

ORIGINAL ARTICLE

Words in Context: The Effects of Length, Frequency, and Predictability on Brain Responses During Natural Reading

Sarah Schuster^{1,2}, Stefan Hawelka^{1,2}, Florian Hutzler^{1,2},
Martin Kronbichler^{1,2,3} and Fabio Richlan^{1,2}

¹Centre for Cognitive Neuroscience, University of Salzburg, 5020 Salzburg, Austria, ²Department of Psychology, University of Salzburg, 5020 Salzburg, Austria, and ³Neuroscience Institute, Christian-Doppler Klinik, Paracelsus Medical University, 5020 Salzburg, Austria

Address correspondence to F. Hutzler, Centre for Cognitive Neuroscience, University of Salzburg, Hellbrunnerstr. 34, 5020 Salzburg, Austria.
Email: florian.hutzler@sbg.ac.at

Abstract

Word length, frequency, and predictability count among the most influential variables during reading. Their effects are well-documented in eye movement studies, but pertinent evidence from neuroimaging primarily stem from single-word presentations. We investigated the effects of these variables during reading of whole sentences with simultaneous eye-tracking and functional magnetic resonance imaging (fixation-related fMRI). Increasing word length was associated with increasing activation in occipital areas linked to visual analysis. Additionally, length elicited a U-shaped modulation (i.e., least activation for medium-length words) within a brain stem region presumably linked to eye movement control. These effects, however, were diminished when accounting for multiple fixation cases. Increasing frequency was associated with decreasing activation within left inferior frontal, superior parietal, and occipito-temporal regions. The function of the latter region—hosting the putative visual word form area—was originally considered as limited to sublexical processing. An exploratory analysis revealed that increasing predictability was associated with decreasing activation within middle temporal and inferior frontal regions previously implicated in memory access and unification. The findings are discussed with regard to their correspondence with findings from single-word presentations and with regard to neurocognitive models of visual word recognition, semantic processing, and eye movement control during reading.

Key words: eye movement control during reading, functional magnetic resonance imaging, lexical processing, semantic processing, visual word form area (VWFA)

Introduction

Reading is perceived as almost effortless, although it requires a considerable number of cognitive operations that proceed within a fraction of time. Specifically, it requires us to relate a given letter sequence to its respective phonology and semantics

and, crucially, to integrate this information in order to comprehend a continuous text. During natural reading, the time our eyes remain on a given word is substantially influenced by the ease with which it can be processed (Rayner 1998). Word length, frequency, and predictability count among the most influential

visuo-orthographic, lexical, and contextual processing factors during visual word recognition (Rayner 1998, 2009). To illustrate, short and/or frequently encountered words are more often skipped (i.e., are not foveated) and receive shorter fixation durations during reading than long and/or infrequently encountered words (e.g., Rayner and Raney 1996; Kliegl et al. 2004, 2006). Likewise, the predictability of a word—based on the preceding sentence context—facilitates word processing inasmuch as contextually predictable words are more often skipped or are fixated shorter than unpredictable words (Balota et al. 1985; Kliegl et al. 2004, 2006). These well-documented behavioral effects have obtained a benchmark status in reading research and are utilized for evaluating the adequacy of computational models of eye movement control (e.g., E-Z Reader model: Reichle et al. 2003; SWIFT model: Engbert et al. 2005). Evidence from neuroimaging regarding these effects, however, is—as yet—scarce and partly inconsistent.

The effects of word length and frequency on brain responses during reading (and reading-related tasks) are predominantly assessed in the context of single-word studies (i.e., studies presenting unrelated words in a serial one-by-one fashion). Undoubtedly, neuroimaging studies presenting context-free single words contributed tremendously to our understanding of the neural mechanisms during visual word recognition (for reviews, see Vigneau et al. 2006; Binder et al. 2009; Price 2012; Taylor et al. 2013; Martin et al. 2015). To what extent these findings generalize to natural reading, however, is an open issue. Moreover, the effect of predictability evolves with increasing contextual information, making context-free single-word presentations an inadequate method for studying its underlying neurocognitive processes. Comparatively few studies investigated participants' brain responses in relation to words, which are presented in context, that is, within sentences or paragraphs (e.g., Mazoyer et al. 1993; Stowe et al. 1998; Vandenberghe et al. 2002; Xu et al. 2005; Brennan and Pylkkänen 2012). To accomplish such investigations, contemporary neuroimaging studies frequently administered a (rapid) serial visual presentation (RSVP). In this paradigm, sentences (or paragraphs) are broken up into a sequence of single words that are presented in (fast) succession. This method makes it possible to analyze (contextual) effects on the word-level (as opposed to model the neural response over whole paragraphs, sentences, or parts of sentences; e.g., Keller et al. 2001; Ikuta et al. 2006; Jobard et al. 2007; Bahlmann et al. 2011; Pallier et al. 2011; Altmann et al. 2014). Within the RSVP paradigm, however, the presentation of the stimuli is externally controlled (for an in-depth discussion, see Hutzler et al. 2007). During natural reading, by contrast, the attention of the reader is deployed endogenously, which is reflected in the eye movement behavior by frequent word skipplings and refixations. Furthermore, the RSVP prevents (parafoveal) preprocessing of upcoming words (e.g., Hutzler et al. 2013), which constitutes a pivotal factor during natural reading (for a review, see Rayner 2009).

Recent studies addressed the gap between the limited ecological validity of contemporary neuroimaging paradigms and natural viewing behavior (Marsman et al. 2012; Richlan et al. 2014; see also Henderson et al. 2015). Marsman et al. (2012) introduced a methodological advancement that allows investigators to infer participants' brain activation in relation to their current fixation. In technical terms, the fixation-related fMRI approach uses the onset of a first fixation on the stimuli as the marker for modeling the haemodynamic brain response; a technique analogous to the well-established fixation-related brain potentials in the context of electroencephalography (EEG;

e.g., Hutzler et al. 2007; Dimigen et al. 2011). In brief, the fixation-related functional magnetic resonance imaging (fMRI) approach allows researchers to analyze effects on the word-level while presenting whole sentences (or paragraphs).

This study investigated the influence of word length, frequency, and predictability on the neural correlates during natural (i.e., self-paced and silent) reading of sentences. For this purpose, we applied the fixation-related fMRI approach to realize an ecologically valid reading situation while analyzing these linguistic variables at the word level. In the following, we briefly sketch out the existing evidence regarding these effects (1. frequency, 2. predictability, 3. length). As described below, findings regarding these effects—especially the effects of frequency and predictability—revealed considerable discrepancies among existing studies which—most probably—can be attributed to differences in methodological aspects (e.g., task demands and the timing of stimulus presentation; see Schuster et al. 2015).

Frequency

Studies comparing the activation elicited by high-frequent versus low-frequent words most consistently revealed a higher activation within the left inferior frontal gyrus (IFG) for low-frequent compared with high-frequent words (Fiebach et al. 2002; Kronbichler et al. 2004; Carreiras et al. 2006, 2009; Yarkoni et al. 2008b). The IFG is considered to be involved in both phonological and semantic processing, which engage in different locations (see Bokde et al. 2001; Devlin et al. 2003; McDermott et al. 2003; Mechelli et al. 2005). Inconsistent findings, however, are reported with regard to the effect of word frequency on the activation pattern of the visual word form area (VWFA), which is localized in the left occipito-temporal sulcus, lateral to the fusiform gyrus (Cohen et al. 2002; Dehaene et al. 2005; Dehaene and Cohen 2011). In its original conceptualization, the role of the VWFA was attributed to sublexical processing only (Dehaene et al. 2002, 2005; McCandliss et al. 2003; Dehaene and Cohen 2011). This notion implies an insensitivity of the VWFA to word frequency (see Chee et al. 2002, 2003; Dehaene et al. 2002; Vinckier et al. 2007 for supportive evidence). By contrast, several other studies reported lower activation of the VWFA in response to high-frequent compared with low-frequent words (Keller et al. 2001; Kuo et al. 2003; Kronbichler et al. 2004; Yarkoni et al. 2008b). To reconcile these findings with the original notion of the VWFA as an area dedicated to sublexical processing, the proponents of the original conceptualization ascribed the higher activation for low-frequent words to increased task-induced top-down activation and (unnaturally) long presentation time (Dehaene and Cohen 2011). Thus, it will be of interest to see whether the VWFA exhibits a frequency effect during self-paced silent reading (in which the average fixation duration is about 250 ms; Rayner 2009) without any additional task demands beyond reading for comprehension (see Schuster et al. 2015 for an in-depth discussion of this issue).

Predictability

Most evidence concerning contextual processing stems from studies using the semantic violation paradigm. In this paradigm, participants read semantically well-formed and semantically anomalous sentences. Findings gained from this paradigm indicate involvements of the left temporal and inferior frontal cortex in semantic processing (for a review, see Lau et al. 2008). To illustrate, the left IFG exhibited elevated activation in relation to semantic violations (e.g., Newman et al. 2001; Kiehl et al. 2002;

Kuperberg et al. 2003, 2008; Hagoort et al. 2004; Dien et al. 2008), which is considered to reflect the top-down mediated retrieval and/or selection of competing semantic information (Thompson-Schill et al. 1997, 1999), or—more recently—as reflecting the unification process that integrates information to form a coherent sentence interpretation (Hagoort 2005, 2013). By contrast, left temporal regions, which encompass posterior parts of the middle temporal gyrus (MTG), the superior temporal sulcus (STS), and the inferior temporal cortex (IT), have been associated with the storage and the access to lexico-semantic information (e.g., Hickok and Poeppel 2007; Lau et al. 2008). The response of the left temporal cortex to semantic violations, however, is less consistently reported by fMRI studies (but see Kuperberg et al. 2003). Notably, studies manipulating the expectancy (i.e., predictability) of target words (in addition to the congruency of the target words) report higher activation for unexpected compared to expected (predictable) words in temporal regions (Baumgaertner et al. 2002; Dien et al. 2008). This finding indicates that the response of the temporal cortex is modulated more by word predictability than by congruency (see Lau et al. 2008 for an in-depth discussion). With regard to this study, the presented sentences did not contain any semantic violations and thus we were able to address the role of predictability during contextual processing. To date, this aspect of sentence processing has been insufficiently addressed in the fMRI literature, although it plays a pivotal role during natural reading (e.g., Balota et al. 1985; Hawelka et al. 2015).

Length

While the effects of frequency and predictability (referred to as expectancy in the aforementioned studies) do not show a conclusive pattern of results, converging evidence suggest that the (linear) effect of word length is mostly restricted to occipital areas including the posterior fusiform and lingual gyri (e.g., Mechelli et al. 2000; Richlan et al. 2010; Schurz et al. 2010). The linear effect of word length is well-established in the context of behavioral studies which have showed that long compared with short words substantially increase participants' response times (i.e., in naming and lexical decision tasks) and viewing times (Vitu et al. 1990; Balota et al. 2004; Juphard et al. 2004). A re-examination of the word length effect on lexical decision reaction times in a large-scale study based on the English Lexicon Project, however, revealed that word length affects response times in a more curvilinear fashion (New et al. 2006). Specifically, medium-length words (i.e., 5–8 letters) elicited the shortest response times while short (i.e., <5 letters) and long words (i.e., 8–13 letters) elicited comparatively prolonged response times. Substantiating this notion, Yarkoni et al. (2008b) showed that word length influences activation in various brain regions in such a curvilinear fashion. Among these regions, the VWFA exhibited a U-shaped modulation by eliciting the least activation to medium-length words (i.e., 7- to 9-letter words) suggesting that this region preferentially responds to words with an “optimal” length (i.e., they can be more efficiently processed which is then reflected by a reduced activation).

