coding; cf. Rao and Ballard, 1999; Friston, 2005). In line with this, several neuro-cognitive models of visual word recognition (e.g., Seidenberg and McClelland, 1989; Carreiras et al., 2014) assume context-based prediction across multiple levels of linguistic processing, and it has been hypothesized that hierarchical predictions during reading involve the pre-activation of visual, pre-lexical (i.e., orthographic or phonological), and lexical-semantic representations of predicted words (Federmeier, 2007; Kuperberg and Jaeger, 2016). However, this proposal has not been systematically tested because currently, available evidence does not unambiguously differentiate between predictive pre-activation of representations at these different linguistic processing stages.

First studies showed that language-related brain regions can be activated before a highly expected stimulus appears (Dikker and Pylkkanen, 2011; Bonhage et al., 2015; Wang et al., 2017), but did not further assess the nature of pre-activated representations. More recent work showed pre-stimulus effects of semantic category (Heikel et al., 2018; Wang et al., 2020) using multivariate pattern analysis techniques (King et al., 2018; Kragel et al., 2018), as well as pre-stimulus effects of word frequency (Fruchter et al., 2015) using linear mixed models (LMMs; Baayen et al., 2008). These studies provide initial evidence that lexical-semantic representations are pre-activated before a predictable word appears. Also, several studies manipulating lexical-semantic context (e.g., the sentence context preceding the target word) found context-based modulations of brain activation in time windows associated with visual or pre-lexical processing (e.g., Lee et al., 2012; Brothers et al., 2015), which however provides only indirect evidence for predictive processing at other levels of linguistic representation than the lexical-semantic level. Studies that investigated pre-lexical context effects more directly found no (Eisenhauer et al., 2019) or only minimal effects (Nieuwland et al., 2018; Nicenboim et al., 2020), indicating that more statistically robust approaches may be needed for identifying the processes involved in pre-lexical pre-activation (see also Nieuwland, 2019, for a review).

Here, we combine behavioral and magnetoencephalography (MEG) data elicited during processing of words and (orthographically legal and pronounceable) pseudowords, to investigate the mechanisms of predictive pre-activation at multiple levels of linguistic representation, i.e., visual, pre-lexical (orthographic and phonological), and lexical-semantic. Contextual predictability was explicitly controlled by using a repetition priming paradigm. In a first step, data from a behavioral experiment were used to determine whether context-dependent facilitation ('priming') interacts with quantitative metrics from psycholinguistics representing different stages of word processing, i.e., visual stimulus complexity (early visual processing; e.g., Pelli et al., 2006), orthographic word similarity (pre-lexical orthographic processing; e.g., Yarkoni et al., 2008), the number of syllables (pre-lexical phonological processing; e.g., Álvarez et al., 2010), and word frequency and stimulus lexicality (lexical-semantic processing; e.g., Forster and Chambers, 1973; Fiebach et al., 2002). Subsequently, MEG activity from an independent experiment was explored strictly for those metrics that interacted in behavior with contextual predictability (primed vs. unprimed), with a particular focus on whether or not neurophysiological signatures of context-dependent pre-activation can be detected in the prime-target delay. This procedure prevented the investigation of neurophysiological effects without a behavioral counterpart, which would be difficult to interpret (Krakauer et al., 2017). High-powered LMMs with crossed random effects (Baayen et al., 2008) were used for both behavioral and MEG analyses. As a critical test case for a mechanistic contribution of the respective linguistic processing stage to context-dependent predictability, we hypothesized that the strength of pre-activation effects in the delay should be inversely related to the strength of neural effects measured during processing of the target.

2 Methods

2.1 Behavioral experiment

2.1.1 Participants

Forty-nine healthy, right-handed, native speakers of German recruited from university campuses (33 females, mean age 24.7 ± 4.9 years, range: 18–39 years) were included in the final data analyses. All participants had normal or corrected-to-normal vision, and normal reading abilities as assessed with the adult version of the Salzburg Reading Screening (the unpublished adult version of Mayringer and Wimmer, 2003). Further participants were excluded before the experiment due to low reading skills (i.e., reading test score below 16th percentile; N = 3) or participation in a similar previous experiment (N = 1), and during the

course of the experiment due to failure to complete the experimental protocol (N = 8) and because of an experimenter error (mix-up of the pseudoword lists during the pre-experiment familiarization procedure; N = 1). All participants gave written informed consent according to procedures approved by the local ethics committee (Department of Psychology, Goethe University Frankfurt, application N° 2015-229) and received 10 € per hour or course credit as compensation.

2.1.2 Stimuli and presentation procedure

60 words and 180 pseudowords (five letters each) were presented (black on white background, 14 pt., 51 cm viewing distance) in a repetition priming experiment consisting of two priming blocks and two non-priming blocks (120 trials each; Fig. 1a) with a total duration of ~20 min. For the present study, we analyzed a subset of 60 pseudowords that were unfamiliar to the participants ('novel pseudoword' condition). However, note that around 90 min prior to the priming experiment described here, these pseudowords had been presented to the participants in another experiment not of interest for the present study, without any learning instruction. In addition, the experiment contained two further lists of 60 pseudowords each, which were familiarized prior to the experiment as described in Eisenhauer et al. (2019; behavioral experiment). The assignment of pseudowords to the novel condition was varied across participants so that three different lists of novel pseudowords were used in the present study.

In priming blocks, a prime and a target stimulus were presented for 800 ms each, separated by a delay period of 800 ms during which a string of five hash marks was presented. Prime and target were identical in 75 % of trials. In non-priming blocks, a single letter string was presented for 800 ms in each trial. The inter-trial-interval was jittered between 800 and 1,200 ms (mean: 1,000 ms). Participants were asked to fixate the space between two vertical black bars at the center of the screen. Upon presentation of a letter string between the two lines, they had to indicate as quickly and accurately as possible whether it had a semantic association or not (which was the case for 50% of items, i.e., for words and for one list of pseudowords that had been semantically familiarized prior to the experiment). For simplicity, these judgments will be called lexical decisions in the following. The novel pseudowords used for the present analyses had no semantic associations. In priming blocks, participants responded only to the second (i.e., the target) letter string. Response hands and the order of blocks were counterbalanced across participants. Each letter string was presented in one priming trial and one non-priming trial.

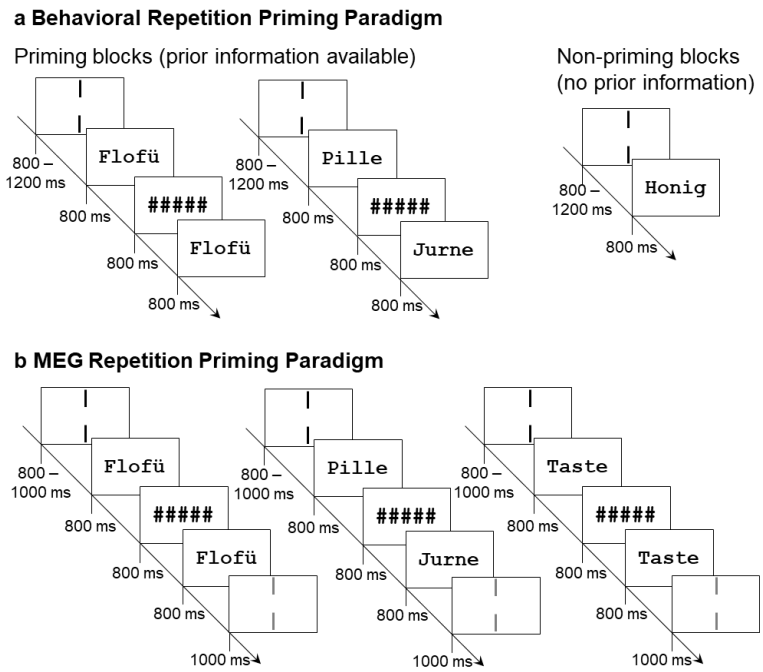

Figure 1. Experimental procedures. (a) Behavioral repetition priming paradigm. Two priming blocks (left) and two non-priming blocks (right) were presented in alternating order. In priming blocks, each trial consisted of a prime and a target stimulus presented for 800 ms each, separated by an interval of 800 ms during which a string of five hash marks was presented. Stimuli could be words or pseudowords. 75 % of trials were repetition trials with identical prime and target, while in the remaining 25 % two different letter strings were presented (non-repetition trials; not analyzed). In this case, prime and target could be from the same or from two different conditions, with all combinations of conditions appearing equally often. In non-priming blocks, only one word or pseudoword was presented in each trial. Participants were instructed to respond on each trial whether or not they had a semantic association with the target in priming blocks or with the isolated item in non-priming blocks. Before onset of the prime or the isolated item, two black vertical bars presented for 800 – 1,200 ms indicated the center of the screen where participants were asked to fixate. Context effects were investigated by comparing isolated items from the non-priming blocks with the targets from the repetition trials. (b) Repetition priming paradigm during MEG recording. The presentation procedure was identical to the priming blocks in (a). There were no non-priming blocks. Additionally, after target offset two grey vertical bars were presented for 1,000 ms indicating a blinking period. Before the onset of the next trial, a blank screen was presented for 500 ms. Participants were instructed to silently read presented letter strings and to respond only to rare catch trials (i.e., presentation of the word *Taste*, Engl. *button*). Context effects were investigated by comparing primes to repeated targets.

Words and pseudowords were matched between lists (i) for orthographic similarity ('word likeness') using the Orthographic Levenshtein Distance 20 (OLD20; Yarkoni et al., 2008) based on the SUBTLEX-DE database (Brysbaert et al., 2011) and (ii) with respect to the number of syllables (computed via Balloon, cf. Reichel, 2012; see also Table 1). Other psycholinguistic metrics of interest for our analyses were logarithmic word frequency and trigram frequency (i.e., the mean frequency of each trigram per word; obtained from SUBTLEX-DE), as well as visual complexity measures (perimetric complexity and the number of simple features), which were obtained for each letter from the GraphCom database

(Chang et al., 2018) and averaged across the five letters of each stimulus (Table 1). The other two visual complexity parameters from GraphCom, i.e., the number of connected points and the number of disconnected components, were not chosen for our analyses, the former due to its high correlation with the number of simple features (>.8 in our stimuli) and the latter due to its low variance across letters of the German language. For parameter correlations within words and pseudowords, see Figure S1 in Supporting Information.

2.1.3 Analyses

Statistical Modeling. A linear mixed model (LMM) was used to investigate the three-way (as well as subordinate two-way) interactions of each of the four visual and pre-lexical word parameters (see below) with the factors context (i.e., the priming effect of primed vs. unprimed stimuli) and stimulus lexicality (words vs. pseudowords) in log-transformed response times, using the lmerTest package (Kuznetsova et al., 2017) of the statistical software package R, version 3.5.3, 2019-03-11 (R Development Core Team, 2008). In the case of the word frequency parameter, only the interaction with context was included, as an interaction with lexicality is not possible (all pseudowords were assigned a word frequency of zero). Note that in the behavioral study, the factor context was operationalized as the contrast between repeated targets in priming blocks vs. single items in non-priming blocks. Therefore, 'context' here represents the effect of the presence vs. absence of contextual information on processing of the target stimulus, whereby only valid contextual information was considered while trials in which the target was preceded by a non-identical prime were discarded (analogous to the MEG experiment; see below). As a consequence, the priming condition had fewer trials (.75 x 60 = 45; minus errors) than the non-priming condition. However, LMMs with crossed random effects are optimal for the analysis of imbalanced data (Baayen et al., 2008). The two-way interaction terms between word parameters (e.g., word frequency) and context were used to determine whether the effect of the respective word parameter was modulated by a predictive context. The three-way interaction with lexicality additionally revealed whether the 'context by word parameter' modulations differed between words and pseudowords. This allowed us to assess whether context-based facilitation at the respective level relies on prior knowledge, which is available for words but not pseudowords. Note that we included these interaction terms, i.e., with context and lexicality, for all word parameters within one single LMM. Trials in which errors occurred (9.8 %) were excluded from analyses. We know from previous experience that trial order can have a strong effect on response times and neuronal activation. To explicitly account for this, trial order was included

into the LMMs as fixed effect. Participant and item were included as random effects on the intercept. For visualization of partial effects, we used the *remef* package (Hohenstein and Kliegl, 2015).

Investigated word parameters. Different descriptors of words were chosen in order to isolate different 'levels' of processing a word. *'Early visual processing'* of a written word depends on the complexity of its physical appearance, which we here characterize, following Chang et al. (2018), using perimetric complexity (describing the density of black pixels in relation to white background which has previously been associated with letter identification efficiency; see Pelli et al., 2006) and the number of simple features that make up a word (i.e., the number of strokes per letter). *'Pre-lexical processing'* of written words comprises phonological and orthographic processing (e.g., see Carreiras et al., 2014). Phonological processing can be captured by the number of syllables of a word (as syllables reflect sublexical units for sequential phonological processing; e.g., Álvarez et al., 2010; Chetail, 2014), and orthographic processing is captured by the orthographic Levenshtein distance 20 (OLD20, Yarkoni et al., 2008) and by trigram frequency (e.g., Colegate and Eriksen, 1972; Chen et al., 2015). Lastly, lexical processes of word identification are often associated with word frequency (e.g., Forster and Chambers, 1973; Fiebach et al., 2002), so that logarithmic word frequency (Brysbaert et al., 2011) is included as a parameter representing *'lexical-semantic processing'*. Pseudowords were assigned a word frequency of zero as they did not appear in the SUBTLEX database. Finally, we included a binary lexicality contrast (words vs. pseudowords) comparing items with and without semantic associations investigating lexical-semantic processing.

