

Evaluating cognitive models of visual word recognition using fMRI:**Effects of lexical and sublexical variables**

Athanassios Protopapas	<u>Eleni Orfanidou</u>
University of Athens	University of Crete
J. S. H. Taylor	<u>Efstratios Karavasilis</u>
Royal Holloway University of London	University of Athens
<u>Efthymia C. Kapnoula</u>	Georgia Panagiotaropoulou
University of Iowa	National Technical University of Athens
Georgios Velonakis	Loukia S. Poulou
University of Athens	University of Athens
<u>Nikolaos Smyrnis</u>	Dimitrios Kelekis
University of Athens	University of Athens

Final draft of 27 November 2015; *NeuroImage*, in press

Author Note

Athanassios Protopapas, Department of Philosophy & History of Science, University of Athens, Greece; Eleni Orfanidou, Department of Psychology, University of Crete, Greece; J. S. H. Taylor, Department of Psychology, Royal Holloway University of London, UK; Efstratios Karavasilis, Georgios Velonakis, Loukia S. Poulou, & Dimitrios Kelekis, Radiology and Medical Imaging Research Unit, University of Athens, Greece; Efthymia C. Kapnoula, Department of Psychology, University of Iowa, USA; Georgia Panagiotaropoulou, School of Electrical and Computer Engineering, National Technical University of Athens, Greece; Nikolaos Smyrnis, Department of Psychiatry, Eginition Hospital, School of Medicine, University of Athens.

Correspondence concerning this article should be addressed to Athanassios Protopapas, MITHE, Ano Ilissia University Campus, GR-157 71 Zografos, Greece. Phone: +30 210 727 5540; fax: +30 210 727 5530; e-mail: aprotopapas@phs.uoa.gr

Abstract

In this study predictions of the dual-route cascaded (DRC) model of word reading were tested using fMRI. Specifically, patterns of co-localization were investigated: (a) between pseudoword length effects and a pseudowords vs. fixation contrast, to reveal the sublexical grapho-phonemic conversion (GPC) system; and (b) between word frequency effects and a words vs. pseudowords contrast, to reveal the orthographic and phonological lexicon. Forty four native speakers of Greek were scanned at 3T in an event-related lexical decision task with three event types: (a) 150 words in which frequency, length, bigram and syllable frequency, neighborhood, and orthographic consistency were decorrelated; (b) 150 matched pseudowords; and (c) fixation. Whole-brain analysis failed to reveal the predicted co-localizations. Further analysis with participant-specific regions of interest defined within masks from the group contrasts revealed length effects in left inferior parietal cortex and frequency effects in the left middle temporal gyrus. These findings could be interpreted as partially consistent with the existence of the GPC system and phonological lexicon of the model, respectively. However, there was no evidence in support of an orthographic lexicon, weakening overall support for the model. The results are discussed with respect to the prospect of using neuroimaging in cognitive model evaluation.

Keywords: reading; visual word recognition; cognitive models; fMRI; lexical decision; Greek

Evaluating cognitive models of visual word recognition using fMRI:**Effects of lexical and sublexical variables**

Computational models of visual word recognition originated in neuropsychology to account for the consequences of damage to the brain. Thus, one might expect to be able to map components of the models more or less transparently to brain regions. A number of neuroimaging studies have sought to localize functions and elements of the connectionist triangle (Harm and Seidenberg, 1999, 2004; Plaut, McClelland, Seidenberg, & Patterson, 1996) and the dual-route cascaded (DRC; Coltheart, Curtis, Atkins, & Haller, 1993; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) models of visual word recognition to particular cortical areas. These models differ in the number of nonsemantic routes available for the computation of phonology from orthography as well as in their approach to the individuation of lexical items.