This Study

To sum up, this study investigated the effects of word length, frequency, and predictability on brain responses during natural reading by means of the recently introduced fixation-related fMRI approach. As aforementioned, the effects of word length, frequency, and predictability on fixation durations and fixation

probabilities play prominent roles in reading research and are well studied in the context of eye movement research (Rayner 1998, 2009; Heister et al. 2012). The body of evidence regarding these effects within the neuroimaging literature, however, is largely based upon single-word presentations. Thus, the question to what extent these findings generalize to natural reading an open issue. Moreover, as discussed above, the effects of frequency and predictability yielded divergent findings. This discrepancy could be—at least partly—attributed to differences in presentation durations and tasks that impose demands above and beyond visuo-orthographic processing (Dehaene and Cohen 2011; Schuster et al. 2015). Fixation-related fMRI offers the possibility to study the effects of word length, frequency, and predictability with “natural” presentation durations (i.e., the individual fixation durations of the participants) and without task demands beyond silent reading for comprehension.

Furthermore, relating participants' brain responses to natural eye movement behavior, which is characterized by frequent word skipplings and refixations, would further substantiate the applicability of ecologically valid scanning procedures as presently administered by means of the fixation-related fMRI approach. Specifically, investigating the impact of word skipplings and refixations in relation to word length, frequency, and predictability on brain responses is—as yet—an unrealizable endeavor in the context of single-word presentations. Thus, it will be of interest to observe whether these eye movement parameters modulate the presently investigated effects of interest and would, thus, further contribute to our understanding of the neural mechanisms during reading. More generally, applying fixation-related fMRI allows us to relate our findings to behavioral evidence and to computational models of eye movement control during reading (e.g., E-Z Reader model: Reichle et al. 2003; SWIFT model: Engbert et al. 2005).

Materials and Methods

Participants

A total of 56 (31 male) undergraduate students participated in the study ($M = 25$ years; $SD = 5$ years). All participants reported no history of neurological or psychiatric disorders and had normal or corrected-to-normal vision. First, the reading skill of the participants was assessed by a measure of words per minute (w.p.m.) read during the course of the present experiment. Second, all participants administered a standardized reading speed test that is currently being developed in our laboratory. The reading speed test required judging the meaningfulness of sentences within a time limit of 3 min (i.e., judging the semantical correctness of sentences). Since judging the meaningfulness of the sentences is very easy ($M < 1$ incorrect answers), the number of correctly marked sentences can be considered as a measure of reading speed. The preliminary norms of the test are based on a sample of 309 university students. Participants who exhibited a reading rate of less than 150 w.p.m. or who marked less than 40 sentences in the reading speed test were excluded from the analysis ($n = 9$). The mean of the reading rate of the final sample was 230 w.p.m. ($SD = 50$) and their mean performance on the reading speed test corresponded to the 72th percentile. Before scanning, participants gave their written informed consent. The experiment was conducted in accordance with the Declaration of Helsinki and was approved by the local ethics committee of the University of Salzburg.

Materials

Participants silently read sentences for comprehension. We utilized the Potsdam Sentence Corpus that originally comprises 144 German sentences which are semantically and syntactically legal (Kliegl et al. 2004). From this corpus, we utilized 117 sentences (ca. 80% of the original Potsdam Sentence Corpus), which did not exceed a total character length of 57 to maintain a visual angle of 0.22 of a single letter (viewing distance was ca. 200 cm). Sentence length ranged from 5 to 11 words with a mean of 7.6 ($SD = 1.2$). Sentences were presented in a bold, monospaced font on an MR-compatible LCD screen (NordicNeuroLab) with a resolution of 1024×768 pixel and a refresh rate of 60 Hz.

For the purpose of the present experiment, we considered three word characteristics, that are word frequency, length, and predictability, as our variables of interest. Word frequency is expressed as the log-transformed (base 10) occurrences per million (range: 0.0–4.4; $M = 2.1$; $SD = 1.3$) and was derived from the CELEX database (Baayen et al. 1993). Word length ranged from 2 to 18 letters ($M = 5.4$; $SD = 2.6$). Similar to previous work, we pooled words consisting of 3 or less letters into a single category and words consisting of 12 or more letters into another single category (Kliegl et al. 2004, 2006; Hawelka et al. 2010). Word predictability is defined as the probability of correctly guessing the upcoming word on the basis of the preceding sentence context. The predictability norms for each word were collected in an independent norming study by the Potsdam group and are based on 272 German native speakers (83 complete predictability protocols; Kliegl et al. 2004). These norms range between 0, which denotes completely unpredictable words, and 1, which denotes the most predictable words ($M = 0.21$; $SD = 0.28$). After excluding sentence initials and closed-class words (i.e., determinators, particles, conjunctions, prepositions, and pronouns), a total of 518 words were used for the eye-tracking as well as the fixation-related fMRI analyses (for a similar approach, see Dimigen et al. 2011). Fixations shorter than 80 ms were excluded from the analysis (4.2%). In total, we observed 14 712 fixations that were used for the eye-tracking analysis. Of these, 11 603 were first fixation cases. Note that a well-known problem in reading research dealing with multiple predictors is the high correlation between variables (i.e., multicollinearity). Reassuringly, the correlations among our target words (i.e., words with valid first fixations on nouns, verbs, adjectives, and adverbs) between word frequency and predictability and between predictability and word length were rather low ($r = 0.21$ and -0.13 , respectively). The size of the correlation between frequency and length was moderate ($r = -0.35$). Due to word skipings, the actual size of the correlations varied for the individual participants (e.g., from $r = -0.29$ to -0.41 for the correlation of word frequency with word length).

Procedure

Before a sentence was presented, two vertically aligned fixation bars appeared at the vertical center near the left border of the screen for a pseudo-randomly chosen duration (ranging from 1000 to 3000 ms with increments of 500 ms). These fixation bars were positioned with respect to the optimal landing position on the first word of a sentence (i.e., in the middle or slightly left to the middle of the first word of a sentence; O'Regan and Lévy-Schoen 1987). While the participants fixated between the bars, a drift correction ($n = 19$ participants) or a fixation control ($n = 28$ participants) was administered by the eye-tracking

system. After a successful drift correction or fixation control, a sentence appeared in the horizontal center of the screen which the participants read silently for comprehension. Fixating a cross at the bottom of the right corner of the screen terminated the presentation of the trial. After approximately 10% of the sentences, participants had to answer a simple 2-alternative forced-choice question regarding the content of the preceding sentence via a button press (12 questions in total). The questions and the alternative choices were presented visually. In addition to the experimental trials (i.e., the presentation of the sentences), 24 null events were implemented during which the fixation bars remained on the screen for 2 s.

Data Acquisition and Analysis

Eye-tracking

Eye movements were recorded with an EyeLink CL system in the long-range setup (SR-Research) with a sampling rate of 1 kHz. The eye-tracker was placed at the rear end of the scanner bore at a distance of approximately 90 cm behind the participant and approximately 120 cm in front of the screen. Recording was monocular (from the right eye), and the participant's head was stabilized in the head coil. A horizontal 3-point calibration routine preceded the experiment. Additionally, the experiment was divided into 3 sessions between each of which the eye-tracker was recalibrated. Each trial (i.e., the presentation of a sentence and the null events) was preceded by a drift correction procedure to confirm the aforementioned calibration parameters ($n = 19$ participants) or a fixation control procedure in which a fixation had to be detected by the eye-tracking system around the fixation bars (40×40 pixels; $n = 28$ participants). If the drift correction or the fixation control procedure failed, the system was recalibrated.

Image Acquisition

Functional imaging data were acquired with a Siemens Magnetom Trio 3 Tesla scanner (Siemens AG) equipped with a 12-channel head coil. Functional images sensitive to blood-oxygen-level dependent (BOLD) contrast were acquired with a T_2^* -weighted gradient echo EPI sequence (TR 2000 ms, TE 30 ms, matrix 64×64 mm, FOV 192 mm, flip angle 80°). Thirty-six slices with a slice thickness of 3 mm and a slice gap of 0.3 mm were acquired within the TR. The scan procedure encompassed 3 sessions with a variable number of scans per session. The exact number of scans depended on the participants' reading speed and potential recalibration procedures and ranged from 106 scans to 437 scans ($M = 152$ and $SD = 39$ scans). In addition to the functional images, a gradient echo field map (TR 488 ms, TE 1 = 4.49 ms, TE 2 = 6.95 ms) and a high-resolution ($1 \times 1 \times 1.2$ mm) structural scan with a T_1 -weighted MPRAGE sequence were acquired from each participant.

fMRI Data Analysis

For preprocessing, we used SPM8, whereas for statistical analysis, we used SPM12 software (<http://www.fil.ion.ucl.ac.uk/spm/>) running in a MATLAB 7.14 environment (Mathworks Inc.). Functional images were corrected for geometric distortions by the use of the FieldMap toolbox, realigned and unwarped, and then coregistered to the high-resolution structural image. Note that due to technical issues, the correction for geometric distortions by means of the respective field map was not viable for one participant. The structural image was normalized to the MNI T1 template image, and the resulting parameters were used for normalization of the functional

images, which were resampled to isotropic $3 \times 3 \times 3$ mm voxels and smoothed with a 6-mm FWHM Gaussian kernel. No slice timing correction was applied.

Statistical analysis was performed by means of computing a fixed effects model on the first level (i.e., single subject level) and a random effects model on the second level (i.e., group level). The BOLD response was related to the eye-tracking data in the specifications of the first-level model. In a first step, each onset of a first fixation on a word (irrespective whether the word received more than one fixation, i.e., multiple fixations) was used to model the canonical hemodynamic response function. Furthermore, we investigated the impact of participants' skipping and refixation behavior (i.e., multiple fixations during first-pass reading) on the parametric modulations of word length, frequency, and predictability. To this end, we contrasted the BOLD response in relation to fixation cases that were either preceded or followed by a word skipping with instances in which the previous or next word was fixated. To investigate the influence of refixations on participants' brain responses, we defined multiple fixation cases (ca. 17.5% of the data) as regressors of no interest in an additional model. In all models, the onsets of a fixation on the first word of each sentence, as well as the onsets and durations of the comprehension questions, were not analyzed further, but coded in separate regressors of no interest. These regressors ensured an unbiased word reading versus fixation baseline contrast. Furthermore, 6 head movement parameters, which were derived from the realignment preprocessing step, were modeled as covariates of no interest. The functional data of these first-level models were high-pass filtered with a cutoff of 128 s and corrected for autocorrelation by an AR(1) model (Friston et al. 2002). The parameter estimates of these first-level models, reflecting signal change for word reading versus baseline (comprising the interstimulus intervals, the null events, and the eye-tracker drift correction/recalibration procedures), were calculated in the context of a General Linear Model (Henson 2004). The three effects of interest, that is, the log-transformed word frequency, word length (linear and quadratic), and word predictability, were added as parametric regressors of the word reading contrast (see Fig. S1 within the Supplementary Material for a prototypical design matrix). We centered and orthogonalized (by means of the "spm_orth" function) word length in order to capture the variance attributed to the quadratic length effect, which is not already explained by the linear length effect. The resultant subject-specific contrast images were then used for the second-level random effects analysis. In the second-level analysis, these subject-specific contrasts were submitted to one-sample *t*-tests. For the unmodulated regressors (i.e., our word reading vs. baseline and word skipping vs. fixation contrasts), we

orthogonalized the parametric modulation regressors with respect to the unmodulated regressors in order to provide an appropriate interpretation of these contrasts (see Mumford et al. 2015). Statistically significant effects on the whole-brain level were identified using a voxel-level threshold of $P < 0.001$ (uncorrected) and a cluster-level threshold of $P < 0.05$ (FWE corrected for multiple comparisons).