Model comparisons. In cases with non-significant interaction effects on one or more processing levels, we repeated the analysis with a simpler model in which only the significant interactions and main effects were included. This sparser model was compared to the full model based on the difference in the Akaike information criterion (AIC). The AIC allows comparing models of different complexity (i.e., with more or fewer parameters included). A significantly lower AIC for a more complex model indicates an increase in model fit with the newly added parameter. If the AIC difference is positive or equal, the sparser model has a better fit, and the addition of the new parameter is not advised. Our results and interpretations will be based on the model that includes the set of parameters that lead to an optimal fit.

2.2 Magnetoencephalography (MEG) experiment

2.2.1 MEG data

We used data from a previously published repetition priming study (Eisenhauer et al., 2019; data publicly available from https://osf.io/fc69p/), comprising MEG recordings from 38 healthy right-handed native speakers of German (26 females) between 18 and 31 years of age and with normal reading abilities. We here provide a short description of the data acquisition and experimental procedures, and refer the reader to our earlier paper for a detailed description. Data were recorded using a MEG system with 269 operating third-order axial gradiometers (Omega 2005; VSM MedTech Ltd., Coquitlan, BC, Canada) at a sampling rate of 1,200 Hz, applying online filtering between 0.1 and 300 Hz. Procedures were approved by the local ethics committees (University Clinic of Goethe University Frankfurt, application N° 107/15; and Department of Psychology, Goethe University Frankfurt, application N° 2015-229).

The experimental paradigm was similar to the behavioral experiment described above: During MEG acquisition, participants silently read words or pseudowords of five letters length (black on white background; 14 pt at 51 cm viewing distance), which were presented in a repetition priming paradigm with 75 % identical prime-target pairings (repetition trials) and 25 % non-identical prime-target pairings (non-repetition trials; Fig. 1b). Prime and target stimuli were presented for 800 ms each, separated by a delay period of 800 ms during which a string of five hash marks was presented. Primes could either be real words, novel pseudowords (as defined above), or familiarized pseudowords, and each of these conditions consisted of 120 items, i.e., 60 stimuli presented twice. To ensure the participants' attention to the presented letter strings, 80 additional catch trials were randomly interspersed into the trial sequence, during which the word *Taste* (Engl. *button)* indicated the requirement to press a button. In total, 440 trials were presented across three blocks with a total duration of ~40 min. Only the word and novel pseudoword conditions were included into the present analyses. From these, 120 trials per letter string condition were considered for analyses of prime and delay intervals whereas 90 items per condition were used for analyses of the target interval (as only 75% of trials were repetition trials; see also below). Words and pseudowords were matched for OLD20 and number of syllables (see Table 1, also including further word parameters, and Fig. S1 in Supporting Information for parameter correlations).

For source localization, structural magnetic resonance (MR) images obtained with a 1.5 T Siemens magnetom Allegra scanner (Siemens Medical Systems, Erlangen, Germany) using a standard T1 sequence (3D MPRAGE, 176 slices, 1 x 1 x 1 mm voxel size, 2.2 s TR,

3.93 ms TE, 9° flip angle) were available for 34 participants (Eisenhauer et al., 2019). The MR images contained fiducial markers for the two auricular MEG head localization coils to improve co-registration with the MEG data.

2.2.2 Preprocessing

Initial preprocessing steps were performed with the FieldTrip toolbox, version 2013 01-06 (http://fieldtrip.fcdonders.nl; Oostenveld et al., 2011) under MATLAB (version 2012b, The MathWorks Inc., Natick, MA). Parallel computations were performed using GNU Parallel (Tange, 2011). MEG data were segmented into epochs of 2,600 ms length, lasting from -200 ms to 2,400 ms with respect to the onset of the prime. Trials contaminated with sensor jump or muscle artifacts, as well as trials in which the head position deviated more than 5 mm from the participant's average head position across all trials, were rejected, and trials contaminated with eye blink, eye movement, or heartbeat artifacts were cleaned using Independent Component Analysis (ICA, Makeig et al., 1996). See Eisenhauer et al. (2019) and our analysis scripts (https://osf.io/fc69p/) for detailed procedures. After preprocessing, an average of 69 trials (range: 26 to 79) per condition remained for analyses involving both repetition and non-repetition trials (see below), and an average of 51 trials (range: 37 to 106) remained for analyses including only repetition trials. Data were low-pass filtered at 20 Hz and baseline corrected using the time window from -110 to -10 ms; all samples between -110 and 2,400 ms were used for subsequent analyses. To reduce computational costs, the MEG data were down-sampled to 40 Hz.

2.2.3 Linear mixed model analyses

As LMMs are beginning to be used also in the analysis of electrophysiological data (e.g., Fruchter et al., 2015; Payne et al., 2015), we analysed the MEG data with LMMs as well. LMMs were computed on event-related field (ERF) values separately for each time point of the baseline, prime, delay, and target time windows. For the analysis of the target time window, we excluded non-repetition trials (in which targets were not predictable) in order to focus on predictable targets only. For the remaining analyses (i.e., concerning prime and delay interval) we included repetition and non-repetition trials as these did not differ between repetition and non-repetition trials. The models were computed using the lmerTest package (Kuznetsova et al., 2014) in R, which allows computing *p-values* for LMMs (Luke, 2017). Resulting *p-values* were Bonferroni-corrected for multiple comparisons to a critical *p-value* of .05/101 = .000495. Following Krakauer et al., (2017), we focused on those effects for which a

behavioral correlate was identified in the independent behavioral experiment, i.e., the *pre-lexical/orthographic level of representation* (reflected in a context x OLD20 interaction) and the *lexical-semantic level of representation* (reflected in context x lexicality and context x word frequency interactions; see Results below). Thus, we included as fixed effects only word frequency and the interaction of OLD20 with lexicality. As in the analysis of the behavioral experiment, trial order was included as additional fixed effect of no interest, and the random effects on the intercept of participant and sensor were estimated. Based on the logic of parsimonious mixed modeling (Bates et al., 2018), the random effect of item was not included, because including it resulted in a high number of un-estimable models (i.e., non-convergence, unidentifiability, or singluar fit). Still, 13.9 % of models (i.e., of the 101 models calculated for the different time points of the time windows of interest) were un-estimable and thus treated as if they were not significant. Word frequency, OLD20, and trial order were *z*-transformed to enhance the likelihood of convergence and the accuracy of parameter estimates (e.g., Harrison et al., 2018). Post hoc LMMs were computed separately for words and pseudowords to resolve interactions with lexicality as well as to obtain the frequency effect for words only; here, 7.9 and 5.0 % of models were un-estimable, respectively.

To explicitly investigate the influence of context (i.e., prime vs. target) on the effects of interest, we combined data from prime and target time windows. We performed this analysis for each of the 32 time points of prime/target presentation, Bonferroni-corrected to a critical *p-value* of .05/32 = 0.0015625. For the interaction of context and lexicality, we computed the full model controlling for effects of the frequency by context interaction and the three-way interaction of lexicality with OLD20 and context (18.8 % un-estimable). For the interaction of context and word frequency or OLD20, we computed separate post hoc models for words and pseudowords (21.9 and 9.4 % un-estimable, respectively). All models included trial order as an additional fixed effect, and participant and sensor as random effects on the intercept.

To assess if the strength of pre-activation (i.e., effects of psycholinguistic metrics detected as significant in the delay period) is related to the processing strength of unpredictable (i.e., the prime) as well as predictable letter strings (i.e., the target) at time windows that show a context effect, (i.e., a change in neural activation from prime to target), we computed pairwise Pearson correlations between significant time points from the delay interval with prime and target time points showing a significant interaction with context (using LMM-based partial effects estimates). In the initial analysis, we estimated the effect of interest (lexicality, word frequency, and OLD20 of words and pseudowords) for the specific

time points at which the interaction with context was significant. We then aggregated the partial effect estimates across participants, sensors, and repeated item presentations before we estimated the correlations. However, in this initial analysis, we noted that almost all correlations were positive, even those between positive and negative effects, indicating that the correlations might be confounded by residual noise not representing the effect of interest. We therefore first performed an LMM analysis modeling the estimated effect of the parameter of interest dependent on the parameter as well as on the random effect of item. Remember that we were not able to include the random effect of item in the original LMM analysis as this resulted in a high proportion of un-estimable models. Next, we re-estimated the effect of interest, partialing out the random effect of item. We then again performed the correlation analysis as described above, using the re-estimated effects. Indeed, in this analysis the sign of the correlation effects represented the sign of the observed effects, i.e., positive effects correlated negatively with negative effects. Obtained *p-values* were Bonferroni-corrected for the number of computed correlations (which could differ between the four psycholinguistic metrics). E.g., for lexicality, six time points were found significant during the delay, and seven time points showed a significant interaction with context, resulting in 190 correlations and consequently a Bonferroni-corrected critical *p-value* of .05/190 = 0.00026.

2.2.4 MEG source localization

We estimated source locations from the ERF difference of words vs. pseudowords, high vs. low-frequency words, and high vs. low OLD20 (the latter separately for words and pseudowords). Word frequency and OLD20 contrasts were based on median splits. As time windows for source analysis we selected the significant time points of lexicality effects, word frequency effects (words only), and the OLD20 effects for words and pseudowords from the LMM analysis (see previous sections). Source localizations of effects obtained in the LMM analysis were conducted on the effects of interest after correcting for all other sources of variance in the model. To this end, we simulated the respective effect of interest from the LMM for each sensor in a way that controls all other effects, using the *remef* package (Hohenstein and Kliegl, 2015). This source localization procedure is optimal when for each time point, all data from all sensors is included in one model, which is possible since we treated the sensors as a random effect. The simulated data from the LMM projects the effect of interest onto the sensors after correcting for all confounding variables, and these simulated sensor data are then subjected to standard source reconstruction procedures as described below. This ‘preprocessing’ procedure for source localization is a central feature of our LMM

analysis and currently, to our knowledge, not used elsewhere.

Source localization was performed using linearly constrained minimum variance (LCMV) spatial filters (Van Veen et al., 1997) in FieldTrip, Version 2016 10-24, for those 34 participants for whom anatomical MR images were available. The procedure closely followed our earlier study (Eisenhauer et al., 2019) and is based on Manahova et al., (2018); see also Dijkstra et al., 2018; Mostert et al., 2018). The approach comprised a permutation procedure with 1,000 analyses on shuffled data to account for noise. The signal of each source location was normalized by its variance to counter the depth bias. Please refer to Eisenhauer et al. (2019) for a detailed description of the general procedure.

For visualization, LMM-based source localizations thresholded at 50 % of the peak activation across all time windows included in the respective contrast, as well as thresholded at 90 % of the individual peak of the respective time point, were plotted on cortical surfaces using MRIcron (Rorden and Brett, 2000). In order to make the OLD20 source activation strengths comparable between words and pseudowords, both were thresholded at 50 % of the peak activation for words. Also, we visualized source overlaps between the identified sources of prime, delay, and target, based on the 50 % threshold. Brain regions were identified from the MNI coordinates of the ≤3 most robust sources per contrast using the AtlasReader (Notter et al., 2019) in Python.

2.3 Code and data accessibility

The analysis code used for behavioral and MEG data analyses, as well as the behavioral data, are publicly available from [to be filled in upon publication]. The MEG and MRI data are publicly available from https://osf.io/fc69p/.

3 Results

3.1 Behavioral experiment: Identifying context-based facilitation effects

In the data from the behavioral experiment, we investigated the influence of a predictive context (i.e., prime absent vs. present) on visual, orthographic, phonological, and lexical-semantic processes during visual word recognition. These different ‘levels’ of processing were approximated in a linear mixed model (LMM) analysis of response times by various psycholinguistic metrics. We thus modeled the influence of context on effects of (i) perimetric complexity and number of simple features of the stimuli (visual processing), (ii) OLD20 and trigram frequency (orthographic processing), (iii) number of syllables (phonological

processing), and (iv) word frequency as well as lexicality (lexical-semantic processing).

As our hypotheses were focused on the interactions of these stimulus characteristics with context-dependent facilitation, we will in the following primarily report results involving context effects; full results of the LMMs can be found in associated Tables. In general, processing of expected items (i.e., targets preceded by an identical prime stimulus) was facilitated, i.e., they were responded to faster than unprimed items (significant main effect of context: $p < 2e\text{-}16$; Table 2). This context-dependent facilitation interacted with parametric predictors representing orthographic (i.e., OLD20) and lexical-semantic (i.e., word frequency and lexicality) but not visual or phonological processing (Table 2; see also Table S1 for the results of the full model). In order to implement an adequate statistical model, we removed the non-significant visual, orthographic (i.e., trigram-frequency), and phonological predictors from the full model and re-estimated the model. This reduced model is presented in Table 2; an explicit model comparison indicated that the Akaike Information Criterion (AIC) of this simpler model was 21 points lower than that of the full model (*chi-square* = 11.051, *df* = 16, *p* = 0.806). Given that these two models did not differ statistically, the simpler model should be preferred (see Methods section).