Specifically, the DRC posits two nonsemantic routes: One route applies sublexical rules to map graphemes (i.e., letters or letter combinations) to phonemes in a serial fashion, from left to right. Another route involves parallel activation of lexical (i.e., word-specific) nodes represented in an orthographic and a phonological lexicon. In this model, words are individuated in the lexical route insofar as each node represents a particular word. In contrast, there are no lexical nodes and no absolute rules in the triangle model. Instead, distributed orthographic representations map onto distributed phonological representations. In this model, words are not individuated by implementation. Rather, support from a distributed semantic route is thought to contribute to the differentiation among orthographic and phonological neighbors. Thus, these models can be potentially distinguished on the basis of the existence of (a) an orthographic lexicon and (b) a serially-operating sublexical grapho-phonemic conversion (GPC) system.

In a meta-analysis of 35 neuroimaging studies, Jobard, Crivello, and Tzourio-Mazoyer (2003) classified contrasts as primarily exposing the lexical or the sublexical route, aiming to identify components of a dual-route reading mechanism. They found little evidence for lexical route components (e.g., an orthographic lexicon). For example, no clusters were activated predominantly by contrasts such as words versus pseudowords. They proposed that access to the meanings and properties of words is provided directly from sublexical to phonological and semantic areas. In contrast, components of a sublexical route were revealed, including clusters with relatively higher involvement in contrasts such as pseudowords versus fixation, located in the superior and middle temporal gyri and in the supramarginal and inferior frontal gyri. Jobard et al. considered these clusters as indicative of a network associated with phonological processing and phonological memory, which are considered necessary for mapping graphemes to phonemes. However, such a general analysis cannot reveal particular properties of reading-specific operations or provide clear support for any one computational model of reading.

Considering the potential for differential predictions of the two models in terms of brain activation, Binder, Medler, Desai, Conant, and Liebenthal (2005) suggested that their difference lies in whether word reading is supported by orthographic representations (i.e., an orthographic lexicon) or by semantic representations. In a neuroimaging study contrasting regular and irregular words and pseudowords, they manipulated word imageability to reveal areas related to semantic processing, and also identified areas modulated by within-condition response time as related to generic task requirements such as memory, attention, and executive control. They found that regions showing more activation for words than for pseudowords were associated with semantic processing rather than with word-level codes, implying a stronger

correspondence with the triangle than with the DRC model, although they noted that the findings were consistent with both models.

A number of studies examining regional activation (e.g., Fiebach, Friederici, Müller, & von Cramon, 2002; Graves, Desai, Humphries, Seidenberg, & Binder, 2010; Mei et al., 2014; but cf. Mechelli, Gorno-Tempini, & Price, 2003), functional connectivity (e.g., Boukrina & Graves, 2013; Levy et al., 2009; Mechelli et al., 2005; Richardson, Seghier, Leff, Thomas, & Price, 2011; Simos, Rezaie, Fletcher, & Papanicolaou, 2013), effects of electrostimulation (Roux et al., 2012), and individual differences in reading skill (e.g., Graves et al., 2014; Reilhac, Peyrin, Démonet, & Valdois, 2013; Seghier, Lee, Schofield, Ellis, & Price, 2008; Wimmer et al., 2010) have identified multiple processing streams associated with reading (see Price, 2012, and Paulesu, Danelli, & Berlingeri, 2014, for reviews). However, these findings cannot be interpreted as preferentially supporting the DRC or the triangle model, because the studies were not designed to assess distinguishing features of the two computational approaches.

More recently, Taylor, Rastle, and Davis (2013) derived neural activation predictions from the two computational models on the basis of an engagement-effort framework, positing that (a) stimuli engage regions with relevant representations and (b) more effort is associated with increased neural activation. This results in an inverted-U-shaped function linking fit to representation with predicted BOLD response. For example, both high and low frequency words would engage regions supporting lexical representations more than pseudowords; but activation is expected to be higher for low frequency words, due to increased effort in retrieving them. Applying this rationale to irregular and regular words and pseudowords being processed by the two models, Taylor et al. derived predictions of activation that would be consistent with particular model components.