Results

Behavioral Results

The analysis of the 2-alternatives forced-choice comprehension questions revealed a close-to-ceiling performance with a mean accuracy of 98.94% (minimum = 11 out of 12 correct). In the following analysis of the eye movement data, we report the main effects of word length, frequency, and predictability within repeated-measures multiple regression analysis of participants' skipping probabilities and log-transformed fixation durations (i.e., first fixation and gaze duration). Specifically, we conducted separate regression equations for each participant as described by Lorch and Myers (1990). The resulting regression coefficients for the linear as well as the quadratic effect of word length and the linear effects of word frequency and predictability were then entered into one-sample *t*-tests. The mean regression coefficients (SDs) and the respective *t* values are presented in Table 1.

As evident from the upper panel of Figure 1, word length and frequency elicited effects on participants' skipping probabilities. The multiple regression analysis revealed that word length was the primary determinant for word skipping. Skipping probabilities asymptotically decreased with increasing word length resulting in a significant linear as well as in a significant quadratic effect of word length. The effect of word frequency on participants' skipping probabilities, that is, increasing skipping probability as a function of increasing frequency, was marginally significant. For word predictability, we likewise observed a marginally significant effect on participants' skipping probabilities, that is, a slight increase of skipping probability with increasing word predictability. Table 1 and Figure 1 further show that word length and frequency elicited substantial effects on all fixation duration measures. Specifically, increasing word length resulted in a curvilinear increase in fixation durations (with the shortest fixation durations for 6- and 7-letter words) resulting in significant linear as well as quadratic effects of word length. Increasing word frequency resulted in a significant decrease of first fixation and gaze durations. Word predictability did not result in significant effects on participants' first fixation and gaze durations.

Table 1 Means, SDs, and corresponding *t* values of the effects of word length, frequency, and predictability on participants' skipping probability and log-transformed first fixation and gaze duration

	Skipping Probability			First Fixation Duration			Gaze Duration		
	M	SD	<i>t</i>	M	SD	<i>t</i>	M	SD	<i>t</i>
Intercept	-0.831	0.472	-12.1**	5.47	0.194	193**	5.56	0.201	190**
Length (linear)	-0.311	0.088	-24.3**	0.008	0.011	5.1**	0.035	0.016	14.7**
Length (quadratic)	0.025	0.020	8.5**	0.001	0.004	1.8 ⁺	0.006	0.004	11.8**
Frequency	0.023	0.091	1.7 ⁺	-0.022	0.022	-6.6**	-0.029	0.026	-7.6**
Predictability	0.197	0.701	1.9 ⁺	0.020	0.134	1.0	-0.006	0.154	-0.3

Note. ** $P < 0.01$; ⁺ $P < 0.10$.

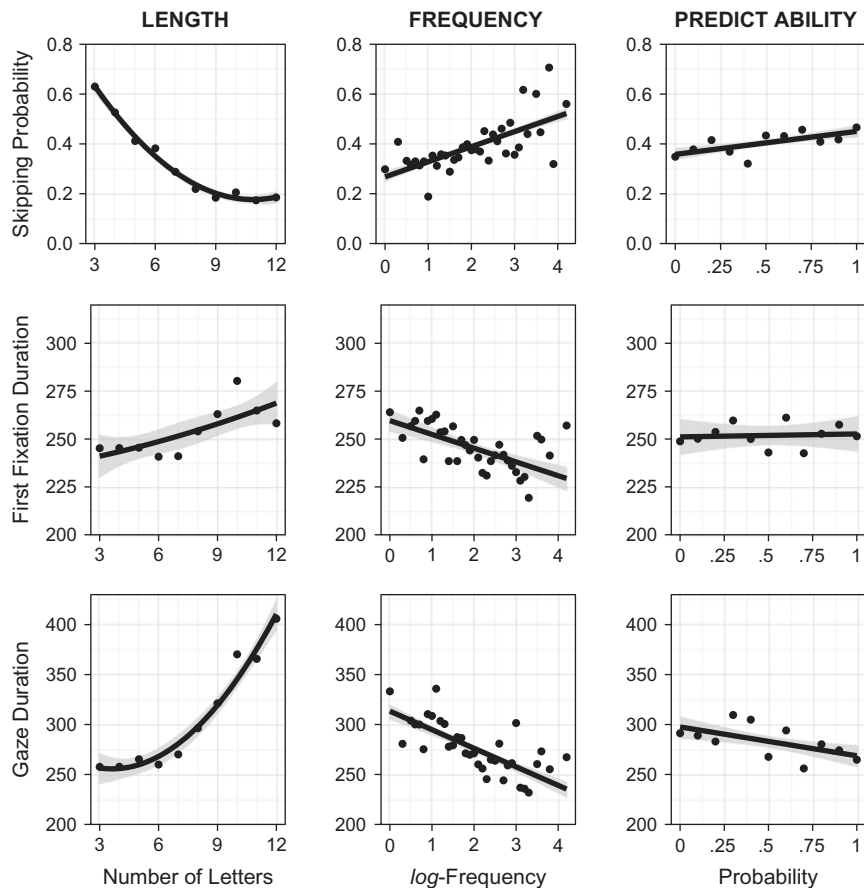


Figure 1. Eye movement parameters in relation to word length, frequency, and predictability. For word frequency and predictability, the lines depict the linear trend of their effects on the eye movement parameters. For word length, the lines show the combined effect of the linear and the quadratic trends. The shaded areas denote 95% pointwise confidence intervals and were derived from the “smooth” function (method = “lm”) of the “ggplot2” package (Wickham 2009) running in the R environment for statistical computing (R Core Team 2015).

fMRI Results

Reading (i.e., the mean activation across fixations) contrasted against baseline elicited activation bilaterally in occipital regions, encompassing the calcarine cortex and the lingual gyri. As illustrated in the upper panel of Figure 2, we observed activation bilaterally in inferior, middle, and superior temporal regions (extending anterior-to-posterior along the superior temporal sulcus), inferior, middle and superior frontal, precentral and posterior parietal regions, cuneus, anterior insula, supplementary motor cortex, and the cerebellum. Furthermore, we observed subcortical activation within the middle cingulate gyrus, hippocampus, thalamus, putamen, caudate, and pallidum. For peak voxels and respective cluster extents, see Table S1 within the Supplementary Material.

The lower panel of Figure 2 illustrates the effects of participants’ skipping behavior on brain activation. As can be seen in Table 2, we observed higher activation within the left posterior middle temporal gyrus when the previous word was skipped as compared with instances in which it received a fixation (see lower left panel of Fig. 2). Furthermore, we observed several regions showing higher activation when the upcoming word was skipped compared with instances in which the upcoming word was fixated. These clusters encompassed bilateral occipital regions including the lingual gyri and the calcarine cortex as well as several left-hemispheric regions including posterior middle temporal gyrus, precentral gyrus, superior parietal

lobule, supplementary motor area, and the cuneus (see lower right panel of Fig. 2). Note that we did not observe any cluster at the whole-brain level for the reverse contrast, that is, no brain region showed higher activation when the previous or upcoming word was fixated as compared with instances in which it was skipped.

Parametric Modulations by Length, Frequency, and Predictability

Before identifying activation clusters that are systematically related to word frequency, length, and predictability, we assessed whether these effects were affected by participants’ skipping behavior. To this end, we compared (by means of paired-sample *t*-tests) the modulation of our effects of interests when the previous or the next word was skipped with instances in which these words were fixated. For the effects of word length and predictability, we observed no significant differences with regard to the issue as to whether the next word received a fixation or not (i.e., no differences for upcoming word skipped > upcoming word fixated nor for upcoming word fixated > upcoming word skipped). For word frequency, we observed higher activation within the left medial frontal cortex when the previous word received a fixation compared with instances in which it was skipped and within bilateral (medial) precentral regions and supplementary motor cortex when the

next word was skipped compared with instances in which it received a fixation. The circumstance as to whether the previous word was fixated or skipped did not affect the effects of word length. For word predictability, we observed higher activation within the right temporo-parietal white matter when the previous word was skipped compared with instances in which it received a fixation. We note that the regions in which the effects of length, frequency and predictability differed with respect to the participants' skipping behavior did not overlap with the regions that we identified by our main models (including fixation cases irrespective whether the previous or the next word were subject to skipping), which we report in the following paragraphs.

The results of our main models, that is, the parametric modulations of word frequency and length, are listed in Table 3. The upper panel of the Table shows the findings with respect to all first fixation cases and the lower panel shows the effects

after controlling for multiple fixation cases. As reasoned in the Introduction section, modulations of word length were tested for both linear and quadratic effects, whereas modulations of word frequency and predictability were tested for linear effects only.

Length

When modeling the hemodynamic response in relation to each first fixation on a word (irrespective of whether it received multiple fixations), we observed an increase in activation with increasing word length in bilateral occipital regions encompassing the lingual gyri and within the right calcarine cortex (see upper panel of Fig. 3). This positive linear modulation, however, ceased to be significant at the whole-brain level when controlling for multiple fixation cases (t values of peak voxels, cluster extents, and cluster-level corrected P values of the left occipital region diminished from t = 4.80, extent = 41, and P = 0.03 to t = 4.31, extent = 11, and P = 0.67, respectively; for the right calcarine cortex, these figures diminished from t = 4.72, extent = 86, and P = 0.001 to t = 3.88, extent = 21, and P = 0.24, respectively). Furthermore, we observed several clusters that exhibited a negative linear relationship, that is, a decrease in activation as a function of increasing word length. These clusters encompassed the right temporo-parietal white matter when modeling brain responses in relation to first fixation cases and, additionally, the left anterior supramarginal gyrus when controlling for multiple fixations. The analysis of the quadratic effect revealed a U-shaped modulation of word length within the brain stem (i.e., in the pons bordering the midbrain; see middle panel of Fig. 3). The specific response of this brain stem structure as a function of word length is illustrated in Figure 4. As can be seen, this structure elicited the least activity in relation to 6- and 7-letter words. We confirmed this apparent pattern by pairwise comparisons of the signal change estimates (for spheres with a radius of 6 mm). Short (<5 letters) as well as long words (>9 letters) elicited a significantly higher activation than words of medium length (6 and 7 letters); t(46) = 2.48, P < 0.05 and t(46) = 2.91, P < 0.01, respectively. Again, this effect was diminished and ceased to be significant at the whole-brain level when controlling for multiple fixations (t values of peak voxels, cluster extents, and cluster-level corrected P values diminished from t = 5.80, extent = 53, and P = 0.01 to t = 4.61, extent = 26, and P = 0.16, respectively).