With respect to *orthographic processing*, a significant 3-way interaction between context, lexicality, and OLD20 ($p = .001$; Table 2; see also Tables 2-2 and 2-3 for resolving this interaction post-hoc via two-way interactions) could be resolved to show significant negative effects of OLD20 on response times for pseudowords, with faster decision times particularly for expected pseudowords that have a high orthographic similarity to words (i.e., high OLD20; see Fig. 2a, Table 3, and Table S2). For words, we observed descriptively a negative effect of OLD20 for words presented in isolation and a positive effect for primed words (Fig. 2a), which resulted in a significant context by OLD20 interaction for words (see post-hoc statistics in Table S2). This interaction could, however, not be further resolved statistically in post-hoc analyses (Table 3).

With respect to *lexical-semantic processing*, a significant interaction between context and word frequency ($p < .001$; Table 2) provided evidence for context-dependent facilitation of lexical-semantic processing. Post-hoc evaluation showed significantly faster decision times to more frequent relative to less frequent words when presented in isolation (i.e., when not predictable), but no significant frequency effect for primed words (given generally faster responses; Fig. 2b; see also Table 3 for word frequency effects of words only, and Table S3 for word frequency effects combined across words and pseudowords, for which a word frequency of zero was assumed). Lexical-semantic processing as one locus of predictive

effects is also supported by the observation of a context by lexicality interaction ($p < .001$; Table 2), which shows that the effect of context-dependent facilitation is generally much stronger for words than for pseudowords (see also Fig. 2a; compare also Eisenhauer et al., 2019). To summarize, these behavioral results suggest that context-based facilitation involves processes at the pre-lexical orthographic and the lexical-semantic stages of visual word recognition, while no evidence was found for context effects at the level of visual or pre-lexical phonological processing. As a consequence, we explore the neurophysiological data with respect to neural signatures of predictive context effects and pre-activation effects of OLD20 (orthographic similarity), word frequency, and lexicality.

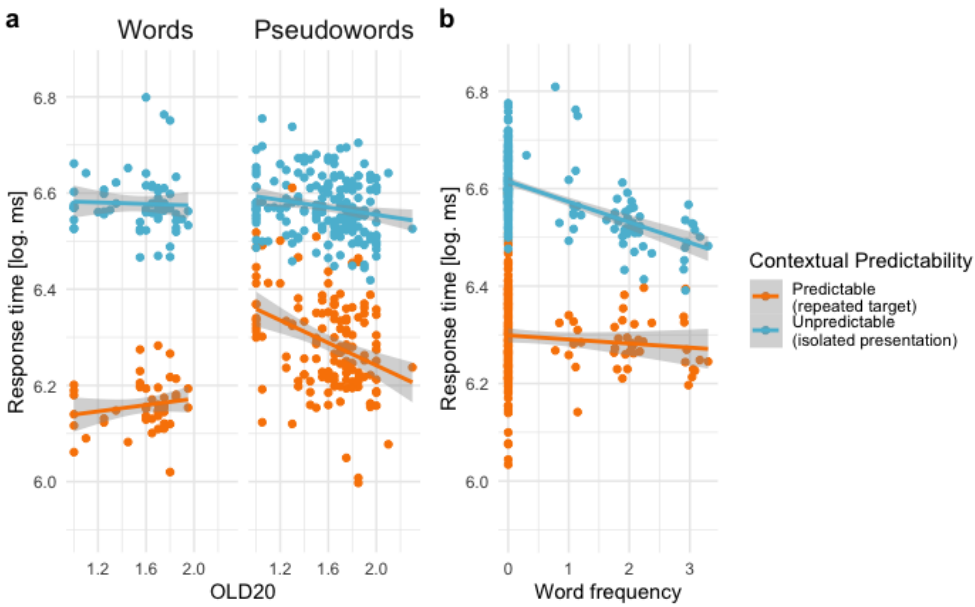

Figure 2. Behavioral results reflecting context-dependent facilitation. Comparison of response times representing (a) orthographic (OLD20 effect separated for words and pseudowords) and (b) lexical processing (word frequency effect) when stimuli were either predictable from context (i.e., repeated targets; orange lines and dots) or not (isolated presentation; blue lines and dots). Note that lower OLD20 values reflect higher word likeness, and, thus, orthographic familiarity. Displayed are logarithmic response times that represent the partial effects estimated from linear mixed models. Dots represent items averaged across participants.

3.2 Investigation of orthographic and lexical-semantic representations in brain data

MEG-measured event-related fields elicited by words and pseudowords were explored in depth in Eisenhauer et al. (2019). The present work goes beyond this initial report and uses LMMs to directly assess (i) context effects during stimulus processing and their effect on orthographic and lexical-semantic levels of linguistic processing (represented by the psycholinguistic metrics OLD20, word frequency, and lexicality), and (ii) their dependency on pre-activation of orthographic and lexical-semantic representations during the prime-target

interval. To this end, we first characterize the temporal dynamics of the effects of OLD20 (reflecting orthographic representations), as well as word frequency and lexicality (both reflecting lexical-semantic representations) by individually analyzing each time point across the entire trial. We then directly compare respective effects identified during processing of prime and target, to quantify context (i.e., repetition) effects at the respective level of representation. Third, we explore pre-activation by testing for the presence of orthographic and lexical-semantic effects in the prime-target delay. In detail, we assess whether these signatures of pre-activation correlate with activation during prime and target processing, primarily to determine how the processing of expected stimuli mechanistically depends on context-dependent pre-activation.

Analogous to the statistical approach used for analyzing the data from the behavioral experiment, we investigated the effects of the three psycholinguistic metrics of interest on brain activation using linear mixed models (LMMs). LMMs model data at the level of individual trials, and thus result in high statistical power (Matuschek et al., 2017). Indeed, while we also investigated brain activation using a multivariate pattern decoding approach (see Supporting Results 2), which has been established in previous work for investigating brain responses in the absence of external stimulation (e.g., Simanova et al., 2015; Heikel et al., 2018), a power analysis indicated higher power and more reliable effect size estimates for LMMs in contrast to multivariate pattern decoding (see Supporting Results 3). Importantly, LMMs can estimate the effects of interest while controlling for confounding variables. Note that LMMs are calculated for every sample time point; the LMM-estimated effect sizes reflect a modulation of the ERF amplitude at the respective time point. In the presence of a negative ERF component, a positive effect size thus indicates that letter strings with a higher value on the predictor variable (e.g., word frequency) have a less negative-going ERF, and vice versa for negative effect sizes.

Effects of orthographic representations. LMMs revealed a significant interaction between OLD20 and lexicality – corresponding to the interaction found in our behavioral data – in several time windows (Bonferroni-corrected; marked with yellow in Fig. 3a,d). The interaction was found (i) during presentation of the prime between 215 and 265 ms and between 540 to 790 ms, (ii) during the prime-target delay at multiple time points between 15 and 740 ms after prime offset (corresponding to 815 to 1540 ms relative to prime onset), and (iii) during presentation of the target, i.e., between 190 and 215 ms, as well as 590 and 690 ms after target onset (corresponding to 1790 to 1815 and 2190 to 2290 ms relative to prime onset). The interaction resulted from predominantly negative effects for words (i.e., greater

negative-going MEG activity the higher the OLD20 parameter, i.e., the less word-like a word was; see detailed effects in Fig. 3a) and mainly positive effects for pseudowords (Fig. 3d). The opposite effects of orthographic similarity (OLD20) on MEG responses elicited during word vs. pseudoword processing are also documented by a negative correlation of the two time courses of the OLD20 effect estimates ($r = -.29$; $t(99) = -3.0$; $p = .003$). Separate investigations of OLD20 effects for words vs. pseudowords revealed an early negative OLD20 effect for pseudowords (i.e., at 40 ms) during prime presentation (Fig. 3d). In contrast, the OLD20 effect for words reached significance considerably later, at 215 ms during prime presentation (Fig. 3a). During target processing, OLD20 effects reached significance shortly before 200 ms, with a slightly earlier onset for pseudowords (165 ms) vs. words (190 ms). In the delay period, significant effects were more abundant for words (40.6 % of investigated time points) vs. pseudowords (18.8 %). When comparing the three time windows, words showed overlapping source activations in bilateral frontal and pseudowords in bilateral parietal regions (Fig. 3a,c vs. d,f).

When quantifying the contextually-mediated effect of orthographic processing by directly contrasting OLD20 effects during prime and target processing, significant differences emerged at 215, 540, and 590 to 615 ms for words and at 115, 190, 265 to 290, and 740 ms for pseudowords, reflecting a reversal of the effect direction from prime to target at these time points (i.e., from negative to positive or vice versa; marked by gray shading in Fig. 3a and Fig. 3d; see also Fig. 3b and Fig. 3e).

To examine the influence of orthographic pre-activation (i.e., the OLD20 effects) on target processing, we correlated significant effect size estimates from the delay with all timepoints showing significant interactions with context (i.e., that show a relevant change from prime to target). First, we expect that time points within the delay period correlate positively when neuronal pre-activation is maintained stable across time (cf. King and Dehaene, 2014). Positive correlations between prime and delay might either reflect a ‘spill-over’ from prime processing, or a re-activation of representations activated while processing the prime, potentially facilitating the processing of the identical target. We expect a negative correlation between delay and target, when the higher delay activation results in more efficient target processing (i.e., as assumed by predictive coding; Friston, 2005). In contrast, if the observed reversal of the effect direction from prime to target already takes place during the delay, this could indicate that more extensive processing of the prime leads to more efficient pre-activation and target processing.

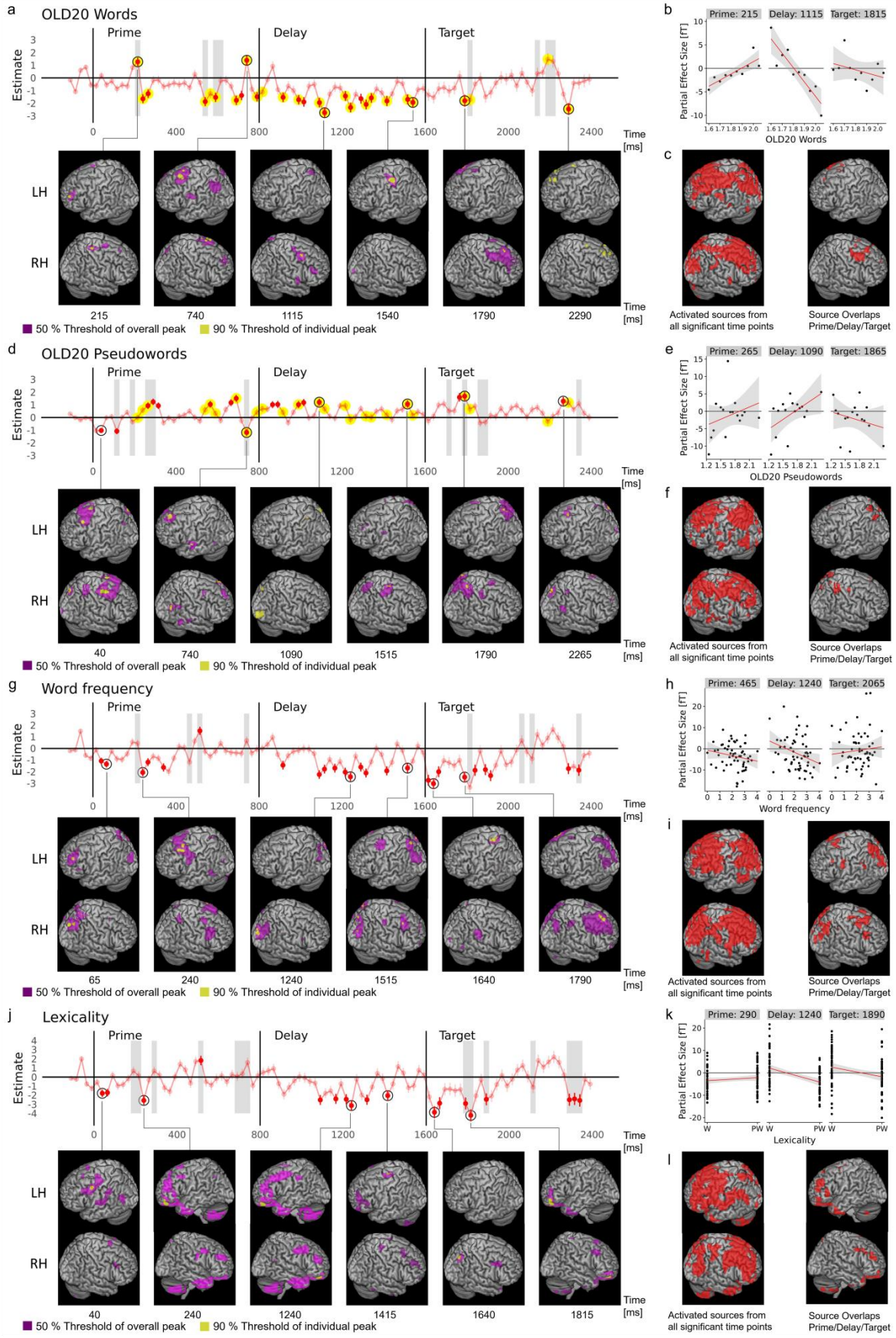

Figure 3. Effects of orthographic and lexical-semantic information across time and their brain localization. (a) Effects of OLD20 during word processing. Displayed are the estimates of linear mixed models ± *SE* across time, including baseline, prime, delay, and target time windows. The

salient red dots represent significant time points, excluding un-estimable models. Gray shading represents time points of a significant interaction between OLD20 and context (prime vs. target). Yellow circles represent time points of a significant interaction between OLD20 and lexicality (words vs. pseudowords; compare panel (b)). Black circles mark time points for which source localizations are shown. Source topographies display sources thresholded at 50 % of the overall peak (maximal activation across all time points; violet), as well as thresholded at 90 % of the individual peak per time point (yellow) within the left (LH) and right (RH) hemisphere. (b) Scatterplots depict model-estimated partial effect sizes for three exemplary time points from prime, delay, and target time windows. In case of prime and target, a time point showing a significant interaction between the respective parameter and the prime/target difference was selected. (c) Activated sources from all significant time points (left) and overlapping sources during prime, delay, and target time windows (right), thresholded at 50 % of the overall peak. (d,e,f), (g,h,i), and (j,k,l) show corresponding results for OLD20 during pseudoword processing, word frequency (for words only), and lexicality, respectively. To facilitate comparability of OLD20 source activation strengths for words and pseudowords, the 50 % threshold of OLD20 source activations for words was used for pseudowords as well. See Table S5 for source coordinates.