In a meta-analysis of 36 studies using activation likelihood estimation (ALE), they identified regions putatively subserving the functions of model components as well as several clusters not corresponding to model components. Predictions for the phonological, orthographic, and semantic layers of the triangle model were consistent with activation in the left inferior frontal gyrus, posterior fusiform and occipitotemporal gyri, and middle temporal and angular gyri, respectively (as also suggested by [Carreiras, Armstrong, Perea, & Frost, 2014](#)). These regions were also associated with the phoneme output, letter input, and semantic systems of the DRC model. As stated earlier, the distinguishing features of the two models are the serial (DRC model) versus parallel (triangle model) spelling-to-sound conversion systems and the implementation of phonological and orthographic lexica in the DRC but not in the triangle model. However, contrasts between words and pseudowords cannot distinguish between serial and parallel processing or between lexical and semantic processing. Thus, the two models were deemed equally plausible given the data included in this meta-analysis.

In the present study we follow up on the localization suggestions of [Taylor et al. \(2013\)](#) and examine predictions from the DRC model using parametric modulation of activation. Specifically, we retain from Taylor et al. (and previous studies) the following assumptions: First, the direct, lexical route must be among the areas significantly activated by word processing. Hence it should be revealed by a contrast between word and pseudoword conditions, especially when the pseudowords have few or no neighbors. And second, the indirect, grapho-phonemic conversion route must be among the areas significantly activated by pseudoword processing. Hence it should be revealed by a contrast between pseudoword and fixation conditions. We chose not to use a contrast between pseudoword and word conditions because the sublexical route may be equally engaged with words and pseudowords (depending on the specific stimuli

used and on the balance between the routes in a given orthography) and thus might not be revealed by this contrast. In addition, we examine the localization suggestions that the orthographic and phonological lexica are located in the anterior fusiform and posterior middle temporal gyri, respectively, whereas the sublexical, serial rule-based GPC system is located in the inferior parietal cortex, dorsal to the angular and supramarginal gyri.

The complementary hypothesis proposed in this study concerns the theoretically fully separable effects of pseudoword length and word frequency. First, according to the DRC model, length effects in pseudoword reading arise from the serial operation of the GPC system. Therefore, modulation of activation by pseudoword length must co-localize with this GPC system. That is, it must overlap with at least one significant cluster from the contrast of pseudowords versus fixation. Based on the conclusions of [Taylor et al. \(2013\)](#), this cluster should be found in the left inferior parietal cortex. Second, in the DRC framework, word frequency effects necessarily arise in the orthographic and the phonological lexicon. Therefore, modulation of activation by word frequency must co-localize with these lexica. That is, it must overlap with at least two significant clusters from the contrast of words versus pseudowords: one for the orthographic and another for the phonological lexicon. Based on the analysis of [Taylor et al. \(2013\)](#), the former cluster should be found in the anterior fusiform gyrus, whereas the latter cluster should be found in the angular or middle temporal gyrus.

In the present study we report results from a lexical decision task using a set of words and pseudowords with decorrelated variables ([Protopapas & Kapnoula, 2013](#)), following [Graves et al. \(2010\)](#) but not limited to monosyllables and the restricted variable ranges they might entail. The choice of task is not expected to limit our conclusions, as it has been demonstrated that patterns of activation are very similar for

lexical decision and reading aloud except for (theoretically less interesting) sensorimotor areas associated with articulation and hearing ([Carreiras, Mechelli, Estévez, & Price, 2007](#)). Moreover, the operation of the DRC model components does not differ between naming and lexical decision tasks. In particular, both routes process stimuli in lexical decision as they do in naming. The only difference between the two tasks as implemented in the DRC is the application of decision criteria, based on lexical activation thresholds, and decreased letter-to-word inhibition (see discussion in [Coltheart et al., 2001](#), pp. 226–231), which are not relevant to our study because we do not simulate response times but only seek to associate neural activation with model components.