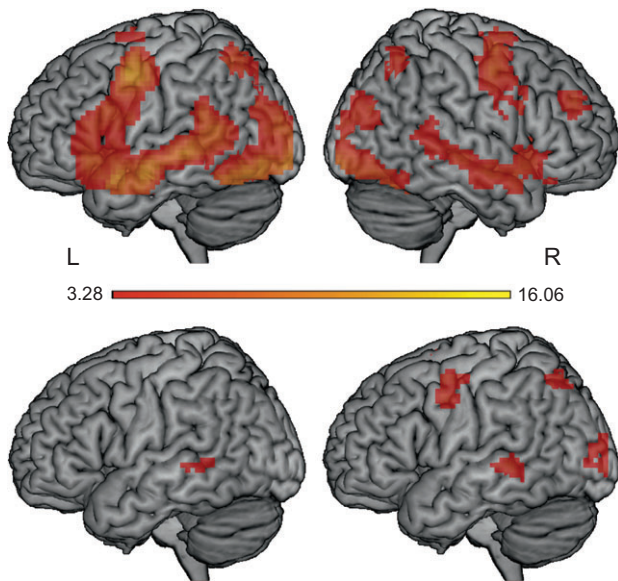


Figure 2. The upper panel depicts regions showing higher activation for word reading versus baseline (presented in left and right view). The lower panel depicts regions showing higher activation when the previous word was skipped as compared with instances in which it received a fixation (left) and when the upcoming word was skipped versus fixated (right).

Table 2 Regions showing higher activation when the previous word was skipped compared with instances in which it was fixated (upper panel) and when the upcoming word was skipped compared with instances in which it was fixated (lower panel)

Region	Voxel extent	MNI coordinates			t
		x	y	z	
Previous word skipped > Previous word fixated					
L posterior middle temporal gyrus	40	-54	-40	1	4.16
Upcoming word skipped > Upcoming word fixated					
L posterior middle temporal gyrus	94	-54	-40	1	5.61
L supplementary motor cortex	80	-3	5	64	5.13
L precentral gyrus	101	-51	-1	40	5.09
R calcarine cortex	337	6	-79	7	4.96
R lingual gyrus		9	-73	-11	4.46
L occipital pole	141	-12	-94	4	4.60
L superior parietal lobule	54	-30	-64	49	4.42
R calcarine cortex	47	18	-64	4	4.36

Note. L, left; R, right.

Table 3 Regions modulated by word length and frequency with respect to first fixation cases (upper panel) and controlled for multiple fixation cases (lower panel) with the corresponding cluster extents and t values

Region	Voxel extent	MNI coordinates			t
		x	y	z	
Model 1: first fixation cases					
Positive linear effect of length					
L occipital pole	41	-12	-97	1	4.80
L lingual gyrus		-9	-85	-11	4.33
R calcarine cortex	86	12	-82	1	4.72
R lingual gyrus		12	-79	-8	3.46
Negative linear effect of length					
R temporo-parietal white matter	41	42	-55	25	4.57
Positive quadratic effect of length					
brain stem	53	6	-34	-20	5.80
Negative linear effect of frequency					
L anterior fusiform gyrus	81	-45	-46	-17	6.27
L middle fusiform gyrus		-48	-58	-20	5.05
L IFG pars triangularis	97	-42	35	13	6.09
L IFG pars opercularis		-42	23	22	3.70
L hippocampus	50	-33	-28	-8	5.40
L superior parietal lobule	99	-27	-64	40	4.47
Model 2: first fixation cases controlled for multiple fixations					
Negative linear effect of length					
R angular gyrus	45	42	-52	28	4.26
L supramarginal gyrus	37	-39	-37	40	4.34
Negative linear effect of frequency					
L anterior fusiform gyrus	79	-45	-46	-17	6.19
L middle fusiform gyrus		-48	-58	-20	5.11
L IFG pars triangularis	89	-45	35	13	5.84
L IFG pars opercularis		-42	23	22	3.54
L anterior cingulate gyrus	42	-9	26	13	5.37
L hippocampus	60	-33	-28	-8	5.19
L superior parietal lobule	107	-27	-64	37	4.49

Note. L, left; R, right.

No cluster exhibited the reverse pattern, that is, an inverted U-shaped modulation as a function of word length.

Frequency

We observed several left-hemispheric clusters in which activation decreased as a function of increasing word frequency (i.e., showing a negative linear effect) in both our models. As can be seen in the lower panel of Figure 3, these clusters encompassed the occipito-temporal cortex, the pars triangularis (extending to the pars opercularis) of the IFG, the superior parietal lobule (including a peak in the intraparietal sulcus), and the left hippocampus. Furthermore, we observed a negative linear relationship within the left anterior cingulate cortex when controlling for multiple fixation cases (see lower panel of Table 3). No cluster showed the opposite effect, that is, an increase in activation in relation to increasing word frequency.

Predictability

For the effect of predictability, we neither observed any positive nor negative modulations at the whole-brain level for both of our models. Note that—contrary to the effect of word length and frequency—word predictability did not elicit a statistically significant effect on participants' fixation durations. Based on previous literature, however, we hypothesized that the effect of predictability would modulate brain activation in a negative linear fashion, that is, brain activation would decrease as a function of increasing predictability. In order to inform future

studies about brain regions activated during contextual processing as indexed by word predictability, we conducted an exploratory analysis using an uncorrected voxel and cluster-level threshold of $P < 0.05$ within our presently obtained contrast testing for negative linear effects. This analysis revealed bilateral modulations within middle and superior temporal regions, inferior frontal (triangular and opercular), and frontal orbital regions, as well as within the insula cortices. Furthermore, we observed a modulation within the left occipito-temporal cortex (see Fig. 5). Note that these modulations were unaffected by the inclusion of multiple fixation cases.

Discussion

This study investigated the effects of word length, frequency, and predictability during natural reading by means of fixation-related fMRI. As yet, evidence regarding these effects is primarily based on serial word-by-word presentations. We assessed to what extent previous findings generalize to natural self-paced silent reading. On the behavioral level, we replicated former findings from eye movement studies by demonstrating that word length and word frequency affect participants' skipping probabilities and fixation durations. Specifically, word length was the primary determinant of word skipping, that is, increasing word length was associated with a (asymptotic) decrease in skipping probability. Furthermore, increasing word length was associated with a curvilinear increase in fixation durations,

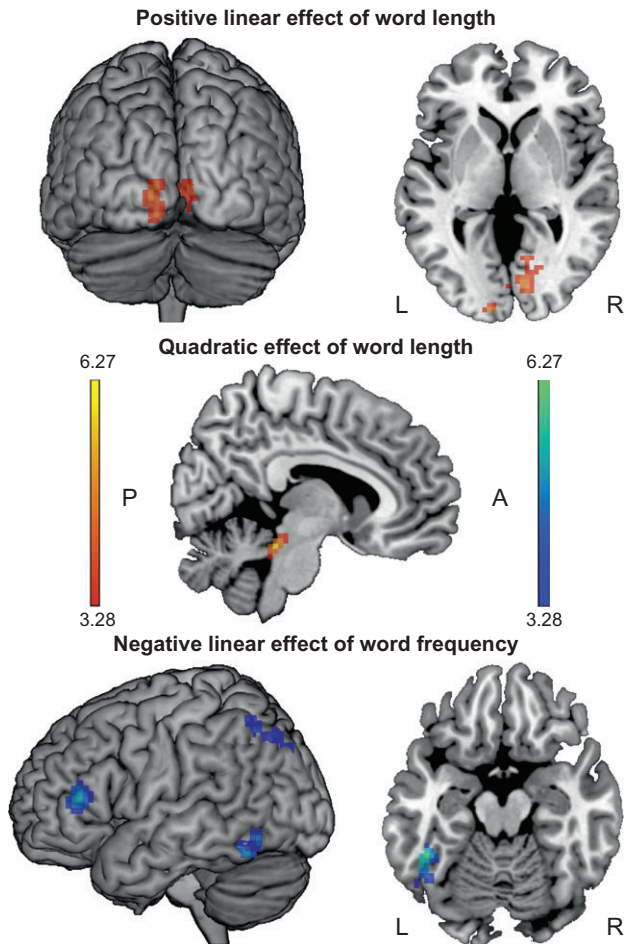


Figure 3. Whole-brain results presented for model 1 (i.e., first fixation cases). The upper and middle panels depict regions that were significantly modulated by word length (presented in posterior, axial at $z = 1$ and sagittal view at $x = 6$). The lower panel depicts regions that were significantly modulated by word frequency (presented in left and axial view at $z = -17$). Orange/yellow activations denote positive effects; blue/turquoise activations denote negative effects.

whereas increasing word frequency was (linearly) associated with shorter fixation durations (in case the word was fixated) (Kliegl et al. 2004, 2006). For word predictability, we observed a tendency towards more frequent skippings as a function of increasing predictability.

On the neuronal level, we observed a positive linear effect of word length, that is, activation increased with increasing word length within bilateral occipital areas linked to visual analysis. Additionally, word length elicited a U-shaped modulation of activation within a brain stem structure located in the pons bordering the midbrain. These effects of word length, however, were diminished when we accounted for multiple fixation cases, suggesting a dependency on participants' fixation behavior (see below). With regard to word frequency, we observed effects within several left-lateralized clusters including occipito-temporal, inferior frontal, superior parietal, and hippocampal regions. These regions exhibited a decrease in activation with increasing frequency. With regard to the effect of predictability, we did not observe any significant modulations at the whole-brain level. Based on previous literature, however, we hypothesized that the effect of predictability would exhibit a negative linear effect on brain activation (i.e., brain activation would decrease as a function of increasing predictability). An

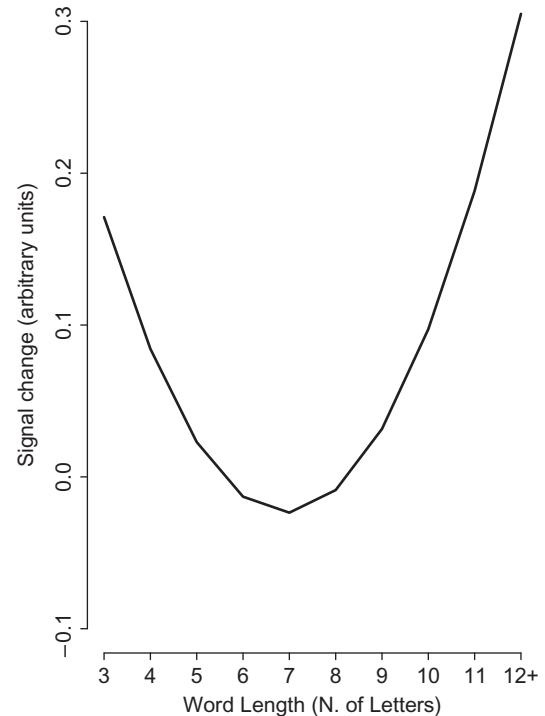


Figure 4. Signal change (in arbitrary units) of the quadratic response of the brain stem region as a function of word length.

exploratory analysis revealed such modulations of word predictability within the left occipito-temporal, bilateral middle and superior temporal, and inferior frontal regions. In the following paragraphs, we discuss the implications of our results with regard to previous findings and to what extent the presently administered fixation-related fMRI approach provides novel insights into the neural underpinnings of visual word recognition. Finally, we discuss how our findings may relate to the proposed neural underpinnings of eye movement control during reading.