Within the delay period (see black triangles in Fig. 4a,b), significant effects of OLD20 are highly correlated, but more sustained across the delay for words than for pseudowords (Fig. 4a,b). This indicates a stable neuronal activation across the delay period for words, while pseudowords showed less orthographic pre-activation during the delay (see above). From prime to delay, the majority of significant correlations for words was positive (74.0 %), while significant correlations for pseudowords did not indicate a stable pattern as only 44.4 % were positive (see the dot-dashed black squares in Fig. 4a,b). Delay-target correlations (see the dashed line black squares in Figs. 4a,b) represent relationships between orthographic pre-activation and processing of the expected item. For words, early target activation (215 ms) was correlated positively, while late target activation (590 to 615 ms) was correlated negatively with orthographic delay activation (Fig. 4a). For pseudowords, the target activation was consistently positively correlated with the delay activation (Fig. 4b). To summarize, these findings indicate more sustained pre-activation of orthographic information for predicted words than pseudowords. However, pre-activated information more consistently resembled target effects for pseudowords compared to words; for words, early target effects resembled delay activation, while late target effects indicated stronger facilitation with stronger pre-activation.

Lexical-semantic representations: Effects of word frequency. As for OLD20, we found significant effects of word frequency (investigated only for the word condition) at several time points across the prime, delay, and target time windows (Fig. 3g,h). Throughout the entire trial, we found predominantly negative effects (i.e., more negative/less positive activation for higher frequency words), with the exception of one single positive effect at 515 ms after prime onset. Specifically, negative effects were found in the time ranges of 40 to 65

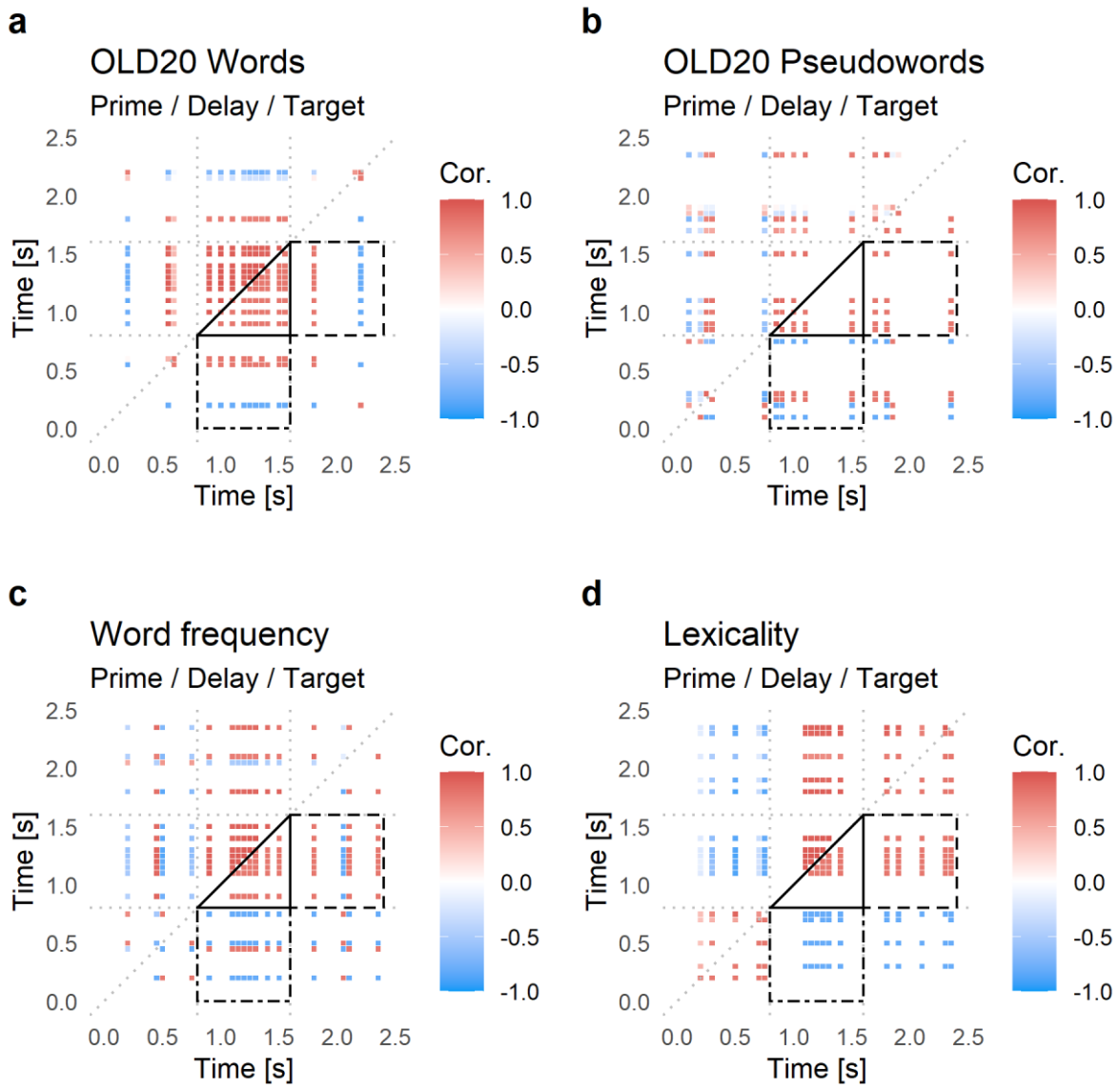

Figure 4. Correlations between LMM-estimated partial effects across prime, delay, and target time points for (a) OLD20 effects for words, (b) OLD20 effects for pseudowords, (c) word frequency, and (d) lexicality. For the delay, correlations were computed for all time points showing a significant effect of the respective letter string characteristic, while for prime and target, time points at which effects changed between prime and target (as reflected in a significant interaction of the characteristic with the context effect) were selected. The time-resolved correlations are down–sampled by a factor of 10. Significant correlations are shown below the diagonal, while all correlations are shown above the diagonal. Black triangles mark within-delay correlations, while the dot-dashed black square marks prime-delay correlations and the dashed line black square marks delay-target correlations.

ms and 240 to 340 ms after prime onset. During the prime-target delay, we observed negative effects at 115 ms and at several time points later in the delay, in the time range from 290 to 715 ms (i.e., 1090 to 1515 ms relative to prime onset). During the target presentation, significant negative frequency effects were found in time windows starting earlier when compared to the prime interval, i.e., between 15 and 65 ms. In addition, negative frequency

effects were found in the range from 190 to 315 ms (1615 to 1665 ms and 1790 to 1815 ms relative to prime onset), but also at late time points (690 and 740 ms, i.e., 2290 and 2340 ms relative to prime onset). Sources of the word frequency effects across the three time windows were localized to frontal and parietal brain regions (Fig. 3i).

To directly assess context-dependent word frequency effects, we compared word frequency effects between prime and target – statistically assessed as a context by frequency interaction. We identified significant interactions at 215, 465, 515, and 740 ms after stimulus onset (gray shading in Fig. 3g), indicating a reversal of the effect direction from prime to target. As for orthographic representations, we assessed correlations among effect strengths involving all time points with significant differences in frequency effects between prime and target (i.e., context x frequency interactions) and all significant time points from the delay period. Analogous to orthographic effects in the word condition, we found highly correlated effects within the delay period (Fig. 4c), suggestive of a consistent neuronal pre-activation of word frequency-related information. We found predominantly negative correlations when prime and delay periods were compared and predominantly positive correlations between pre-activation effects in the delay and target processing (75.0 % of significant correlations each). These findings suggest a reversal of the word frequency effect direction from prime to delay but relatively consistent neuronal activation patterns throughout the delay and target periods.

Lexical-semantic representations: Effects of lexicality. The time course of effect estimates for the second psycholinguistic metric of lexical-semantic processing, i.e., stimulus lexicality (word vs. pseudoword), was very similar to that of word frequency ($r = .94$; $t(99) = 26.7$; $p < .001$; compare Figs. 3g vs. j). In detail, during the presentation of the prime, negative lexicality effects (reflecting more negative activation for pseudowords relative to words) reached significance at 40 to 65 and at 240 ms, and we found a positive effect at 515 ms post-stimulus onset. During the delay, significant negative lexicality effects were found for several time points in the range from 290 to 615 ms (corresponding to 1090 to 1415 ms relative to prime onset). During the presentation of the target, significant negative effects were found at 40 to 65, between 190 and 290 ms, and from 690 to 740 ms.

Context by lexicality interaction effects partly overlapped with context interactions involving word frequency: Lexicality effects differed between prime and target at 190 to 215, 290, 515, and 690 to 740 ms after stimulus onset (gray shading in Fig. 3j; see also Fig. 3k). In the correlation matrices we found a very clear picture (Fig. 4d). First, all effects within the delay were highly positively inter-correlated. Again, the prime and delay period effects showed negative correlations and delay and target period effects showed positive correlations.

Across prime, delay, and target time windows, the lexicality effects were localized to frontal, temporal pole, and cerebellar regions. Thus, despite similar time courses, there is an apparent dissociation from the frequency effect (cf. Fig. 3g,i vs. j,l).

4 Discussion

We investigated the role of visual, orthographic, phonological, and lexical-semantic representations for context-dependent (i.e., predictive) facilitation of visual word recognition. Consistent with abundant previous literature (e.g., Tulving and Schacter, 1990; Kutas and Federmeier, 2011), context-dependent facilitation is reflected in faster response times and an amplitude reduction in the event-related fields (ERFs) for repeated relative to initial stimulus presentation (see Eisenhauer et al., 2019, Fig. 4, for ERFs from the present dataset). Our behavioral data indicate that orthographic and lexical-semantic, but not visual or phonological processing, interact with context-dependent facilitation. The MEG data show differential effects of orthographic word similarity, word frequency, and lexicality on prime vs. target processing. Importantly, all three parameters also modulated brain activity in the delay interval, providing evidence for pre-activation of orthographic and lexical-semantic representations in predictive contexts. This finding extends previous evidence that had indicated predictive pre-activation of lexical-semantic representations (Fruchter et al., 2015; Wang et al., 2020) without establishing corresponding evidence for orthography. A change in the direction of estimated effects reflected a change in the quality of pre-activated lexical-semantic representations from prime to delay, while delay and target effects were positively correlated.

Context-dependent facilitation of visual word recognition

Contextual facilitation was modulated by orthographic word similarity. As expected (e.g., Balota and Chumbley, 1984; Fiebach et al., 2007), we found fast non-word responses when orthographic familiarity is low. Predictability amplified this effect and reversed the weak OLD20 effect for words, suggesting an interactive role of orthographic processing and context-dependent facilitation. Regarding lexical-semantic processing, we found stronger behavioral priming effects for words than pseudowords (replicating, e.g., Fiebach et al., 2005) and reduced word frequency effects for expected words (replicating, e.g., Forster and Davis, 1984; Kinoshita, 1995). In contrast to the context-dependent amplification of orthographic non-word processing, contextual predictability abolished the word frequency effect. As word

frequency effects are typically interpreted as reflecting differences in the ease of lexical access (e.g., Coltheart et al., 2001), this (together with the generally faster target response times) indicates that the effort of accessing meaning is reduced in highly predictive contexts. Consistent with current literature, we found no evidence for context-based facilitation of visual (e.g., Slattery and Yates, 2018) or phonological processes (e.g., Nieuwland et al., 2018).

Neurophysiological Mechanisms of Context-Dependent Facilitation

Pre-lexical orthographic and lexical-semantic representations modulated brain activation during prime, delay, and target intervals. Orthographic familiarity elicited negative and positive effects for words and pseudowords, respectively, at around 200 and 500 ms. Lexical-semantic effects were identified earlier (before 100 ms) and around 250 ms, both showing a negative modulation (e.g., more positive activation for words than pseudowords) independent of predictability. This result (and its localization to frontal brain areas) replicates early frontal MEG effects obtained with similar stimuli (e.g., Wheat et al., 2010; Woodhead et al., 2014). Context-dependent changes of representations (i.e., from prime to target) started at ~100 ms post-stimulus onset for orthographic (in pseudowords) and ~200 ms for lexical-semantic representations. These prime-target differences also involved qualitative changes, with larger activations for words vs. pseudowords at the prime and the inverse effect at the target. These findings suggest that context-dependent facilitation is implemented at orthographic and lexical-semantic processing levels (e.g., Almeida and Poeppel, 2013; Brothers et al., 2015).