The study of activation modulation by lexical and sublexical variables in reading tasks is not novel. Effects on activation related to word processing by variables such as frequency and length have been reported in several studies of orthographic and lexical processing in visual word recognition. Hauk, Davis, and Pulvermüller (2008) combined correlated sets of variables into three main groups, namely frequency, typicality (based on bigram and trigram frequencies), and length. They found effects of all groups over several regions but no overlap between frequency effects and activity for words versus a baseline condition. Yarkoni, Speer, Balota, McAvoy, and Zacks (2008) used a word-by-word text reading task and identified several regions that were sensitive to differences in multiple lexical variables—namely, word frequency, length, orthographic neighborhood size, lexical decision and naming latency, and position in sentence—in addition to a contrast between reading and fixation. These regions included the left inferior frontal and middle temporal gyri, temporoparietal and inferotemporal cortex, as well as the precuneus, posterior cingulate, and dorsomedial prefrontal cortex. [Graves et al. \(2010\)](#) used a set of 465 monosyllabic words, selected to decorrelate frequency,

consistency, imageability, and bigram and biphone frequency, in a reading task. They also examined the correlation of activation with response latency, aiming to disentangle specific reading processes from those related to more general aspects of performance. They found an overlap between frequency and imageability effects, indicative of lexical processing, bilaterally in the precuneus and angular gyri, unmodulated by response time. Other studies have also reported effects of variables such as frequency (Fiebach et al., 2002; Kronbichler et al., 2004), orthographic neighborhood size (Fiebach, Ricker, Friederici, & Jacobs, 2007), length (Schurz et al., 2010), and orthographic typicality (Woollams, Silani, Okada, Patterson, & Price, 2011). Beyond various inconsistencies among the findings, these studies show that investigation of parametric modulation by lexical and sublexical variables is practically feasible and may provide useful constraints for the evaluation of reading models. However, these studies were not specifically designed to test predictions from computational models of reading.

In the analyses presented below, we focus on the left hemisphere only, following Taylor et al. (2013) and reviews of language processing and reading networks (Carreiras et al., 2014; Paulesu et al., 2014; Price, 2012). We work in the Greek language, which has a well understood, relatively transparent orthographic system (Protopapas & Vlahou, 2009), aiming to extend the range of investigation towards more inclusive cross-linguistic validity. Our main question is: Can we identify the purported orthographic lexicon and grapho-phonemic converter of the DRC based on their properties as postulated by the model?

Method

Participants

The sample included 44 adults (31 women), recruited through the University community (age $M = 29.5$ years, $SD = 5.8$, range 19–47). All were right-handed, native

speakers of Greek who attended Greek school since Grade 1, could read comfortably without glasses (or with contact lenses) and reported no history of reading difficulties.

Materials

The stimuli included 150 words and 150 pseudowords matched on length, bigram and syllable frequency, and orthographic transparency. Words 2–5 syllables long were selected from the C corpus of the ILSP PsychoLinguistic Resource (speech.ilsp.gr/iplr; Protopapas, Tzakosta, Chalamandaris, & Tsakoulis, 2012), spanning a wide range over several target variables. Table 1 lists their descriptive statistics in comparison to the entire corpus. Orthographic and phonological syllable and bigram frequency refer to the mean logarithmic token frequency of (position-independent) syllables or symbol pairs (letters or phonemes), respectively, in occurrences per million tokens. Orthographic and phonological neighborhood counts refer to Coltheart's N, that is, the number of words with the same length that differ by only one letter or phoneme, respectively ([Coltheart, Davelaar, Jonasson, & Besner, 1977](#)). Graphophonemic consistency was computed as the logarithmic mean of nondirectional token "sonograph" probabilities, that is, ratios of specific grapheme-phoneme mappings over the total number of grapheme-phoneme tokens ([Spencer, 2009](#)). In an iterative process, items were selected and a nonparametric index of association (Spearman's ρ) among all variables was calculated; the process terminated when groups of qualitatively distinct variables were not significantly correlated. The final intercorrelations among variables in the selected items are shown in Table 2.