Length

Increasing activation in occipital fusiform regions has been linked to the local feature processing demands for longer words. Increasing activation in the lingual gyri has been linked to the global shape processing demands (see Mechelli et al. 2000). Accordingly, our results revealed a positive linear relationship between brain activation and word length within bilateral occipital regions including the lingual gyri and within the right calcarine cortex (Mechelli et al. 2000; Wydell et al. 2003; Valdois et al. 2006; Schurz et al. 2010). In addition, this study revealed that these effects were qualified by the participants' fixation behavior in such a way that the effects of word length were only significant when we modeled brain activation in relation to each first fixation irrespective of whether a word received multiple fixations. When we included multiple fixations as an additional regressor into our model, then the effects were diminished. This indicates that the increase in activation in visual areas when processing long compared with short words reflects the visual processing demands over the whole course of first-pass reading (i.e., the sum of all fixations during the first encounter of a word). That multiple fixations account for the length effect on the neural level is in line

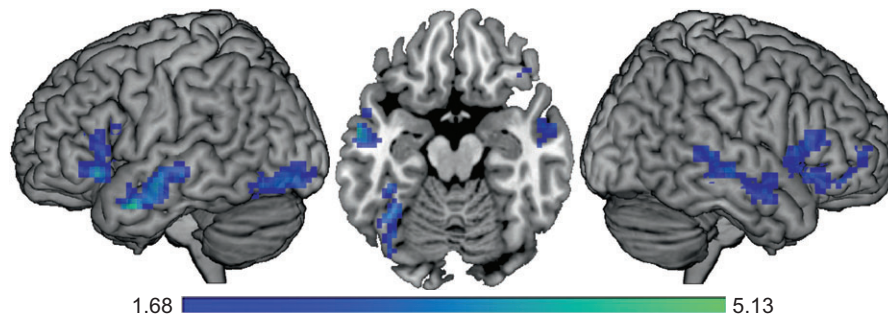


Figure 5. Exploratory analysis of the negative linear effect of word predictability presented in left lateral, ventral (at $z = -17$), and right lateral view. The threshold for both, the voxel level and the cluster level, was $P < 0.05$.

with the observation from eye movement studies that single fixations and first fixations are relatively unaffected by word length (up to 7 letters; see Fig. 1 and Hawelka et al. 2010). In addition to the linear effect of word length, we observed a U-shaped modulation within a brain stem structure located in the pons bordering the midbrain. This area showed the least activation in response to 6- and 7-letter words and higher activation in response to short (<5 letters) and long words (>9 letters). Again, when we accounted for multiple fixations, the effect was diminished (and ceased to be significant). This quadratic effect of word length within the brain stem (and its dependency on fixation behavior) will be discussed below (see “Neural correlates linked to eye movement control during reading”).

Frequency

Increasing word frequency was associated with a decrease in activation within several left-lateralized clusters that have been related to language processing (Jobard et al. 2003; Price 2012; Taylor et al. 2013), visuospatial attention (Corbetta and Shulman 2002; Dosenbach et al. 2007, 2008; Power et al. 2011), and memory (formation and) retrieval (Suzuki et al. 2000, 2014; Davachi 2004; Duff and Brown-Schmidt 2012). As described in the Introduction section, studies investigating the effect of word frequency most consistently revealed higher activation within the left IFG in response to low-frequent compared with high-frequent words (e.g., Fiebach et al. 2002; Kronbichler et al. 2004; Carreiras et al. 2006, 2009; Hauk et al. 2008; Yarkoni et al. 2008b). This IFG activation is considered to reflect both semantic as well as phonological processing that engage in different locations (Bokde et al. 2001; Devlin et al. 2003; McDermott et al. 2003; Mechelli et al. 2005). To illustrate, anterior parts of the left IFG were reported to be more active during semantic than during phonological judgements, whereas posterior parts of the left IFG showed the reverse pattern (see Devlin et al. 2003). This study, which investigated self-paced silent reading with little demands on (explicit) phonological processing, revealed a modulation in response to word frequency primarily within an anterior IFG cluster (i.e., IFG pars triangularis) extending to more posterior sites (i.e., IFG pars opercularis). When controlling for time-on-task effects by means of including the fixation duration as an additional regressor (in our model in which we already controlled for multiple fixation cases), this activation cluster was limited to anterior parts only (see Table S2 of the Supplementary Material), that is, those areas that were previously associated with semantic processing.

Previous studies yielded inconsistent findings regarding the effect of word frequency on the activation pattern of the left OTC, which encompasses the (putative) VWFA (Cohen et al. 2002). The role of the VWFA was originally attributed to sublexical processing—implying an insensitivity to word frequency (see Dehaene et al. 2002; Vinckier et al. 2007). Several studies, however, found lower activation of the VWFA in response to high-frequency compared with low-frequency words (Keller et al. 2001; Kuo et al. 2003; Kronbichler et al. 2004; Yarkoni et al. 2008b)—a finding which we replicated by means of fixation-related fMRI during self-paced silent reading. One interpretation of the sensitivity of the VWFA to word frequency is that it is indicative of the functioning of a mental lexicon comprising representations of frequently encountered words (Kronbichler et al. 2004, 2007, 2009; Glezer et al. 2009; Richlan et al. 2010; Schurz et al. 2010; Wimmer et al. 2010; Ludersdorfer et al. 2013, 2016). According to this notion, the lower activation for high-frequency words can be considered as reflecting “facilitated access” to the representation of frequently encountered words as, for example, envisioned by the dual-route cascaded model of visual word recognition (e.g., Coltheart et al. 2001; but see Protopapas et al. 2016).

Alternatively, the frequency effect in the left OTC could be interpreted within the framework of the Interactive Account of the region’s functioning (Price and Devlin 2011). This explanation—contrary to assuming a selective tuning to orthographic input as purported by the aforementioned VWFA account—suggests that the activation within the left OTC can be attributed to an integrative process, which is characterized by combining the low-level visual information with high-level semantic and phonological information (Price and Devlin 2011). This conceptualization is based on the predictive coding framework postulating reciprocal activations of sensory cortices and higher-order processing regions that convey automatically generated, experience-dependent “predictions” regarding the identity of a stimulus (Rao and Ballard 1999; Friston 2010). Higher activation in response to low-frequent words compared with high-frequent words in the left OTC—as observed in this study—would be interpreted as the result of greater prediction errors for the former than the latter type of words, because previous “experience” with low-frequent words is (per definition) limited. The notion that the higher activation of the VWFA in response to unfamiliar stimuli (e.g., pseudo-words) is caused by automatically generated predictions, and the resultant prediction errors, however, was recently put into perspective by a study from our laboratory (Schuster et al. 2015).

Yet another explanation for the higher activation for unfamiliar (i.e., low-frequency or pseudo) words was recently attributed to complex (artificial) task demands and inadequately long exposure durations (Dehaene and Cohen 2011). Crucially, this study investigated the effect of word frequency during self-paced, silent reading with minimal task demands and natural exposure durations (as realized by the fixation-related fMRI approach). Thus, the effect of word frequency within the VWFA is hardly attributable to such methodological aspects (see also Schuster et al. 2015). Furthermore, when we controlled for time-on-task effects by means of including the fixation duration as an additional regressor (in our model in which we already controlled for multiple fixation cases), the effect within the VWFA remained significant (albeit reduced in extent; see Table S2 within the Supplementary Material). Thus, the presently observed word frequency effect within the VWFA (and other regions) cannot be attributed to differences in processing times for low-frequent versus high-frequent words.

Predictability

For the effect of predictability, we did not observe any effects at the whole-brain level. This was, however, not utterly unexpected, since recent behavioral evidence suggests that a word's predictability has a comparatively small effect in proficient readers (compared with less proficient readers; see Ashby et al. 2005; Hawelka et al. 2015). In order to give insights about brain regions activated during contextual processing as indexed by word predictability, which is—to date—rarely reported in literature (but see Baumgaertner et al. 2002; Dien et al. 2008), we administered an exploratory analysis. A more confirmatory analysis of the effect of predictability (and its relation to reading proficiency) is an issue for future research. Our exploratory analysis revealed that increasing predictability was associated with a decrease in activation within bilateral IFG and MTG, as well as within left OTC regions.

As noted in the Introduction section, most studies on semantic processing of words presented in context administered the semantic violation paradigm. These studies consistently revealed higher left IFG activation in response to semantically incongruent compared with congruent words (Newman et al. 2001; Kiehl et al. 2002; Kuperberg et al. 2003, 2008; Hagoort et al. 2004; Dien et al. 2008; Zhu et al. 2012, 2013). This activation pattern was ascribed to the controlled selection and/or retrieval of semantic information (Thompson-Schill et al. 1997, 1999). More recently, this recruitment of the left IFG was attributed to a unification process, that is, the integration of lexical representations into a coherent multiword representation (Hagoort 2005, 2013). Left MTG activation—which is associated with the storage and retrieval of such lexico-semantic representations (e.g., Hickok and Poeppel 2007)—is less consistently reported by fMRI studies (but see Kuperberg et al. 2003). This is surprising since source-localization by magnetoencephalography consistently pointed towards the left temporal cortex as a major contributor to the well-documented semantic N400 effect (Simos et al. 1997; Helenius et al. 1998; Halgren et al. 2002). The N400 is a negative deflection of the EEG signal in response to semantic violations or unexpected words (Kutas and Hillyard 1980, 1984). Lau et al. (2008) suggested that this discrepancy may be due to differences in the manipulation of the sentential context. To illustrate, most fMRI studies primarily manipulated congruency, whereas most EEG studies also induced strong expectancies as to the identity of the target words. Our findings indicate that differences in predictability

(i.e., expectancy) modulates MTG activation, which supports the notion that the MTG activation does not necessarily reflect “semantic anomaly” signaling, but the extent of pre-activation of lexico-semantic representations within predictive contexts (DeLong et al. 2005; Van Berkum et al. 2005; Lau et al. 2008).

The role of generating predictions during reading has recently gained much interest in light of the Bayesian brain hypothesis (Friston 2003, 2010). In short, it is assumed that our brain aims to minimize surprise by making predictions as to the identity of upcoming sensory events (Bar et al. 2006; Bar 2007). Neuronally, this is achieved by reciprocal activations between lower and higher levels within the cortical hierarchy (Rao and Ballard 1999). The feedforward activation within the visual cortex is considered to convey the “residual errors” between the generated predictions and the actual sensory input. Accordingly, we observed a reduction in activation with increasing word predictability in visual areas extending to anterior OTC regions—indicating reduced prediction errors for predictable words (Hofmann et al. 2014; Willems et al. 2015). As opposed to the conceptualization of predictive coding in the context of word frequency (Price and Devlin 2011), one may argue that the role of predictions might be situated on the sentence level (e.g., Altmann and Kamide 1999; Bonhage et al. 2015) rather than on the level of a single word's familiarity (i.e., frequency).