Orthographic as well as lexical-semantic effects were consistently positively correlated during the delay period, indicating the high stability of activated representations. Effects elicited during the delay reflect, by necessity, either a 'spill-over' from processing the prime or predictive pre-activation of target-associated linguistic representations. Lexicality and word frequency effects during prime processing were mainly anticorrelated with delay and target effects, while delay and target effects were positively correlated. This finding indicates qualitative changes from prime to delay (which seems incompatible with spill-over effects) and pre-activation of target-associated representations. Early orthographic effects on word primes were anticorrelated with later prime effects (as well as delay and most of the target representations), suggesting that orthographic representations for words are relatively quickly transformed into preparatory codes (i.e., after lexical access which is achieved before 500 ms post-stimulus onset; e.g., Duñabeitia and Molinaro, 2014). The direct demonstration of orthographic pre-activation constitutes a major theoretical advance of the present study, as

previous evidence for predictive processing at the orthographic level had only been inferred indirectly from predictability effects after word onset (e.g., DeLong et al., 2019).

Source localization revealed correlates of the orthographic effects primarily in frontal and parietal brain regions. We localized lexical-semantic effects to frontal regions (replicating Fiebach et al., 2002; Binder et al., 2003); lexicality effects were furthermore also associated with temporal pole (previously implicated in lexical-semantic facilitation; Lau et al., 2013; 2016) and cerebellum activity, while word frequency effects also involved parietal regions (cf. Desai et al., 2018). The finding that similar brain regions were activated during stimulus processing and pre-activation replicates previous observations (Dikker and Pylkkänen, 2013; Fruchter et al., 2015; Wang et al., 2017).

Implications for Models of Predictive Coding

The activation of linguistic representations before an expected linguistic stimulus is in line with the assumption of active prediction (as specified in predictive coding theory, e.g., Rao and Ballard, 1999). Kuperberg and Jaeger (2016) recently proposed that language comprehension involves predictive processes at multiple levels of the linguistic hierarchy. The present data support this hypothesis for orthographic and lexical-semantic, but not visual and phonological levels of word processing. Neither language-specific (Kuperberg and Jaeger, 2016) nor general theories of predictive coding (Friston, 2005) have discussed restrictions of predictive processing to certain levels of representation to the best of our knowledge. Therefore, this possibility should be explored further in future studies (see Gagl et al., 2020, for a recent example from pre-lexical orthographic processing).

Our data also indicate more temporally extended effects of orthographic pre-activation for words than pseudowords. This finding suggests that the presence of prior stimulus knowledge modulates predictive pre-activation. For example, it may be easier to activate (parts of) pre-existing mental representations in a top-down guided manner, rather than activate a newly assembled set of orthographic features of a previously unknown pseudoword. This notion is also compatible with the absence of sustained orthographic effects for pseudowords in the second half of the delay. In that case, orthographic delay effects for pseudowords could reflect residual activation from processing the prime, but on the other hand, this is not directly compatible with the observed change in the direction of the correlation between prime and delay (as discussed above). In this context, it is also interesting to note that pre-activation effects for lexical-semantic representations in the delay seem to temporally coincide with orthographic pre-activation for words but not pseudowords.

Predictive coding theories assume that internal models can ‘explain away’ the known part of expected sensory inputs and that this part of a stimulus as a consequence does not have to be processed. Accordingly, brain responses elicited during perception should carry more information about the stimulus when perceiving unpredictable in contrast to predictable stimuli (Kok et al., 2012; Blank and Davis, 2016). We found this pattern for OLD20, where more significant effects were observed on the prime than the target. For lexical-semantic metrics, we observed more significant effects on the target than on the prime. This finding may be more compatible with ‘sharpening’, an alternative account of repetition priming effects according to which noisy parts of the signal are suppressed when sensory input is predictable, while the signal’s informative components are sharpened (e.g., Kok et al., 2012). The sharpening theory does not rule out predictive processing per se but represents an alternative mechanism of how predictive processing may be implemented (Kok et al., 2012; Walsh et al., 2020).

Also, the nature of correlations between delay and target seems to pose a challenge for predictive coding models. Predictive coding (Friston, 2005) assumes that the prediction contains relevant information about a stimulus, which is then suppressed when processing the expected stimulus. The strong and sustained positive correlations between delay-period pre-activation and orthographic and lexical-semantic effects during the predictable target's processing indicate that similar representations are active before and during predictable stimulus processing, while representations during unpredictable prime processing were more variable (i.e., of positive and negative effect direction). Again, this pattern suggests a sharpening mechanism that reduces the variability of neuronal responses for predictable compared to unpredictable contexts starting at the delay period.

Potential Limitations

Two methodological considerations may influence the interpretation of our findings: First, the fixed stimulus length (five letters) and resulting restriction of variance in the phonological predictor (number of syllables) may have limited our ability to observe behavioral effects related to visual and phonological processing levels. An unbiased investigation of potential contributions of these processing levels in future work may require more variable stimulus materials. Second, all results of the present study depend on the adequacy of the predictor variables for describing the representations at different processing levels. However, all stimulus measures used here are very well-established in psycholinguistics, and thus appear to be the best possible choices.

Conclusion

We used linear mixed model analyses to show that in behavior and brain activation, context-dependent facilitation of visual word recognition relies most strongly on orthographic and lexical-semantic processes. At these two levels of linguistic representation, information is represented consistently and in a temporally coordinated manner in brain signals before the stimulus is presented. Our findings indicate that neuronal language processing systems actively prepare for the upcoming word and thereby increase word recognition efficiency via a sharpening mechanism.

References

Almeida, D., & Poeppel, D. (2013). Word-specific repetition effects revealed by MEG and the implications for lexical access. *Brain and Language*, *127*. https://doi.org/10.1016/j.bandl.2013.09.013

Álvarez, C.J., Carreiras, M., & Perea, M. (2004). Are syllables phonological units in visual word recognition? *Language and Cognitive Processes*, *19*(3), 427–452. https://doi.org/10.1080/01690960344000242

Assadollahi, R., & Pulvermüller, F. (2003). Early influences of word length and frequency: A group study using MEG. *Neuroreport*, *14*(8), 1183–7. https://doi.org/10.1097/00001756-200306110-00016

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(3), 340–357. https://doi.org/10.1037/0096-1523.10.3.340

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). Parsimonious Mixed Models. *arXiv:1506.04967 [stat]*. http://arxiv.org/abs/1506.04967

Binder, J. R., McKiernan, K. A., Parsons, M. E., Westbury, C. F., Possing, E. T., Kaufman, J. N., & Buchanan, L. (2003). Neural Correlates of Lexical Access during Visual Word Recognition. *Journal of Cognitive Neuroscience*, *15*(3), 372–393. https://doi.org/10.1162/089892903321593108

Blank, H., & Davis, M. H. (2016). Prediction Errors but Not Sharpened Signals Simulate Multivoxel fMRI Patterns during Speech Perception. *PLOS Biology*, *14*(11), e1002577. https://doi.org/10.1371/journal.pbio.1002577

Bonhage, C. E., Mueller, J. L., Friederici, A. D., & Fiebach, C. J. (2015). Combined eye tracking and fMRI reveals neural basis of linguistic predictions during sentence comprehension. *Cortex*, *68*, 33–47. https://doi.org/10.1016/j.cortex.2015.04.011

Brothers, T., Swaab, T. Y., & Traxler, M. J. (2015). Effects of prediction and contextual support on lexical processing: Prediction takes precedence. *Cognition*, *136*, 135–149. https://doi.org/10.1016/j.cognition.2014.10.017

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, *58*(5), 412–424. https://doi.org/10.1027/1618-3169/a000123

Carrasco-Ortiz, H., Midgley, K.J., Grainger, J., & Holcomb, P.J. (2017). Interactions in the neighborhood: Effects of orthographic and phonological neighbors on N400 amplitude. *Journal of Neurolinguistics*, *41*, 1–10. https://doi.org/10.1016/j.jneuroling.2016.06.007

Carreiras, M., Armstrong, B. C., Perea, M., & Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences*, *18*(2), 90–98. https://doi.org/10.1016/j.tics.2013.11.005

Chang, L.-Y., Chen, Y.-C., & Perfetti, C. A. (2018). GraphCom: A multidimensional measure of graphic complexity applied to 131 written languages. *Behavior Research Methods*, *50*(1), 427–449. https://doi.org/10.3758/s13428-017-0881-y

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*(9), 1738–1751. https://doi.org/10.1162/jocn_a_00815

Chetail, F. (2014). Effect of number of syllables in visual word recognition: New insights from the lexical decision task. *Language, Cognition and Neuroscience*, *29*(10), 1249–1256. https://doi.org/10.1080/23273798.2013.876504

Colegate, R. L., & Eriksen, C. W. (1972). Form of redundancy as a determinant of tachistoscopic word recognition. *Perception & Psychophysics*, *12*(6), 477–481. https://doi.org/10.3758/BF03210939

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204

DeLong, K. A., Chan, W., & Kutas, M. (2019). Similar time courses for word form and meaning preactivation during sentence comprehension. *Psychophysiology*, *56*(4), e13312. https://doi.org/10.1111/psyp.13312

Desai, R. H., Choi, W., & Henderson, J. M. (2018). Word frequency effects in naturalistic reading. *Language, Cognition and Neuroscience*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1080/23273798.2018.1527376

Dijkstra, N., Mostert, P., Lange, F. P. de, Bosch, S., & van Gerven, M. A. (2018). Differential temporal dynamics during visual imagery and perception. *eLife*, *7*, e33904. https://doi.org/10.7554/eLife.33904

Dikker, S., & Pylkkanen, L. (2011). Before the N400: Effects of lexical–semantic violations in visual cortex. *Brain and Language*, *118*(1), 23–28. https://doi.org/10.1016/j.bandl.2011.02.006

Duñabeitia, J. A., & Molinaro, N. (2013). The wide-open doors to lexical access. *Frontiers in psychology*, *4*, 471. https://doi.org/10.3389/fpsyg.2013.00471

Eisenhauer, S., Fiebach, C. J., & Gagl, B. (2019). Context-Based Facilitation in Visual Word Recognition: Evidence for Visual and Lexical But Not Pre-Lexical Contributions. *eNeuro*, *6*(2). https://doi.org/10.1523/ENEURO.0321-18.2019

Embick, D., Hackl, M., Schaeffer, J., Kelepir, M., & Marantz, A. (2001). A magnetoencephalographic component whose latency reflects lexical frequency. *Cognitive brain research*, *10*(3), 345–348. https://doi.org/10.1016/s0926-6410(00)00053-7

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, *44*(4), 491–505. https://doi.org/10.1111/j.1469-8986.2007.00531.x

Fiebach, C. J., Friederici, A. D., Müller, K., & von Cramon, D. Y. (2002). FMRI evidence for dual routes to the mental lexicon in visual word recognition. *Journal of Cognitive Neuroscience*, *14*(1), 11–23. https://doi.org/10.1162/089892902317205285

Fiebach, C. J., Gruber, T., & Supp, G. G. (2005). Neuronal Mechanisms of Repetition Priming in Occipitotemporal Cortex: Spatiotemporal Evidence from Functional Magnetic Resonance Imaging and Electroencephalography. *Journal of Neuroscience*, *25*(13), 3414–3422. https://doi.org/10.1523/JNEUROSCI.4107-04.2005

Fiebach, C. J., Ricker, B., Friederici, A. D., & Jacobs, A. M. (2007). Inhibition and facilitation in visual word recognition: Prefrontal contribution to the orthographic neighborhood size effect. *NeuroImage*, *36*(3), 901–911. https://doi.org/10.1016/j.neuroimage.2007.04.004

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning & Verbal Behavior*, *12*(6), 627–635. https://doi.org/10.1016/S0022-5371(73)80042-8

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(4), 680–698. https://doi.org/10.1037/0278-7393.10.4.680

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Fruchter, J., Linzen, T., Westerlund, M., & Marantz, A. (2015). Lexical Preactivation in Basic Linguistic Phrases. *Journal of Cognitive Neuroscience*, *27*(10), 1912–1935. https://doi.org/10.1162/jocn_a_00822

Gagl, B., Richlan, F., Ludersdorfer, P., Sassenhagen, J., & Fiebach, C.J. (2016). The lexical categorization model: A computational model of left ventral occipito-temporal cortex activation in visual word recognition. *BioRxiv*, 085332. https://doi.org/10.1101/085332

Gagl, B., Sassenhagen, J., Haan, S., Gregorova, K., Richlan, F., & Fiebach, C. J. (2020). An orthographic prediction error as the basis for efficient visual word recognition. *NeuroImage*, *214*, 116727. https://doi.org/10.1016/j.neuroimage.2020.116727

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Parkkonen, L., & Hämäläinen, M. S. (2014). MNE software for processing MEG and EEG data. *NeuroImage*, *86*, 446–460. https://doi.org/10.1016/j.neuroimage.2013.10.027

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *6*, e4794. https://doi.org/10.7717/peerj.4794

Heikel, E., Sassenhagen, J., & Fiebach, C. J. (2018). Decoding semantic predictions from EEG prior to word onset. *BioRxiv*, 393066. https://doi.org/10.1101/393066

Hohenstein, S., & Kliegl, R. (2015). remef: Remove Partial Effects. R Package Version 1.0.6.9000. Available at: https://github.com/hohenstein/remef/

King, Jean-Rémi, Gwilliams, L., Holdgraf, C., Sassenhagen, J., Barachant, A., Engemann, D., Larson, E., & Gramfort, A. (2018). *Encoding and Decoding Neuronal Dynamics: Methodological Framework to Uncover the Algorithms of Cognition*. https://hal.archives-ouvertes.fr/hal-01848442