A set of 150 pseudowords were constructed to resemble the words in basic phonological and orthographic structure and letter and phoneme distribution. The pseudowords were indistinguishable from the words in the target variables, as verified by the Anderson-Darling test for equality of distributions (from R package kSamples;

Scholz & Zhu, 2015). The results of these tests are also listed in Table 1 and the intercorrelations are shown in Table 2. See Protopapas and Kapnoula (in press) for more information regarding stimulus selection and distribution.

Pseudoword neighborhoods were excluded from matching and intercorrelation requirements because neighbors of long pseudowords would look distractingly similar to existing words. Moreover, due to the rich inflectional morphology of Greek, pseudoword neighbors would be likely to be inflectional variants of a single base form, likely resulting in undue activation of specific lexemes by pseudoword stimuli. Thus pseudowords were constructed with as few neighbors as possible, to minimize lexical activation from pseudoword stimuli that might obscure the word-pseudoword contrast.

Behavioral validation for this stimulus set and the expected effects on response times of the selected variables from a relatively large and diverse adult participant group can be found in Protopapas and Kapnoula (2013, in press).

Procedure

Stimulus presentation and response collection was controlled by DMDX ([Forster & Forster, 2003](#)). Participants performed a lexical decision task in the scanner, in three blocks of 150 trials each. In addition to the 150 words and 150 pseudowords there were also 150 “fixation” trials consisting in a pair of crosses (plus signs). Participants were instructed to respond to words by pressing a button with their right index finger and to pseudowords or crosses by pressing another button with their right middle finger. They were asked to respond as soon as possible and not to dwell on or try to pronounce the stimuli either aloud or silently. The order of trials was randomized for each participant.

Each trial was initiated by a volume onset trigger from the scanner. After a random duration ranging continuously between 0–4 s, a stimulus (word, pseudoword,

or cross pair) was projected onto a semi-translucent PVC board mounted near the participant's feet, at a distance of 218 cm (measured from the eyes), in a fixed-width font (Courier New). Each letter occupied 2.5 cm, corresponding to 0.66 degrees of visual angle at center field. The stimulus remained on the screen until the participant's response, for a maximum period of 5 s, at which point a timeout was recorded and the procedure moved to the next trial. Figure 1 displays a schematic of the trial sequence. The time interval between successive trial onsets was just over 5 s on average (participant mean, $M = 5148$ ms, $SD = 143$, range 4905–5528). The entire duration of the study (all 3 blocks) was approximately 40 min (participant mean, $M = 41.5$, $SD = 1.3$, range 39–45).

MRI data acquisition

Data were collected in a 3T Philips Achieva TX (Philips Healthcare, Best, The Netherlands) whole-body MR scanner using an 8-channel phased-array head coil with SENSE factor 2.0. Functional scans were collected in 3 successive runs using gradient-echo EPI sequences acquired with k-space matrix size 64×80 , field of view (FOV) 192×240 mm 2 , including 36 sequential (bottom up) transverse slices parallel to the anterior commissure–posterior commissure line, with TR = 2000 ms, TE = 30 ms, flip angle = 90°, slice thickness 3 mm, interslice gap 0 mm.

After the functional runs, a high-resolution structural anatomical image was acquired using a 3D IR prepared turbo field echo T1 weighted scan, in the same orientation as the EPI scans, with 140 slices, TR = 9.9 ms, TE = 4.6 ms, flip angle = 8°, FOV = 240×240 mm 2 , at 1×1×1 mm resolution.

Scanner trigger signals and participant responses were input to the experimental computer via a Cedrus Lumina LSC-400(A) serial response box equipped with LU400-PAIR two-button response pads.