To conclude, the majority of these findings are in line with evidence from previous studies investigating the neural underpinnings of visual word recognition by means of single-word presentations (e.g., Jobard et al. 2003; Vigneau et al. 2006; Price 2012; Taylor et al. 2013; Martin et al. 2015). This general accordance shows that findings from “traditional” experimental designs do generalize to natural reading. Specifically, we replicated previous findings indicating involvements of the left IFG during lexical processing as indexed by the effect of word frequency (Fiebach et al. 2002; Kronbichler et al. 2004; Carreiras et al. 2006, 2009; Hauk et al. 2008; Yarkoni et al. 2008b) and semantic processing as indexed by the effect of word predictability (albeit within an exploratory analysis; Hagoort 2005, 2013; Lau et al. 2008). Furthermore, the presently observed effects of word length, frequency, and predictability were unaffected by participants' skipping behavior. Put differently, word skipping did not affect those regions that we observed when modeling brain activation in relation to each first fixation on the words. Additionally, the effects of word frequency and predictability were likewise relatively unaffected by rfixations (i.e., multiple fixations on a word) that further substantiate the notion that findings from traditional experimental designs (e.g., RSVP) seem to generalize to natural reading.

Administering an ecologically valid scanning procedure by means of fixation-related fMRI, however, also shed new light on the neural correlates of the effect of word length. For the linear as well as the quadratic effect of word length (discussed below), we observed a dependency on participants' fixation behavior in such a way that the increase in activation associated with processing long words within bilateral occipital regions and the brain stem was diminished when accounting for multiple fixations. A further merit of an ecologically valid scanning procedure, as realized in this study, is the possibility of investigating these word-level effects with minimal task demands. Furthermore, the procedure ensures a natural “exposure duration” (if one equates fixation times with presentation duration). As previously suggested, such methodological aspects might elevate task-related top-down processing which could, for example, modulate the response of the VWFA above

and beyond the intrinsic requirements of reading (Dehaene and Cohen 2011; Schuster et al. 2015). Moreover, this procedure allows us to relate our findings to the potential neural underpinnings of eye movement control during reading.

Neural Correlates Linked to Eye Movement Control During Reading

Since some of our observed effects can be attributed to the unconfined natural eye movement behavior of the participants, we further discuss (some of) our results with regard to pertinent models of eye movement control during reading. Models of eye movement control ascribe a pivotal role to the allocation of visual attention during natural reading, although having different notions about the exact nature of its influence. That is, either by means of “sequential attention shifts” (e.g., E-Z reader model: Reichle et al. 2003) or “guidance by attentional gradients” (e.g., SWIFT: Engbert et al. 2005). On the neural level, such visuospatial attentional processes are presumably guided by the dorsal attention network (Corbetta and Shulman 2002). More recently, the fronto-parietal control network, which includes parts of the dorsal attention network as proposed by Corbetta and Shulman (2002), has been associated with the initiation and adjustment of top-down control (Dosenbach et al. 2007, 2008). With respect to these findings, we observed activation within the superior parietal lobule extending into the intraparietal sulcus—a region that is considered to constitute a part of this network. Specifically, this region exhibited a decrease in activation as a function of increasing word frequency and showed higher activation when the upcoming word is about to be skipped. Findings from eye movement studies on reading suggest that the disengagement of visuospatial attention from the currently fixated word occurs earlier for words that are easier to process such as high-frequency words (e.g., Reichle et al. 2003). One could hypothesize that 1) the earlier disengagement for high-frequency words is reflected by lower activation and that 2) the higher affordances of attentional shifting for word skipplings is reflected by higher activation of this part of the attention network.

Neural circuits within the brain stem are supposedly linked to the final stage of eye movement execution (Reichle et al. 2003; Engbert et al. 2005). This study revealed that word length modulates the activation within the brain stem in a U-shaped fashion, that is, we observed the lowest activation in response to medium-length words. This observed brain stem cluster presumably encompasses the paramedian pontine reticular formation (PPRF). It has been suggested that neurons within the PPRF give rise to the final command of executing voluntary horizontal saccades (Leigh and Zee 1999; Sparks 2002). Eye movement studies on reading consistently revealed that the average saccade length during reading is 7–9 characters (including inter-word spaces; see, e.g., Rayner 2009). Thus, one could speculate that the reduced activation of the brain stem in response to medium-length words reflects this preferred saccadic length during reading. Furthermore, as described in the Introduction section, a large-scale behavioral study revealed that participants exhibit the shortest response latencies for medium-length words (New et al. 2006; see also Ferrand et al. 2010). In accordance with our interpretation of the presently observed brain stem activation in terms of saccadic preference, New et al. (2006) argued (as one of several conceivable causes) that the shortest response latencies for medium-length words reflects their high probability of being the target of a single fixation during natural reading (whereas short words are

frequently skipped and long words are typically processed with multiple fixations; e.g., Kliegl et al. 2004). Critically, this quadratic effect of word length was diminished when accounting for such multiple fixations indicating that the increase in activation in response to long words is—at least partly—driven by such multiple fixation cases.

Furthermore, applying fixation-related fMRI allowed us to investigate participants’ brain responses in relation to their skipping behavior, which constitutes an important factor during natural reading. To illustrate, adult proficient readers skip approximately one-third of the words while reading for comprehension (see Rayner 1998, 2009). With regard to these findings, contrasting the activation in response to skipped against fixated words revealed higher activation within the left MTG when the previous word was skipped compared with instances in which the previous word was fixated. This finding might suggest that previously skipped words are further processed at the lexico-semantic level (e.g., Hickok and Poeppel 2007; Lau et al. 2008). In the same vein, we observed higher activation within the left MTG when the upcoming word is about to be skipped, suggesting that lexico-semantic information is likewise accessed for non-foveated words. Furthermore, bilateral occipital areas showed enhanced activation in instances before a word was skipped as compared with instances before a word was fixated, which might indicate that about to be skipped words are subject to a more pronounced visual analysis (during parafoveal preprocessing) than words that will receive a fixation.

Limitations

This study investigated the neurocognitive correlates of natural reading while participants silently read unconnected sentences for comprehension. Our work was primarily inspired by contemporary models of eye movement control during reading and hence focused on word-level effects. In research on eye movement control during reading, the presentation of unconnected sentences is common practice. One may argue, however, that presenting connected text passages would mimic our daily reading experience even more closely and would have allowed us to investigate more complex linguistic processes such as (long-distance) reference resolution (e.g., McMillan et al. 2012) or elaborative inferences (e.g., Kuperberg et al. 2006). Furthermore, single sentences probably do not induce such strong contextual effects (i.e., expectancies about upcoming words) as whole paragraphs would do. Thus, this study probably did not capture the full potential effect that predictability may exert on brain activation. A further limitation of this study is that it focused exclusively on the effects of word length, frequency, and predictability, because these effects have a benchmark status in eye movement research on reading (for reviews, see Rayner 1998, 2009). From a wealth of studies we know, however, that a great variety of lexical as well as sublexical word characteristics potentially influence visual word recognition (e.g., Graf et al. 2005). For example, a re-analysis of a large-scale study on behavioral data (i.e., lexical decision and naming latencies) revealed that orthographic similarity (indexed by Levenshtein distance) account for a substantial part of both the linear and the quadratic effect of word length (Yarkoni et al. 2008a). This indicates that the effect of word length is a composite of serial processing (of long words) and competition among orthographically similar representations (which is more pronounced for short words; Whitney 2008; Ferrand et al. 2010). Thus, future fixation-related fMRI studies on reading may

consider additional variables (e.g., orthographic similarity, frequency of local combinations such as bigrams, concreteness, imaginability, or emotional valence—to name but a few). From this perspective, this study should be considered as a first step in bridging the gap between behavioral, eye movement, and neuroimaging research on reading.

Conclusion

This study administered fixation-related fMRI in order to investigate the effects of word length, frequency, and predictability during natural reading of whole sentences. The majority of these findings are in line with evidence from previous studies investigating the neural underpinnings of visual word recognition by means of single-word presentations, suggesting that findings gained from commonly used paradigms (e.g., RSVP paradigm) generalize to natural reading. The effects of word length, however, showed a dependency on participants' refixation behavior that might indicate that increasing activation associated with processing long words reflects the visual processing demands over the whole course of first-pass reading. Furthermore, we observed a frequency effect in the VWFA whose role was originally attributed to sublexical processing (Dehaene et al. 2005). We tentatively interpreted this finding as indicative that the area acts as storage for abstract, frequency-sensitive whole-word recognition units (e.g., Wimmer et al. 2010). An exploratory analysis of the effect of word predictability indicated effects in middle temporal regions—implicated in memory retrieval and pre-activation of memory-based lexico-semantic representations (e.g., Lau et al. 2008)—and inferior frontal regions—implicated in the unification of lexical representations into a coherent multiword representation (e.g., Hagoort 2005, 2013).

Supplementary Material

Supplementary material can be found at: <http://www.cercor.oxfordjournals.org/>.

Funding

Austrian Science Fund (FWF P 25799-B23); Austrian Agency for International Cooperation in Education and Research (OeAD PL 11/2015).

Notes

We thank Julian Wenzel, Lisa Wiesner, and Nicola Jacobi for their help with data acquisition, Mario Braun and Nicole A. Himmelstoss for helpful discussions, and Franziska A. Fowles for proof-reading. We are grateful to Reinhold Kliegl for providing us the Potsdam Sentence Corpus. *Conflict of Interest*: None declared.

References

Altmann U, Bohrn IC, Lubrich O, Menninghaus W, Jacobs AM. 2014. Fact vs fiction—how paratextual information shapes our reading processes. *Soc Cogn Affect Neurosci*. 9:22–29.

Altmann GT, Kamide Y. 1999. Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*. 73:247–264.

Ashby J, Rayner K, Clifton C. 2005. Eye movements of highly skilled and average readers: differential effects of frequency and predictability. *Q J Exp Psychol A*. 58:1065–1086.

Baayen RH, Piepenbrock R, van Rijn H. 1993. The CELEX lexical database (CD-ROM). Philadelphia (PA): Linguistic Data Consortium, University of Pennsylvania.

Bahlmann J, Mueller JL, Makuuchi M, Friederici AD. 2011. Perisylvian functional connectivity during processing of sentential negation. *Front Psychol*. 2:104.

Balota DA, Cortese MJ, Sergent-Marshall SD, Spieler DH, Yap M. 2004. Visual word recognition of single-syllable words. *J Exp Psychol Gen*. 133:283–316.

Balota DA, Pollatsek A, Rayner K. 1985. The interaction of contextual constraints and parafoveal visual information in reading. *Cogn Psychol*. 17:364–390.

Bar M. 2007. The proactive brain: using analogies and associations to generate predictions. *Trends Cogn Sci*. 11: 280–289.