King, J.-R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal generalization method. *Trends in Cognitive Sciences*, *18*(4), 203–210. https://doi.org/10.1016/j.tics.2014.01.002

Kinoshita, S. (1995). The word frequency effect in recognition memory versus repetition priming. *Memory & Cognition*, *23*(5), 569–580. https://doi.org/10.3758/BF03197259

Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron*, *75*(2), 265–270. https://doi.org/10.1016/j.neuron.2012.04.034

Kragel, P. A., Koban, L., Barrett, L. F., & Wager, T. D. (2018). Representation, Pattern Information, and Brain Signatures: From Neurons to Neuroimaging. *Neuron*, *99*(2), 257–273. https://doi.org/10.1016/j.neuron.2018.06.009

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron*, *93*(3), 480–490. https://doi.org/10.1016/j.neuron.2016.12.041

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, *31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161–163. https://doi.org/10.1038/307161a0

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(1), 1–26. https://doi.org/10.18637/jss.v082.i13

Lau, E. F., Gramfort, A., Hämäläinen, M. S., & Kuperberg, G. R. (2013). Automatic Semantic Facilitation in Anterior Temporal Cortex Revealed through Multimodal Neuroimaging. *Journal of Neuroscience*, *33*(43), 17174–17181. https://doi.org/10.1523/JNEUROSCI.1018-13.2013

Lau, E. F., Weber, K., Gramfort, A., Hämäläinen, M. S., & Kuperberg, G. R. (2016). Spatiotemporal Signatures of Lexical–Semantic Prediction. *Cerebral Cortex*, *26*(4), 1377–1387. https://doi.org/10.1093/cercor/bhu219

Lee, C.-Y., Liu, Y.-N., & Tsai, J.-L. (2012). The Time Course of Contextual Effects on Visual Word Recognition. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00285

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, *49*(4), 1494–1502. https://doi.org/10.3758/s13428-016-0809-y

Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. J. (1996). Independent Component Analysis of Electroencephalographic Data. In D. S. Touretzky, M. C. Mozer, & M. E. Hasselmo (Eds.), *Advances in Neural Information Processing Systems 8* (p. 145–151). MIT Press. http://papers.nips.cc/paper/1091-independent-component-analysis-of-electroencephalographic-data.pdf

Manahova, M. E., Mostert, P., Kok, P., Schoffelen, J.-M., & de Lange, F. P. (2018). Stimulus Familiarity and Expectation Jointly Modulate Neural Activity in the Visual Ventral Stream. *Journal of Cognitive Neuroscience*, *30*(9), 1366–1377. https://doi.org/10.1162/jocn_a_01281

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Mayringer, H., & Wimmer, H. (2003). Salzburger Lese-Screening für die Klassenstufen 1-4 (SLS 1-4). Bern: Verlag Hans Huber.

Mostert, P., Albers, A. M., Brinkman, L., Todorova, L., Kok, P., & Lange, F. P. de. (2018). Eye Movement-Related Confounds in Neural Decoding of Visual Working Memory Representations. *eNeuro*, *5*(4). https://doi.org/10.1523/ENEURO.0401-17.2018

Nicenboim, B., Vasishth, S., & Rösler, F. (2020). Are words pre-activated probabilistically during sentence comprehension? Evidence from new data and a Bayesian random-effects meta-analysis using publicly available data. *Neuropsychologia*, *142*, 107427. https://doi.org/10.1016/j.neuropsychologia.2020.107427

Nieuwland, M. S. (2019). Do ‘early’ brain responses reveal word form prediction during language comprehension? A critical review. *Neuroscience & Biobehavioral Reviews*, *96*, 367–400. https://doi.org/10.1016/j.neubiorev.2018.11.019

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Von Grebmer Zu Wolfsthurn, S., Bartolozzi, F., Kogan, V., Ito, A., Mézière, D., Barr, D. J., Rousselet, G. A., Ferguson, H. J., Busch-Moreno, S., Fu, X., Tuomainen, J., Kulakova, E., Husband, E. M., … Huettig, F. (2018). Large-scale replication study

reveals a limit on probabilistic prediction in language comprehension. *eLife*, *7*, e33468. https://doi.org/10.7554/eLife.33468

Notter, M., Gale, D., Herholz, P., Markello, R., Notter-Bielser, M.-L., & Whitaker, K. (2019). AtlasReader: A Python package to generate coordinate tables, region labels, and informative figures from statistical MRI images. *Journal of Open Source Software*, *4*(34), 1257. https://doi.org/10.21105/joss.01257

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 156869. https://doi.org/10.1155/2011/156869

Payne, B. R., Lee, C.-L., & Federmeier, K. D. (2015). Revisiting the incremental effects of context on word processing: Evidence from single-word event-related brain potentials. *Psychophysiology*, *52*(11), 1456–1469. https://doi.org/10.1111/psyp.12515

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, *46*(28), 4646–4674. https://doi.org/10.1016/j.visres.2006.04.023

R Development Core Team. (2008). R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology. Human Perception and Performance*, *37*(2), 514–528. https://doi.org/10.1037/a0020990

Reichel, U. D. (2012). PermA and Balloon: Tools for string alignment and text processing. *Proc. Interspeech*. Interspeech, 13th Annual Conference of the International Speech Communication Association, Portland, USA. https://doi.org/10.5282/ubm/epub.18042

Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioural Neurology*, *12*(4), 191–200. https://doi.org/10.1155/2000/421719

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523–568. https://doi.org/10.1037/0033-295x.96.4.523

Simanova, I., van Gerven, M. A., Oostenveld, R., & Hagoort, P. (2015). Predicting the semantic category of internally generated words from neuromagnetic recordings. *Journal of cognitive neuroscience*, *27*(1), 35–45. https://doi.org/10.1162/jocn_a_00690

Slattery, T. J., & Yates, M. (2018). Word skipping: Effects of word length, predictability, spelling and reading skill: *Quarterly Journal of Experimental Psychology*. https://journals.sagepub.com/doi/10.1080/17470218.2017.1310264

Tange, O. (2011). GNU Parallel: The Command-Line Power Tool. *login: The USENIX Magazine*, *36*(1), 42–47.

Tikhonov, A.N., Arsenin, V.I., & John, F. (1977). Solutions of ill-posed problems, Vol. 14. Washington, DC: Winston.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science*, *247*(4940), 301–306. https://doi.org/10.1126/science.2296719

Van Veen, B. D., van Drongelen, W., Yuchtman, M., & Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Bio-Medical Engineering*, *44*(9), 867–880. https://doi.org/10.1109/10.623056

Vergara-Martínez, M., & Swaab, T. Y. (2012). Orthographic neighborhood effects as a function of word frequency: an event-related potential study. *Psychophysiology*, *49*(9), 1277–1289. https://doi.org/10.1111/j.1469-8986.2012.01410.x

Walsh, K. S., McGovern, D. P., Clark, A., & O'Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Annals of the New York Academy of Sciences*, *1464*(1), 242–268. https://doi.org/10.1111/nyas.14321

Wang, L., Hagoort, P., & Jensen, O. (2017). Language Prediction Is Reflected by Coupling between Frontal Gamma and Posterior Alpha Oscillations. *Journal of Cognitive Neuroscience*, *30*(3), 432–447. https://doi.org/10.1162/jocn_a_01190

Wang, L., Wlotko, E., Alexander, E., Schoot, L., Kim, M., Warnke, L., & Kuperberg, G. R. (2020). Neural Evidence for the Prediction of Animacy Features during Language Comprehension: Evidence from MEG and EEG Representational Similarity Analysis.

Journal of Neuroscience, *40*(16), 3278–3291. https://doi.org/10.1523/JNEUROSCI.1733-19.2020

Wheat, K. L., Cornelissen, P. L., Frost, S. J., & Hansen, P. C. (2010). During Visual Word Recognition, Phonology Is Accessed within 100 ms and May Be Mediated by a Speech Production Code: Evidence from Magnetoencephalography. *Journal of Neuroscience*, *30*(15), 5229–5233. https://doi.org/10.1523/JNEUROSCI.4448-09.2010

Woodhead, Z. V. J., Barnes, G. R., Penny, W., Moran, R., Teki, S., Price, C. J., & Leff, A. P. (2014). Reading front to back: MEG evidence for early feedback effects during word recognition. *Cerebral Cortex*, *24*(3), 817–825. https://doi.org/10.1093/cercor/bhs365

Yarkoni, T. (2009). Big Correlations in Little Studies: Inflated fMRI Correlations Reflect Low Statistical Power-Commentary on Vul et al. (2009). *Perspectives on psychological science*, *4*(3), 294–298. https://doi.org/10.1111/j.1745-6924.2009.01127.x

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*(5), 971–979. https://doi.org/10.3758/PBR.15.5.971

Author Notes

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2013) under grant agreement n° 617891 awarded to C.J.F. and a Marie Curie Fellowship under grant agreement n° 707932 awarded to B.G. We thank Isabelle Ehrlich, Anne Hoffmann, Jan Jürges, Rebekka Tenderra, and Stefanie Wu for their help with data acquisition. The authors declare no competing financial interests.

Author contributions: S.E., B.G., and C.J.F. designed research; S.E. and B.G. performed research; S.E. analyzed data with contributions from B.G.; S.E., B.G., and C.J.F. wrote the paper.

Tables

Table 1. Stimulus parameters of words and pseudowords (PW) in the behavioral and MEG experiments

	Minimum	1st Quartile	Median	3rd Quartile	Maximum	Mean	SE
Behavior: Logarithmic word frequency							
Words	0.000	1.518	1.971	2.189	3.301	1.933	0.095
Behavior: OLD20							
Words	1.000	1.288	1.650	1.750	1.950	1.538	0.038
PW set 1	1.000	1.500	1.650	1.762	2.000	1.605	0.032
PW set 2	1.000	1.288	1.650	1.850	2.100	1.542	0.045
PW set 3	1.000	1.337	1.700	1.850	2.300	1.596	0.044
Behavior: Trigram frequency							
Words	1.806	2.556	2.793	3.162	3.773	2.837	0.056
PW set 1	1.342	2.623	2.856	3.067	3.514	2.826	0.047
PW set 2	1.176	2.567	2.833	3.136	3.136	2.792	0.063
PW set 3	1.437	2.576	2.844	3.237	3.602	2.808	0.061
Behavior: Number of syllables							
Words	1.00	2.00	2.00	2.00	3.00	1.833	0.059
PW set 1	1.00	2.00	2.00	2.00	2.00	1.95	0.028
PW set 2	1.00	2.00	2.00	2.00	2.00	1. 967	0.023
PW set 3	1.00	2.00	2.00	2.00	2.00	1.9	0.039
Behavior: Perimetric complexity							
Words	4.400	5.800	6.400	6.600	8.200	6.354	0.093
PW set 1	5.000	5.800	6.300	6.800	7.800	6.320	0.083
PW set 2	5.200	6.200	6.600	7.000	7.800	6.553	0.083
PW set 3	5.000	5.950	6.600	7.000	7.800	6.457	0.090
Behavior: Number of simple features							
Words	1.400	1.950	2.000	2.200	2.600	2.031	0.031
PW set 1	1.200	1.800	2.000	2.200	2.600	1.950	0.038
PW set 2	1.200	1.800	2.000	2.200	2.600	2.003	0.040
PW set 3	1.400	1.800	2.000	2.200	2.400	1.973	0.040
MEG: Logarithmic word frequency							
Words	0.000	1.512	2.229	2.858	4.032	2.137	0.115
MEG: OLD20							
Words	1.600	1.750	1.850	1.900	2.050	1.825	0.013
PW	1.250	1.637	1.750	1.863	2.300	1.743	0.027
MEG: Trigram frequency							
Words	1.773	2.573	2.785	3.070	3.778	2.804	0.049
PW	1.176	2.402	2.684	2.915	3.597	2.649	0.060
MEG: Number of syllables							

Words	1.00	2.00	2.00	2.00	2.00	1.817	0.050
PW	1.00	2.00	2.00	2.00	2.00	1.933	0.032
MEG: Perimetric complexity							
Words	5.000	5.800	6.400	6.800	7.600	6.260	0.086
PW	5.000	5.800	6.400	6.800	8.400	6.323	0.091
MEG: Number of simple features							
Words	1.400	2.000	2.200	2.400	2.800	2.160	0.038
PW	1.200	1.800	2.000	2.200	2.400	1.983	0.040

See Figure S1 for parameter correlations.

Table 2. Results of the linear mixed model (LMM) investigating context effects (repeated target vs. isolated presentation) on word frequency, OLD20, and lexicality in behavioral response times

	FE	*SE*	*t*	*p*
Context	**-0.278**	**0.013**	**-21.951**	**< 2e-16**
Word frequency	**-0.047**	**0.014**	**-3.312**	**0.001112**
Lexicality	-0.019	0.027	-0.715	0.476
OLD20	**-0.013**	**0.006**	**-2.078**	**0.038**
Trial order	**-0.025**	**0.002**	**-10.03**	**< 2e-16**
Word frequency x Context	**0.039**	**0.012**	**3.316**	**0.000918**
Lexicality x Context	**-0.131**	**0.023**	**-5.601**	**2.19e-08**
OLD20 x Context	**-0.018**	**0.007**	**-2.596**	**0.009450**
OLD20 x Lexicality	0.007	0.011	0.603	0.547129
OLD20 x Context x Lexicality	**0.035**	**0.011**	**3.233**	**0.001230**

Bold numerals indicate significant effects, i.e., $p < .05$. *FE* = fixed effect estimates, *SE* = standard error. Word frequency for pseudowords was set to zero. No significant difference was found in a model comparison of this model with the full model shown in Table S1. Post hoc tests are shown in Table S2 and Table S3.