Behavioral data analysis

Response times, for correct responses only, were logarithmically transformed (to better approximate the normal distribution; [Baayen & Milin, 2010](#)) and analyzed with general linear mixed-effects models with crossed random effects for participants and items ([Baayen, 2008](#); [Baayen, Davidson, & Bates, 2008](#)) using function lmer of the lme4 package ([Bates, Maechler, Bolker, & Walker, 2013](#)) in R ([R Core Team, 2013](#)). The random effects structure included per-participant slopes for block, within-block trial order, and RT to the preceding item; as well as per-item slopes for block. In addition to the experimental set of 6 lexical and sublexical parameters, the fixed effects model specification included the trial order interacting with block and the RT to the preceding item interacting with its lexicality (word, pseudoword, crosses). All numeric predictors were centered. Nominal variables (block and previous stimulus type) were deviation coded using contr.sum. Effect significance was evaluated based on Satterthwaite's approximation using package lmerTest ([Kuznetzova, Brockhoff, & Christensen, 2013](#)).

MRI data analysis

Data was processed using SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK; <http://www.fil.ion.ucl.ac.uk/spm/software/spm8/>) with default parameters except as noted below. Quality checking and preprocessing also involved scripts from the ArtRepair toolbox ([Mazaika, Hoeft, Glover, & Reiss, 2009](#); [Mazaika, Whitfield-Gabrieli, & Reiss, 2007](#)) and Alphascript (Marzelli, Hoeft, Mazaika, & Sheau, 2009; provided by P. Mazaika, personal communication, 15 November 2013).

Raw slices were examined with art_movie, in high contrast view, to identify major artifacts. No participants were rejected. Slice correction was then performed using alpha_art_slice (reference volume 10, rejection threshold 5), rejecting a small proportion of slices (per participant $M = 0.6\%$, $SD = 1.7$, range=0.0–15%), followed by

SPM slice timing correction (microtime resolution 36, reference frame 18) and spatial realignment (estimate and reslice). Realigned images were voxelwise clipped to 8% and high-pass filtered with a 37-tap filter to remove slow variations using art_despike (filter type 2). Subsequently, images were smoothed with a 4-mm FWHM Gaussian kernel and voxelwise signal intensities were adjusted by regressing onto six motion correction regressors, applied more strongly near image edges, using alpha_motionJ. The T1 image was coregistered to the mean EPI image using SPM coregister (estimate) and then simultaneously segmented, bias-corrected, and normalized to the MNI template using SPM8 default tissue probability maps. Deformation fields were then applied to the motion-adjusted EPI scans to bring them into registration with the MNI space, resampled at 3×3×3 resolution. Finally, the normalized images were smoothed with a 7-mm FWHM Gaussian kernel.

In the first-level analysis for each participant, each stimulus was modeled as an event of the corresponding type (word, pseudoword, fixation) with duration equal to its presentation time (i.e., the participant's response time for the specific item). Temporal and dispersion derivatives for each event type were included in the model in addition to the SPM canonical (dual-gamma) hemodynamic response function (hrf). Motion regressors were not included because residual motion artifacts were removed in preprocessing. Examination of a basic contrast ($0.5 \times \text{word} + 0.5 \times \text{pseudoword} - 1.0 \times \text{fixation}$) confirmed significant activation clusters for each participant (at $p < .001$, uncorrected), typically including ventral occipitotemporal as well as frontal regions. Subsequently, two sets of contrasts were taken to the second (group) level: pseudowords vs. fixation and words vs. pseudowords. In addition, a second set of first-level analyses included one of the six experimental parameters as a parametric modulator on the corresponding (word or pseudoword) event. All parameters were

centered prior to the analysis. For each of these contrasts the parametric modulation on the hrf was taken to the group level.

For graphical display of the results, spatially normalized data (T maps) were visualized onto an inflated cortical surface map from the Human PALS-B12 Atlas (Van Essen, 2005; Van Essen & Dierker, 2007) using CARET v5.65 (Harwell, Drury, Hanlon, & Van Essen, 2012; Van Essen et al., 2001). Anatomical labels were provided by the automated anatomical labeling (aal) toolbox (Tzourio-Mazoyer et al., 2002) through xjView (<http://www.alivelearn.net/xjview>).