Bar M, Kassam KS, Ghuman AS, Boshyan J, Schmid AM, Dale AM, Hämäläinen MS, Marinkovic K, Schacter DL, Rosen BR, et al. 2006. Top-down facilitation of visual recognition. *Proc Natl Acad Sci USA*. 103:449–454.

Baumgaertner A, Weiller C, Büchel C. 2002. Event-related fMRI reveals cortical sites involved in contextual sentence integration. *Neuroimage*. 16:736–745.

Binder JR, Desai HR, Graves WW, Conant LL. 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex*. 19:2767–2796.

Bokde AL, Tagamets MA, Friedman RB, Horwitz B. 2001. Functional interactions of the inferior frontal cortex during the processing of words and word-like stimuli. *Neuron*. 30: 609–617.

Bonhage CE, Mueller JL, Friederici AD, Fiebach CJ. 2015. Combined eye tracking and fMRI reveals neural basis of linguistic predictions during sentence comprehension. *Cortex*. 68:33–47.

Brennan J, Pykkänen L. 2012. The time-course and spatial distribution of brain activity associated with sentence processing. *Neuroimage*. 60:1139–1148.

Carreiras M, Mechelli A, Price CJ. 2006. Effect of word and syllable frequency on activation during lexical decision and reading aloud. *Hum Brain Mapp*. 27:963–972.

Carreiras M, Riba J, Vergara M, Heldmann M, Münte TF. 2009. Syllable congruency and word frequency effects on brain activation. *Hum Brain Mapp*. 30:3079–3088.

Chee MW, Hon NH, Caplan D, Lee HL, Goh J. 2002. Frequency of concrete words modulates prefrontal activation during semantic judgments. *Neuroimage*. 16:259–268.

Chee MW, Lee HL, Soon CS, Westphal C, Venkatraman V. 2003. Reproducibility of the word frequency effect: comparison of signal change and voxel counting. *Neuroimage*. 18:468–482.

Cohen L, Lehericy S, Chochon F, Lemer C, Rivaud S, Dehaene S. 2002. Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*. 125:1054–1069.

Coltheart M, Rastle K, Perry C, Langdon R, Ziegler J. 2001. DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol Rev*. 108:204–256.

Corbetta M, Shulman GL. 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci*. 3: 201–215.

Davachi L. 2004. The ensemble that plays together, stays together. *Hippocampus*. 14:1–3.

Dehaene S, Cohen L. 2011. The unique role of the visual word form area in reading. *Trends Cogn Sci*. 15:254–262.

Dehaene S, Cohen L, Sigman M, Vinckier F. 2005. The neural code for written words: a proposal. *Trends Cogn Sci*. 9: 335–341.

- Dehaene S, Le Clec'h G, Poline JB, Le Bihan D, Cohen L. 2002. The visual word form area: a prelexical representation of visual words in the fusiform gyrus. *Neuroreport*. 13:321–325.
- DeLong KA, Urbach TP, Kutas M. 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nat Neurosci*. 8:1117–1121.
- Devlin JT, Matthews PM, Rushworth MF. 2003. Semantic processing in the left inferior prefrontal cortex: a combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *J Cogn Neurosci*. 15:71–84.
- Dosenbach NU, Fair DA, Cohen AL, Schlaggar BL, Petersen SE. 2008. A dual-networks architecture of top-down control. *Trends Cogn Sci*. 12:99–105.
- Dosenbach NU, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RA, Fox MD, Snyder AZ, Vincent JL, Raichle ME, et al. 2007. Distinct brain networks for adaptive and stable task control in humans. *Proc Natl Acad Sci USA*. 104:11073–11078.
- Dien J, Franklin MS, Michelson CA, Lemen LC. 2008. fMRI characterization of the language formulation area. *Brain Res*. 1229:179–192.
- Dimigen O, Sommer W, Hohlfeld A, Jacobs AM, Kliegl R. 2011. Coregistration of eye movements and EEG in natural reading: analyses and review. *J Exp Psychol Gen*. 140:552–572.
- Duff MC, Brown-Schmidt S. 2012. The hippocampus and the flexible use and processing of language. *Front Hum Neurosci*. 6:69.
- Engbert R, Nuthmann A, Richter EM, Kliegl R. 2005. SWIFT: a dynamical model of saccade generation during reading. *Psychol Rev*. 112:777–813.
- Ferrand L, New B, Brysbaert M, Keuleers E, Bonin P, Méot A, Augustinova M, Pallier C. 2010. The French Lexicon Project: lexical decision data for 38 840 French words and 38 840 pseudowords. *Behav Res Methods*. 42:488–496.
- Fiebach CJ, Friederici AD, Müller K, von Cramon DY. 2002. fMRI evidence for dual routes to the mental lexicon in visual word recognition. *J Cogn Neurosci*. 14:11–23.
- Friston K. 2003. Learning and inference in the brain. *Neural Netw*. 16:1325–1352.
- Friston K. 2010. The free-energy principle: a unified brain theory? *Nat Rev Neurosci*. 11:127–138.
- Friston KJ, Glaser DE, Henson RN, Kiebel S, Phillips C, Ashburner J. 2002. Classical and Bayesian inference in neuroimaging: applications. *Neuroimage*. 16:484–512.
- Glezer LS, Jiang X, Riesenhuber M. 2009. Evidence for highly selective neuronal tuning to whole words in the “visual word form area”. *Neuron*. 62:199–204.
- Graf R, Nagler M, Jacobs AM. 2005. Faktorenanalyse von 57 Variablen der visuellen Worterkennung. *Zeitschrift für Psychologie* 213:205–218.
- Hagoort P. 2005. On Broca, brain, and binding: a new framework. *Trends Cogn Sci*. 9:416–423.
- Hagoort P. 2013. MUC (Memory, Unification, Control) and beyond. *Front Psychol*. 4:416.
- Hagoort P, Hald L, Bastiaansen M, Petersson KM. 2004. Integration of word meaning and world knowledge in language comprehension. *Science*. 304:438–441.
- Halgren E, Dhond RP, Christensen N, Van Petten C, Marinkovic K, Lewine JD, Dale AM. 2002. N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *Neuroimage*. 17:1101–1116.
- Hauk O, Davis MH, Pulvermüller F. 2008. Modulation of brain activity by multiple lexical and word form variables in visual word recognition: a parametric fMRI study. *Neuroimage*. 42:1185–1195.
- Hawelka S, Gagl B, Wimmer H. 2010. A dual-route perspective on eye movements of dyslexic readers. *Cognition*. 115:367–379.
- Hawelka S, Schuster S, Gagl B, Hutzler F. 2015. On forward inferences of fast and slow readers. An eye movement study. *Sci Rep*. 5:8432.
- Heister J, Würzner K, Kliegl R. 2012. Analysing large datasets of eye movements during reading. In: Adelman J.S., editor. *Visual word recognition, volume 2: Meaning and context, individuals and development*. Hove (UK), England: Psychology Press. p. 102–131.
- Helenius P, Salmelin R, Service E, Connolly JF. 1998. Distinct time courses of word and context comprehension in the left temporal cortex. *Brain*. 121:1133–1142.
- Henderson JM, Choi W, Luke SG, Desai RH. 2015. Neural correlates of fixation duration in natural reading: evidence from fixation-related fMRI. *Neuroimage*. 119:390–397.
- Henson RNA. 2004. Analysis of fMRI time series: Linear time-invariant models, event-related fMRI and optimal experimental design. In: Frackowiak RSJ, Friston KJ, Frith C, Dolan R, Price CJ, Zeki S, Ashburner JT, Penny WD, editors. *Human brain function*. 2nd ed. London (UK): Academic Press. p. 793–822.
- Hickok G, Poeppel D. 2007. The cortical organization of speech processing. *Nat Rev Neurosci*. 8:393–402.
- Hofmann MJ, Dambacher M, Jacobs AM, Kliegl R, Radach R, Kuchinke L, Plichta MM, Fallgatter AJ, Herrmann MJ. 2014. Occipital and orbitofrontal hemodynamics during naturally paced reading: an fNIRS study. *Neuroimage*. 94:193–202.
- Hutzler F, Braun M, Vö ML, Engl V, Hofmann M, Dambacher M, Leder H, Jacobs AM. 2007. Welcome to the real world: validating fixation-related brain potentials for ecologically valid settings. *Brain Res*. 1172:124–129.
- Hutzler F, Fuchs I, Gagl B, Schuster S, Richlan F, Braun M, Hawelka S. 2013. Parafoveal X-masks interfere with foveal word recognition: evidence from fixation-related brain potentials. *Front Syst Neurosci*. 7:33.
- Ikuta N, Sugiura M, Sassa Y, Watanabe J, Akitsuki Y, Iwata K, Miura N, Okamoto H, Watanabe Y, Sato S, et al. 2006. Brain activation during the course of sentence comprehension. *Brain Lang*. 97:154–161.
- Jobard G, Crivello F, Tzourio-Mazoyer N. 2003. Evaluation of the dual route theory of reading: a meta-analysis of 35 neuroimaging studies. *Neuroimage*. 20:693–712.
- Jobard G, Vigneau M, Mazoyer B, Tzourio-Mazoyer N. 2007. Impact of modality and linguistic complexity during reading and listening tasks. *Neuroimage*. 34:784–800.
- Juphard A, Carbonnel S, Valdois S. 2004. Length effect in reading and lexical decision: evidence from skilled readers and a developmental dyslexic participant. *Brain Cogn*. 55:332–340.
- Keller TA, Carpenter PA, Just MA. 2001. The neural bases of sentence comprehension: a fMRI examination of syntactic and lexical processing. *Cereb Cortex*. 11:223–237.
- Kiehl KA, Laurens KR, Liddle PF. 2002. Reading anomalous sentences: an event-related fMRI study of semantic processing. *Neuroimage*. 17:842–850.
- Kliegl R, Grabner E, Rolfs E, Engbert R. 2004. Length, frequency, and predictability effects of words on eye movements in reading. *Eur J Cogn Psychol*. 16:262–284.
- Kliegl R, Nuthmann A, Engbert R. 2006. Tracking the mind during reading: the influence of past, present, and future words on fixation durations. *J Exp Psychol Gen*. 135:12–35.