Table 3. Results of post hoc linear mixed models (LMMs) investigating word frequency and OLD20 effects on behavioral response times separately for repeated and isolated words and pseudowords

	Words							
	Repeated				***Isolated***			
	FE	*SE*	*t*	*p*	*FE*	*SE*	*t*	*p*
Word frequency	0	0.012	-0.041	0.967	**-0.038**	**0.011**	**-3.499**	**0.00099**
OLD20	0.003	0.012	0.264	0.793	-0.005	0.011	-0.473	0.638
Trial order	**-0.017**	**0.006**	**-2.749**	**0.0060**	**-0.023**	**0.004**	**-6.388**	**1.99e-10**
	Pseudowords							
	Repeated				***Isolated***			
	FE	*SE*	*t*	*p*	*FE*	*SE*	*t*	*p*
OLD20	**-0.032**	**0.011**	**-2.877**	**0.0047**	**-0.012**	**0.006**	**-2.095**	**0.038**
Trial order	**-0.027**	**0.007**	**-3.805**	**0.00015**	**-0.027**	**0.004**	**-6.298**	**3.53e-10**

Bold numerals indicate significant effects, i.e., $|t| > 2$. *FE* = fixed effect estimates, *SE* = standard error.

Supporting Information

Supporting Materials. Correlations between stimulus parameters

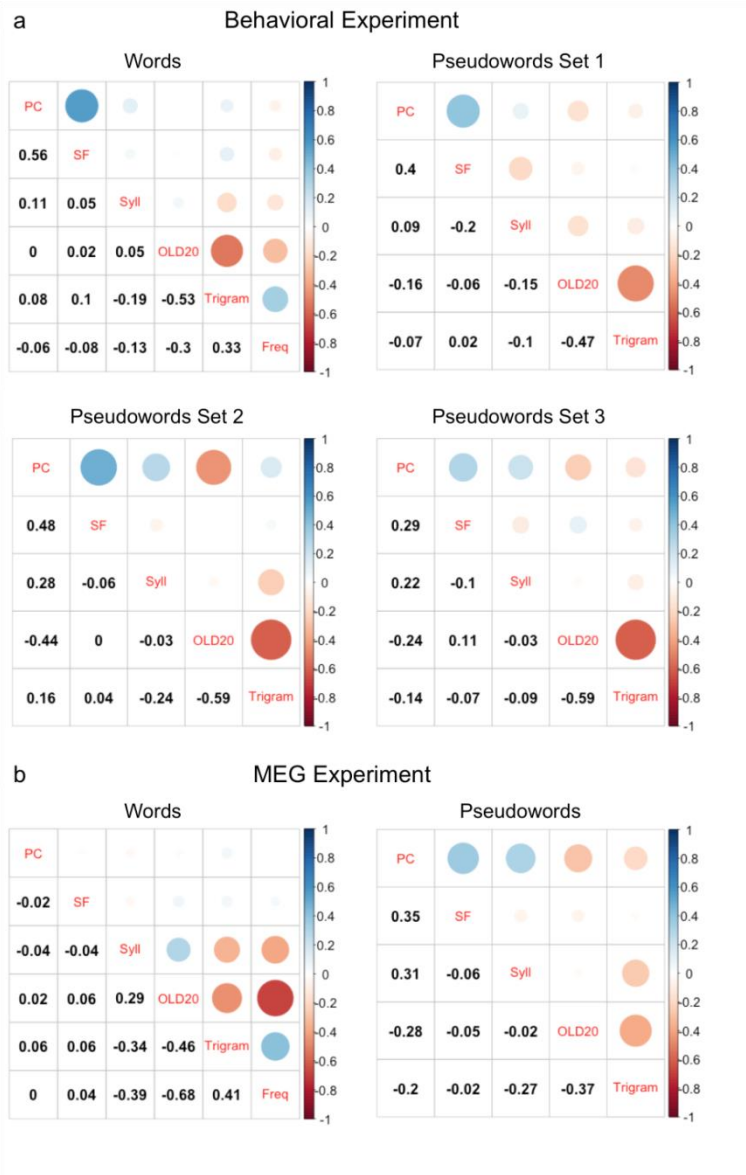

Figure S1. Correlations between stimulus parameters for words and pseudowords of (a) the behavioral and (b) the MEG experiment. PC: Perimetric complexity. SF: Number of simple features. Syll: Number of syllables. OLD20: Orthographic Levenshtein Distance 20. Trigram: Logarithmic trigram frequency. Freq: Logarithmic word frequency.

Supporting Results 1. Additional linear mixed modeling results for behavioral data

Table S1. Results of the full linear mixed model (LMM) of behavioral response times including – besides word frequency and OLD20 – also interactions of context (repeated target vs. isolated presentation) and lexicality with perimetric complexity, the number of simple features, the number of syllables, and trigram frequency

	FE	*SE*	*t*	*p*
Context	**-0.275**	**0.013**	**-20.618**	**< 2e-16**
Word frequency	**-0.049**	**0.015**	**-3.271**	**0.00128**
Lexicality	-0.014	0.028	-0.503	0.616
OLD20	-0.006	0.008	-0.801	0.424
Perimetric complexity	0.007	0.008	0.899	0.369
Trigram frequency	0.008	0.008	0.966	0.335
Simple features	-0.004	0.006	-0.601	0.548
Syllables	-0.004	0.010	-0.429	0.669
Trial order	**-0.025**	**0.002**	**-10.081**	**< 2e-16**
Word frequency x Context	**0.043**	**0.013**	**3.411**	**0.00065**
Lexicality x Context	**-0.139**	**0.025**	**-5.539**	**3.13e-08**
OLD20 x Context	-0.014	0.009	-1.538	0.124
OLD20 x Lexicality	0.002	0.013	0.117	0.907
Perimetric complexity x Context	0.012	0.009	1.345	0.179
Perimetric complexity x Lexicality	0.001	0.013	0.084	0.933
Trigram frequency x Context	0.001	0.009	0.146	0.884
Trigram frequency x lexicality	-0.006	0.014	-0.438	0.662
Simple features x Context	-0.002	0.007	-0.265	0.791
Simple features x Lexicality	-0.007	0.014	-0.508	0.612
Syllables x Context	0.000	0.011	0.042	0.967
Syllables x Lexicality	0.001	0.012	0.121	0.904
OLD20 x Context x Lexicality	**0.036**	**0.013**	**2.760**	**0.00579**
Perimetric complexity x Context x Lexicality	-0.014	0.013	-1.106	0.269
Trigram frequency x Context x Lexicality	0.005	0.013	0.371	0.711
Simple features x Context x Lexicality	0.018	0.013	1.384	0.167
Syllables x Context x Lexicality	0.009	0.013	0.660	0.509

Bold numerals indicate significant effects, i.e., $p < .05$. *FE* = fixed effect estimates, *SE* = standard error. Word frequency for pseudowords was set to zero.

Table S2. Results of post hoc linear mixed models (LMMs) resolving the three-way interaction of context with OLD20 and lexicality, as well as assessing the 'classical' word frequency effect which is typically estimated on words only, by estimating the two-way interactions of context with OLD20 and word frequency separately for words and pseudowords

	Words				***Pseudowords***			
	FE	*SE*	*t*	*p*	*FE*	*SE*	*t*	*p*
Context	**-0.375**	**0.007**	**-56.192**	**< 2e-16**	**-0.314**	**0.008**	**-40.853**	**< 2e-16**
Word frequency	**-0.031**	**0.008**	**-3.715**	**0.00042**	-	-	-	-
OLD20	-0.006	0.008	-0.74	0.462	**-0.013**	**0.006**	**-2.004**	**0.046**
Trial order	**-0.02**	**0.003**	**-6.249**	**4.5e-10**	**-0.03**	**0.004**	**-7.925**	**2.84e-15**
Word frequency x Context	**0.025**	**0.007**	**3.69**	**0.00023**	-	-	-	-
OLD20 x Context	**0.017**	**0.007**	**2.32**	**0.020**	**-0.018**	**0.008**	**-2.378**	**0.017**

Bold numerals indicate significant effects, i.e., |*t*| >2. *FE* = fixed effect estimates, *SE* = standard error.

Table S3. Results of post hoc linear mixed models (LMMs) resolving the interaction of context with word frequency as well as the three-way interaction of context with OLD20 and lexicality in behavioral response times by separately estimating word frequency effects and the OLD20 by lexicality interaction for repeated targets and isolated letter strings

	Repeated target				***Isolated presentation***			
	FE	*SE*	*t*	*p*	*FE*	*SE*	*t*	*p*
Word frequency	-0.001	0.025	-0.043	0.966	**-0.055**	**0.015**	**-3.772**	**0.00023**
Lexicality	**-0.166**	**0.047**	**-3.549**	**0.00057**	-0.001	0.028	-0.031	0.976
OLD20	**-0.03**	**0.01**	**-3.085**	**0.0023**	**-0.012**	**0.006**	**-1.995**	**0.047**
Trial order	**-0.022**	**0.005**	**-4.587**	**4.63e-06**	**-0.026**	**0.003**	**-8.826**	**< 2e-16**
OLD20 x Lexicality	0.033	0.02	1.707	0.090	0.007	0.011	0.588	0.558

Bold numerals indicate significant effects, i.e., |*t*| >2. *FE* = fixed effect estimates, *SE* = standard error. Word frequency for pseudowords was set to zero.

Supporting Results 2. Multivariate pattern decoding analysis and results

We had initially favored multivariate pattern analysis (MVPA) as method of choice for the investigation of pre-activation effects in the MEG data, as previous studies had successfully decoded linguistic information from M/EEG data even in the absence of external stimulation (Simanova et al., 2015; Heikel et al., 2018). However, with this procedure we observed that the statistical power was limited and only resulted in significant findings for the strong and categorical lexicality effect during prime processing (Fig. S2). The absence of decodability of OLD20 and word frequency, i.e., the two investigated parametric variables, from MEG responses was unexpected (i) as both effects were found reliable in our behavioral data, and (ii) given that word frequency and orthographic similarity have in previous work been found to modulate M/EEG responses to words and pseudowords (e.g., Embick et al., 2001; Assadollahi and Pulvermüller, 2003; Vergara-Martínez and Swaab, 2012; Dufau et al., 2015; Gagl et al., 2016; Carrasco-Ortiz et al., 2017). Effects of OLD20 and word frequency should thus at least be found to reliably modulate brain activity during the presentation of the prime. As this was not the case, we reasoned that the MVPA procedure might either lack the statistical power to capture these effects, or that confounding effects (e.g., due to correlated word characteristics) may have obscured the effects of interest. To overcome these potential limitations, we investigated effects of the three stimulus characteristics of interest on MEG-measured brain activation using linear mixed models (LMMs), analogous to the statistical approach used for analyzing the data from the behavioral experiment. To compare the two analysis approaches, a power analysis was calculated (Fig. S3).

MVPA was performed using the SlidingEstimator function of the open-source package MNE-Python (https://martinos.org/mne/stable/index.html; Gramfort et al., 2014). Analyses involving the target time window were conducted exclusively on repetition trials (i.e., trials with identical primes and targets), whereas analyses of baseline, prime, and delay intervals included all available data (i.e., non-repetition and repetition trials). Decoding analyses were performed for each participant on the z-transformed multivariate pattern of neuronal responses measured at each of the 269 MEG sensors, which served as features. We used linear decoders (cf. King et al., 2020) with default parameters (alpha = 1) implemented in the Python-based scikit-learn package (Pedregosa et al., 2011). For the binary lexicality contrast (i.e., words vs. pseudowords), a Logistic Regression decoder was used on data stratified to contain 50% of trials per condition, while continuous parameters were predicted using a Ridge Regression decoder (Tikhonov et al., 1977). In a 5-fold cross-validation

approach, the decoder was trained, in each iteration, on 80 % of the data and then used to predict the remaining 20 % of unseen data, such that each trial was used for testing only once. In detail, trials were randomly partitioned into five folds (with the constraint for the lexicality contrast that training and testing data both contained 50% of trials from each letter string condition), each of which served as testing data once after the decoder had been trained on data from the remaining four folds, thereby allowing us to derive a prediction for all trials in the dataset. To counteract a potential influence of the assignment of trials into folds, this procedure was repeated 100 times with trials randomly assigned to the five partitions of data.