Results

Behavioral data

Table 3 lists descriptive statistics for speed and accuracy of responding to the stimuli by the 44 participants. For words, the analysis indicated that frequency ($\beta = -2.52 \times 10^{-2}$, $t = -6.64$, $p < .001$), length ($\beta = 1.61 \times 10^{-2}$, $t = 4.34$, $p < .001$), neighborhood size ($\beta = -9.82 \times 10^{-3}$, $t = -2.01$, $p = .040$), and syllable frequency ($\beta = 2.75 \times 10^{-3}$, $t = 2.27$, $p = .025$) had a significant effect on word lexical decision RTs, whereas bigram frequency ($\beta = 1.45 \times 10^{-2}$, $t = 0.54$, $p = .591$) and orthographic transparency ($\beta = 5.89 \times 10^{-2}$, $t = 0.68$, $p = .498$) did not.

For pseudowords, length ($\beta = 3.73 \times 10^{-2}$, $t = 10.06$, $p < .001$) and neighborhood size ($\beta = 1.28 \times 10^{-2}$, $t = 2.28$, $p = .024$) had a significant effect on lexical decision RTs, whereas syllable frequency ($\beta = 1.52 \times 10^{-3}$, $t = 1.38$, $p = .171$), bigram frequency ($\beta = 1.01 \times 10^{-2}$, $t = 0.55$, $p = .583$) and orthographic transparency ($\beta = -5.54 \times 10^{-2}$, $t = -0.76$, $p = .452$) did not.

Even though the pseudowords were constructed based on real words and were matched to the words in bigram and syllable frequency, the fact that they were not matched in neighborhood size leaves open the possibility that they may have been

sufficiently dissimilar to the words to bias the lexical decision task and allow it to be performed without reliance on lexical activation. To alleviate this concern we have used the LD1NN algorithm to quantify the bias inherent in our stimulus set (Keuleers & Brysbaert, 2011), applied to the specific stimulus sequence delivered to each of the 44 participants using R package vwr (Keuleers, 2013). No significant bias was detected (mean odds = 0.80, range 0.69–0.94; mean $z = -1.25$, range -1.99 to -0.40). In particular, cumulative average bias was negligible for words, whereas pseudowords were somewhat biased toward words, confirming their word-like construction.

Including the word probabilities estimated by the LD1NN algorithm in the RT analysis models produced no significant effects for words ($\beta = -2.62 \times 10^{-3}$, $t = -0.40$, $p = .693$) and no interaction with trial order ($\beta = 1.60 \times 10^{-3}$, $t = 0.274$, $p = .784$). In contrast, there was a significant effect of estimated word probability on pseudoword RT ($\beta = 2.58 \times 10^{-2}$, $t = 3.64$, $p < .001$) and an interaction with trial order ($\beta = 1.22 \times 10^{-2}$, $t = 2.05$, $p = .041$), consistent with slower responses to the most word-like pseudowords, especially in later trials. Thus, we feel confident that performing our lexical decision task did require lexical activation, giving us adequate power to localize the lexicon.

Whole-brain MRI data

Voxelwise significance was corrected for familywise error (FWE) at $p < .05$; associated statistical inference information is shown in Table 4. Note that, due to signal dropout, no activity can be detected at the ventral temporal pole or orbitofrontal cortex by our analyses. Table 5 lists the significant clusters (of minimum size 4 voxels) and peak coordinates (up to 3 peaks per cluster) for the two main contrasts, namely pseudowords vs. fixation and words vs. pseudowords, and parametric modulators, namely number of letters (modulating pseudowords) and frequency (modulating words). Note that the extent of differentially activated regions is larger than usually

reported. This is because of the greater sample size, consistent with the demonstration of [Thyreau et al. \(2012\)](#) that the proportion of gray matter exceeding a given threshold increases as a function of sample size. There were no significant effects of orthographic neighborhood, bigram frequency, graphophonemic transparency, or pseudoword syllable frequency, so these variables will not be considered further.