- Kronbichler M, Bergmann J, Hutzler F, Staffen W, Mair A, Ladurner G, Wimmer H. 2007. Taxi vs. Taksi: on orthographic word recognition in the left ventral occipitotemporal cortex. *J Cogn Neurosci*. 19:1584–1594.
- Kronbichler M, Hutzler F, Wimmer H, Mair A, Staffen W, Ladurner G. 2004. The visual word form area and the frequency with which words are encountered: evidence from a parametric fMRI study. *Neuroimage*. 21:946–953.
- Kronbichler M, Klackl J, Richlan F, Schurz M, Staffen W, Ladurner G, Wimmer H. 2009. On the functional neuroanatomy of visual word processing: effects of case and letter deviance. *J Cogn Neurosci*. 21:222–229.
- Kutas M, Hillyard SA. 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*. 207:203–205.
- Kutas M, Hillyard SA. 1984. Brain potentials during reading reflect word expectancy and semantic association. *Nature*. 307:161–163.
- Kuo WJ, Yeh TC, Lee CY, Yu Wu, Chou CC, Ho LT, Hung DL, Tzeng OJ, Hsieh JC. 2003. Frequency effects of Chinese character processing in the brain: an event-related fMRI study. *Neuroimage*. 18:720–730.
- Kuperberg GR, Holcomb PJ, Sitnikova T, Greve D. 2003. Distinct patterns of neural modulation during the processing of conceptual and syntactic anomalies. *J Cogn Neurosci*. 15:272–293.
- Kuperberg GR, Lakshmanan BM, Caplan DN, Holcomb PJ. 2006. Making sense of discourse: an fMRI study of causal inferencing across sentences. *Neuroimage*. 33:343–361.
- Kuperberg GR, Sitnikova T, Lakshmanan BM. 2008. Neuroanatomical distinctions within the semantic system during sentence comprehension: evidence from functional magnetic resonance imaging. *Neuroimage*. 40:367–388.
- Lau EF, Phillips C, Poeppel D. 2008. A cortical network for semantics: (de)constructing the N400. *Nat Rev Neurosci*. 9:920–933.
- Leigh RJ, Zee DS. 1999. *The neurology of eye movements*. New York (NY): Oxford University Press.
- Lorch RF Jr, Myers JL. 1990. Regression analyses of repeated measures data in cognitive research. *J Exp Psychol Learn Mem Cogn*. 16:149–157.
- Ludersdorfer P, Schurz M, Richlan F, Kronbichler M, Wimmer H. 2013. Opposite effects of visual and auditory word-likeness on activity in the visual word form area. *Front Hum Neurosci*. 7:491.
- Ludersdorfer P, Wimmer H, Richlan F, Schurz M, Hutzler F, Kronbichler M. 2016. Left ventral occipitotemporal activation during orthographic and semantic processing of auditory words. *Neuroimage*. 124:834–842.
- Marsman JB, Renken R, Velichkovsky BM, Hooymans JM, Cornelissen FW. 2012. Fixation based event-related fMRI analysis: using eye fixations as events in functional magnetic resonance imaging to reveal cortical processing during the free exploration of visual images. *Hum Brain Mapp*. 33:307–318.
- Martin A, Schurz M, Kronbichler M, Richlan F. 2015. Reading in the brain of children and adults: a meta-analysis of 40 functional magnetic resonance imaging studies. *Hum Brain Mapp*. 36:1963–1981.
- Mazoyer BM, Tzourio N, Frak V, Syrota A, Murayama N, Levrier O, Salamon G, Dehaene S, Cohen L, Mehler J. 1993. The cortical representation of speech. *J Cogn Neurosci*. 5:467–479.
- McCandliss BD, Cohen L, Dehaene S. 2003. The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn Sci*. 7:293–299.
- McDermott KB, Petersen SE, Watson JM, Ojemann JG. 2003. A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. *Neuropsychologia*. 41:293–303.
- McMillan CT, Clark R, Gunawardena D, Ryant N, Grossman M. 2012. fMRI evidence for strategic decision-making during resolution of pronoun reference. *Neuropsychologia*. 50:674–687.
- Mechelli A, Crinion J, Long S, Friston KJ, Lambon-Ralph MA, Patterson K, McClelland JL, Price CJ. 2005. Dissociating reading processes on the basis of neuronal interactions. *J Cogn Neurosci*. 17:1753–1765.
- Mechelli A, Humphreys GW, Mayall K, Olson A, Price CJ. 2000. Differential effects of word length and visual contrast in the fusiform and lingual gyri during reading. *Proc Biol Sci*. 267:1909–1913.
- Mumford JA, Poline JB, Poldrack RA. 2015. Orthogonalization of regressors in fMRI models. *PLoS One*. 10:e0126255.
- New B, Ferrand L, Pallier C, Brysbaert M. 2006. Reexamining the word length effect in visual word recognition: new evidence from the English Lexicon Project. *Psychon Bull Rev*. 13:45–52.
- Newman AJ, Pancheva R, Ozawa K, Neville HJ, Ullman MT. 2001. An event-related fMRI study of syntactic and semantic violations. *J Psycholinguist Res*. 30:339–364.
- O'Regan JK, Lévy-Shoen A. 1987. Eye-movement strategy and tactics in word recognition and reading. *Attention Perform*. 12:363–383.
- Pallier C, Devauchelle AD, Dehaene S. 2011. Cortical representation of the constituent structure of sentences. *Proc Natl Acad Sci USA*. 108:2522–2527.
- Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, et al. 2011. Functional network organization of the human brain. *Neuron*. 72:665–678.
- Price CJ. 2012. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*. 62:816–847.
- Price CJ, Devlin JT. 2011. The interactive account of ventral occipitotemporal contributions to reading. *Trends Cogn Sci*. 15:246–253.
- Protopapas A, Orfanidou E, Taylor JS, Karavasilis E, Kapnoula EC, Panagiotaropoulou G, Velonakis G, Poulou LS, Smyrnis N, Kelekis D. 2016. Evaluating cognitive models of visual word recognition using fMRI: effects of lexical and sublexical variables. *Neuroimage*. 128:328–341.
- R Core Team. 2015. *R: A language and environment for statistical computing*. Austria, Vienna: R Foundation for Statistical Computing. URL: <https://www.R-project.org/>.
- Rao RP, Ballard DH. 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci*. 2:79–87.
- Rayner K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychol Bull*. 124:372–422.
- Rayner K. 2009. Eye movements and attention in reading, scene perception, and visual search. *Q J Exp Psychol (Hove)*. 62:1457–1506.
- Rayner K, Raney GE. 1996. Eye movement control in reading and visual search: effects of word frequency. *Psychon Bull Rev*. 3:245–248.

- Reichle ED, Rayner K, Pollatsek A. 2003. The E-Z Reader model of eye-movement control in reading: comparisons to other models. *Behav Brain Sci.* 4:445–476.
- Richlan F, Gagl B, Hawelka S, Braun M, Schurz M, Kronbichler M, Hutzler F. 2014. Fixation-related fMRI analysis in the domain of reading research: using self-paced eye movements as markers for hemodynamic brain responses during visual letter string processing. *Cereb Cortex.* 24:2647–2656.
- Richlan F, Sturm D, Schurz M, Kronbichler M, Ladurner G, Wimmer H. 2010. A common left occipito-temporal dysfunction in developmental dyslexia and acquired letter-by-letter reading? *PLoS One.* 5:e12073.
- Schurz M, Sturm D, Richlan F, Kronbichler M, Ladurner G, Wimmer H. 2010. A dual-route perspective on brain activation in response to visual words: evidence for a length by lexicality interaction in the visual word form area (VWFA). *Neuroimage.* 49:2649–2661.
- Schuster S, Hawelka S, Richlan F, Ludersdorfer P, Hutzler F. 2015. Eyes on words: a fixation-related fMRI study of the left occipito-temporal cortex during self-paced silent reading of words and pseudowords. *Sci Rep.* 5:12686.
- Simos PG, Basile LF, Papanicolaou AC. 1997. Source localization of the N400 response in a sentence-reading paradigm using evoked magnetic fields and magnetic resonance imaging. *Brain Res.* 762:29–39.
- Sparks DL. 2002. The brainstem control of saccadic eye movements. *Nat Rev Neurosci.* 3:952–964.
- Stowe LA, Broere CA, Paans AM, Wijers AA, Mulder G, Vaalburg W, Zwarts F. 1998. Localizing components of a complex task: sentence processing and working memory. *Neuroreport.* 9:2995–2999.
- Suzuki WA, Eichenbaum H. 2000. The neurophysiology of memory. *Ann N Y Acad Sci.* 911:175–191.
- Suzuki WA, Naya Y. 2014. The perirhinal cortex. *Annu Rev Neurosci.* 37:39–53.
- Taylor JS, Rastle K, Davis MH. 2013. Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychol Bull.* 139:766–791.
- Thompson-Schill SL, D’Esposito M, Aguirre GK, Farah MJ. 1997. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc Natl Acad Sci USA.* 94:14792–14797.
- Thompson-Schill SL, D’Esposito M, Kan IP. 1999. Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron.* 23:513–522.
- Valdois S, Carbonnel S, Juphard A, Baciou M, Ans B, Peyrin C, Segebarth C. 2006. Polysyllabic pseudo-word processing in reading and lexical decision: converging evidence from behavioral data, connectionist simulations and functional MRI. *Brain Res.* 1085:149–162.
- Van Berkum JJ, Brown CM, Zwitserlood P, Kooijman V, Hagoort P. 2005. Anticipating upcoming words in discourse: evidence from ERPs and reading times. *J Exp Psychol Learn Mem Cogn.* 31:443–4467.
- Vandenberghe R, Nobre AC, Price CJ. 2002. The response of left temporal cortex to sentences. *J Cogn Neurosci.* 14:550–560.
- Vigneau M, Beaucousin V, Hervé PY, Duffau H, Crivello F, Houdé O, Mazoyer B, Tzourio-Mazoyer N. 2006. Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage.* 30:1414–1432.
- Vinckier F, Dehaene S, Jobert A, Dubus JP, Sigman M, Cohen L. 2007. Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron.* 55:143–156.
- Vitu F, O’Regan JK, Mittau M. 1990. Optimal landing position in reading isolated words and continuous text. *Percept Psychophys.* 47:583–600.
- Whitney C. 2008. Supporting the serial in the SERIOL model. *Lang Cognit Process.* 23:824–865.
- Wickham H. 2009. *ggplot2: elegant graphics for data analysis.* New York (NY): Springer-Verlag.
- Willems RM, Frank SL, Nijhof AD, Hagoort P, van den Bosch A. 2015. Prediction during natural language comprehension. *Cereb Cortex.* 26:2506–2516.
- Wimmer H, Schurz M, Sturm D, Richlan F, Klackl J, Kronbichler M, Ladurner G. 2010. A dual-route perspective on poor reading in a regular orthography: an fMRI study. *Cortex.* 46:1284–1298.
- Wydell TN, Vuorinen T, Helenius P, Salmelin R. 2003. Neural correlates of letter-string length and lexicality during reading in a regular orthography. *J Cogn Neurosci.* 15:1052–1062.
- Xu J, Kemeny S, Park G, Frattali C, Braun A. 2005. Language in context: emergent features of word, sentence, and narrative comprehension. *Neuroimage.* 25:1002–1015.
- Yarkoni T, Balota D, Yap M. 2008a. Moving beyond Coltheart’s N: a new measure of orthographic similarity. *Psychon Bull Rev.* 15:971–979.
- Yarkoni T, Speer NK, Balota DA, McAvoy MP, Zacks JM. 2008b. Pictures of a thousand words: investigating the neural mechanisms of reading with extremely rapid event-related fMRI. *Neuroimage.* 42:973–987.
- Zhu Z, Feng G, Zhang JX, Li G, Li H, Wang S. 2013. The role of the left prefrontal cortex in sentence-level semantic integration. *Neuroimage.* 76:325–331.
- Zhu Z, Hagoort P, Zhang JX, Feng G, Chen HC, Bastiaansen M, Wang S. 2012. The anterior left inferior frontal gyrus contributes to semantic unification. *Neuroimage.* 60:2230–2237.