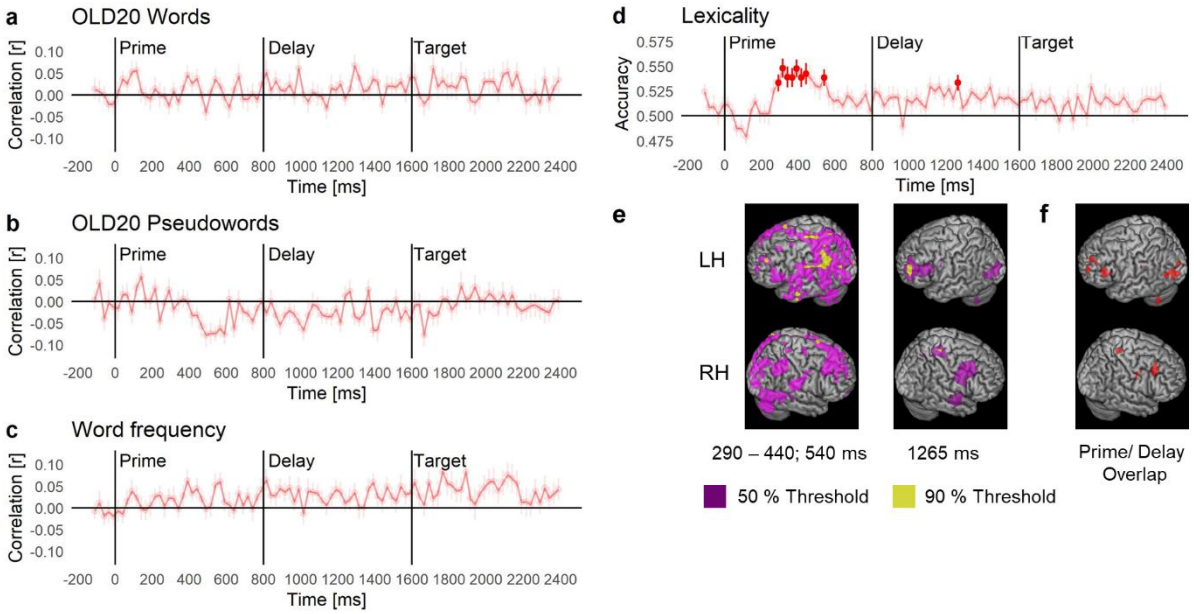

Figure S2. Multivariate pattern analysis (MVPA) decoding performance quantified as the Pearson correlation coefficient between actual and predicted parameter values from ridge regression decoders ± SEM (a-b) or the area under the receiver operating characteristic curve from logistic regression decoders (d) for (a) OLD20 in words, (b) OLD20 in pseudowords, (c) word frequency (in words only), and (d) lexicality. Salient dots represent significant time points (Bonferroni-corrected one-sided Wilcoxon signed-rank test against chance, i.e., a correlation coefficient of zero or an accuracy of .5). The difference between prime and target did not reach significance for any contrast or time point (Bonferroni-corrected). (e) Source topographies display event-related field based sources of significant time points of the lexicality contrast, thresholded at 50 % of the overall maximal activation (violet) within the left (LH) and right (RH) hemisphere. Additionally, source activations thresholded at 90 % of the individual peak per time point are shown (yellow). For the prime contrast, source activations were averaged across significant time points. Note that no control of confounding factors, as for the LMM-based source analysis, was possible. (f) Overlapping sources during prime and delay time windows across all significant time points, thresholded at 50 % of the overall peak. See Table S4 for source coordinates.

Table S4. MNI coordinates and source location of the strongest clusters as well as the overlapping clusters for the lexicality contrast presented in Figure S2

Time [ms]	Cluster	Peak MNI coordinates			Region of cluster
		x	y	z	
					Prime
290 to 440;	1	-53.5	-52.5	38.5	left angular gyrus/ inferior parietal lobule
540	2	-30.5	-32.5	68.5	left postcentral gyrus
	3	2.5	11.5	68.5	right superior frontal gyrus
					Delay
1265	1	-44.5	50.5	19.5	left frontal pole
	2	30.5	-30.5	69.5	right postcentral gyrus
					Prime/Delay Overlap
	1	-39.5	-80.5	4.5	left lateral inferior occipital cortex
	2	44.5	20.5	40.5	right middle frontal gyrus
	3	-51.5	29.5	14.5	left inferior frontal gyrus pars triangularis

Prime and delay clusters are based on the source activation thresholded at 90 % of the individual peak. Overlap estimation included all significant sources of the respective time windows, thresholded at 50 % of the overall peak.

Supporting Results 3. Power comparison between linear mixed modeling and MVPA decoding analysis

To estimate the statistical power for LMM and MVPA analyses, we implemented a Monte-Carlo simulation-based power analysis in R: As a first step, data (behavioral as well as MEG, independent analyses) were randomly sampled from the original datasets (the baseline in case of the MEG data), and split into two halves respecting the random effect structure of the dataset (e.g., the number of participants, and, in case of the LMM-based analysis, sensors). For the LMM-based analysis, we then added to one half of the dataset a specific value reflecting the to-be-tested effect size and conducted statistical analyses as for the original dataset (including Bonferroni correction). For the MVPA analysis, we sampled from the classification accuracies of lexicality, added the critical effect size value, and then tested statistically against .5, i.e., the chance level for this analysis. We re-estimated these analyses 10,000 times for each effect size level (i.e., d values from 0.1 to 1), dataset (behavioral and MEG), and statistical analysis (LMM and MVPA), and recorded the p-values and estimated effect sizes from all tests. The probability of a test being significant at any given effect size represents the power. The mean of all effect size estimates from the significant tests allowed us to estimate the effect size bias, i.e., the likelihood that a significant finding represents an overestimated effect size. Inspection of the power curves (Fig. S3a) reveals that for the LMM analysis of the behavioral data, the present study achieved a power of greater than .9 at small effect sizes of .3. Thus, the likelihood of obtaining a significant result if a small effect exists in the population was greater than 90%. This clearly reflects the benefit of modeling data at the level of the individual trial (as opposed to the 'classical' approach of aggregating trials by subject and condition before calculating statistics; see also Baayen et al., 2008). For the same effect size, the LMM analysis of brain data resulted in a power of around .6 and, alarmingly, for the MVPA analysis, the power of obtaining a statistically significant result was lower than .1. This finding confirms our intuition after the inspection of our brain activation results, i.e., that the MVPA analysis is much less powered. The consequences are severe, because there is a substantial possibility that significant MVPA findings may vastly overestimate the size of the true effect (Fig. S3b). This phenomenon is a frequent but unwanted effect of the correct practice of neglecting non-significant tests, so that significant effects of an underpowered study are often over-estimated in their effect size (e.g., Yarkoni, 2009). LMM analysis is much less susceptible to such overestimations, mainly because it can explicitly estimate crossed random effects (such as subjects and items) by modeling all data at the same time,

while MVPA (and other methods) requires aggregation of data points (for more detailed discussions, see Baayen et al., 2008; Bates et al., 2018). The higher power and reduced bias of the behavioral compared to the MEG analyses further supports our procedure of constraining MEG investigations based on behavioral results; this reduces the possibility of false-positive effects, which is generally high in very high-dimensional datasets.

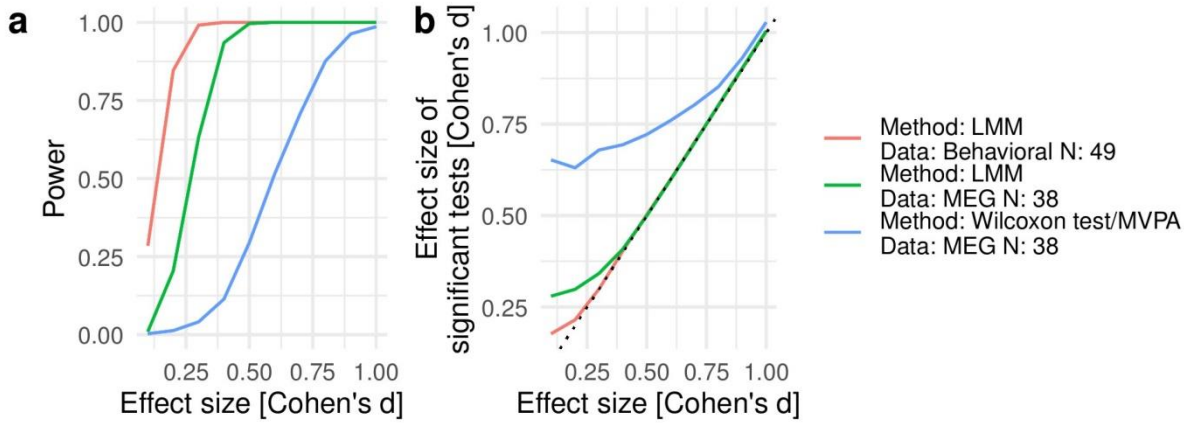

Figure S3. (a) Power curve and (b) effect size inflation for the behavioral analysis and LMM as well as MVPA procedures used to analyze the MEG data.

Supporting Results 4. MNI coordinates and source locations of sources presented in Figure 3

Table S5. MNI coordinates and source locations of the strongest clusters (thresholded at 90 % of the individual peak) presented in Figure 3, as well as of the largest overlapping clusters thresholded at 50 % of the overall peak) across prime, delay and target

Contrast	Time [ms]	Cluster	Peak MNI coordinates			Region of cluster
			x	y	z	
				Prime		
Words vs. PW	40	1	-54.5	0.5	49.5	left precentral gyrus
		2	-34.5	-10.5	69.5	left precentral gyrus
Words vs. PW	240	1	-22.5	52.5	-22.5	left frontal pole
Word frequency	65	1	30.5	-84.5	44.5	right lateral superior occipital cortex
		2	42.5	-72.5	52.5	right lateral superior occipital cortex/ angular gyrus
		3	-42.5	42.5	32.5	left frontal pole
Word frequency	240	1	-39.5	10.5	50.5	left middle frontal gyrus
		2	-30.5	10.5	64.5	left middle frontal gyrus
		3	40.5	9.5	64.5	right middle frontal gyrus
Word OLD20	215	1	-42.5	54.5	20.5	left frontal pole
		2	30.5	-30.5	69.5	right postcentral gyrus
Word OLD20	740	1	-24.5	19.5	60.5	left superior frontal gyrus
		2	-9.5	10.5	64.5	left superior frontal gyrus
PW OLD20	40	1	40.5	-0.5	64.5	right precentral/ middle frontal gyrus
		2	19.5	-82.5	52.5	right lateral superior occipital cortex/ superior parietal lobule
		3	-0.5	0.5	69.5	left juxtapositional lobule/ superior frontal gyrus
PW OLD20	740	1	-22.5	42.5	52.5	left frontal pole
		2	54.5	-69.5	29.5	right lateral superior occipital cortex
		3	-59.5	-9.5	-0.5	left anterior superior temporal gyrus
				Delay		
Words vs. PW	1240	1	-22.5	52.5	-22.5	left frontal pole
		2	20.5	50.5	-24.5	right frontal pole
Words vs. PW	1415	1	-20.5	-19.5	69.5	left precentral gyrus
Word frequency	1240	1	30.5	-94.5	20.5	right occipital pole
		2	24.5	-92.5	29.5	right occipital pole
Word frequency	1515	1	19.5	-82.5	52.5	right lateral superior occipital cortex/ superior parietal lobule

		2	-20.5	-69.5	60.5	left lateral superior occipital cortex/ superior parietal lobule
		3	44.5	20.5	50.5	right middle frontal gyrus
Word OLD20	1115	1	40.5	9.5	64.5	right middle frontal gyrus
Word OLD20	1540	1	-44.5	-21.5	60.5	left postcentral gyrus
		2	-49.5	-29.5	54.5	left postcentral gyrus
PW OLD20	1090	1	40.5	-80.5	0.5	right lateral inferior occipital cortex/ occipital pole
		2	9.5	-82.5	52.5	right lateral superior occipital cortex/ superior parietal lobule
		3	-42.5	-52.5	62.5	left superior parietal lobule
PW OLD20	1515	1	30.5	-10.5	69.5	right precentral gyrus
				Target		
Words vs. PW	1640	1	42.5	-72.5	52.5	right lateral superior occipital cortex/ angular gyrus
Words vs. PW	1815	1	-22.5	52.5	-22.5	left frontal pole
		2	20.5	50.5	-24.5	right frontal pole
Word frequency	1640	1	-20.5	-39.5	69.5	left postcentral gyrus
Word frequency	1790	1	19.5	30.5	64.5	right superior frontal gyrus
Word OLD20	1790	1	19.5	30.5	64.5	right superior frontal gyrus
Word OLD20	2290	1	20.5	39.5	54.5	right frontal pole
		2	-22.5	42.5	52.5	left frontal pole
		2	-0.5	54.5	40.5	right/left superior frontal gyrus
PW OLD20	1790	1	30.5	-50.5	69.5	right superior parietal lobule
		2	-19.5	-64.5	69.5	left lateral superior occipital cortex/ superior parietal lobule
		3	-19.5	-69.5	64.5	left lateral superior occipital cortex/ superior parietal lobule
PW OLD20	2265	1	20.5	-84.5	54.5	right lateral superior occipital cortex/ superior parietal lobule
		2	-29.5	10.5	60.5	left superior frontal gyrus
		3	19.5	59.5	29.5	right frontal pole
				Prime/Delay/Target Overlap		
Words vs. PW		1	-27.5	35.5	-18.5	left frontal pole
		2	20.5	40.5	-19.5	right frontal pole
		3	-38.5	46.5	21.5	left frontal pole
Word frequency		1	-22.5	-79.5	41.5	left lateral superior occipital cortex/ superior parietal lobule
		2	35.5	-69.5	48.5	right lateral superior occipital cortex/ angular gyrus
		3	22.5	17.5	61.5	right superior frontal gyrus

Word OLD20	1	34.5	9.5	57.5	right middle frontal gyrus
	2	-22.5	41.5	51.5	left superior frontal gyrus
	3	12.5	-0.5	68.5	right superior frontal gyrus
PW OLD20	1	36.5	-45.5	62.5	right superior parietal lobule
	2	-28.5	-73.5	50.5	left superior occipital cortex/ superior parietal lobule
	3	-8.5	-60.5	66.5	left precuneus/ superior occipital cortex

PW = pseudowords. Up to three clusters per contrast are shown.