Presentation and discussion of results will henceforth consider left-hemisphere lateral and ventral cortex only, as this is where the “reading network” is thought to be located and where previous attempts to localize cognitive model components have focused. Figure 2 (top row) displays significant contrast clusters (FWE-corrected), for the left hemisphere only, on lateral and ventral views. To explore the possibility that clusters may have been missed due to stringent statistical correction, Figure 2 also shows clusters thresholded by an uncorrected criterion ($p < .001$).

Contrasts between stimulus types. Significantly greater signal for pseudowords, compared to fixation, was detected in large clusters, posteriorly over both dorsal and ventral visual streams, from the calcarine sulcus through the lingual, inferior occipital, fusiform, and inferior temporal gyri, as well as through middle and superior occipital and inferior and superior parietal gyri; and anteriorly, over the inferior frontal gyrus, extending to the anterior insula, anterior Rolandic operculum, precentral and postcentral gyri. Significantly less signal for pseudowords than fixation was detected in the middle occipital, middle temporal, angular, supramarginal and adjacent inferior parietal gyri, as well as in the middle and superior frontal gyri, and in the posterior insula, posterior Rolandic operculum, and superior temporal gyrus.

In the words vs. pseudowords contrast, significantly greater signal for words was detected in the middle occipital, inferior, middle, and superior temporal, angular, supramarginal and adjacent inferior parietal gyri, in the middle and superior frontal

gyri, in the postcentral and adjacent superior parietal gyri, as well as in the insula and inferior frontal gyrus. Significantly less signal for words than pseudowords was detected only in the precentral and postcentral gyri.

Parametric modulation. The parametric modulation analysis revealed significant effects for number of letters in pseudowords. Positive effects, that is, greater signal for longer pseudowords, were detected in the calcarine sulcus and the lingual, inferior and middle occipital, and fusiform gyri, as well as in the precentral and postcentral gyri. An additional cluster in the middle occipital gyrus exceeded the uncorrected ($p < .001$) threshold (peak at $-30 -82 +10$). No negative effects exceeded the FWE-corrected threshold; however, a cluster in the middle temporal gyrus ($-54 -58 -5$) and another in the postcentral gyrus ($-51 -25 +49$) exceeded the uncorrected threshold.

There were no positive effects for word frequency (i.e., greater signal for higher-frequency words) at the FWE-corrected threshold. However, a cluster in the middle temporal gyrus ($-48 -61 +4$) and another in the inferior parietal lobule ($-39 -43 +58$) exceeded the uncorrected threshold. Negative word frequency effects were found in the inferior frontal gyrus.

Effect overlap. Significant clusters from the pseudoword vs. fixation contrast, including positive and negative effects, were used to form an inclusive mask image, within which we evaluated the number of letters contrasts. Likewise, significant clusters from the word vs. pseudoword contrast, including positive and negative effects, were used to form an inclusive mask image, within which we evaluated the word frequency contrasts. Common regions between pseudowords vs. fixation and number of letters in pseudowords included a 167-voxel cluster in the calcarine sulcus, lingual, inferior and middle occipital, and fusiform gyri, and a 38-voxel cluster in the precentral

and postcentral gyri. There were no common regions at the FWE-corrected threshold between words vs. pseudowords and word frequency, except for a single voxel in the left insula.

Duration modeling. To examine the potential effect of duration modeling in masking time-sensitive regions involved in word processing (cf. Taylor, Rastle, & Davis, 2014), Figure 3(A) displays the (FWE-corrected) significant clusters from an additional analysis without duration modeling (duration simply set to 0 for all three types of events), superimposed over the results with duration modeling. The only substantial discrepancies were seen in the words vs. pseudowords contrast, in which duration modeling apparently masked a significant 27-voxel cluster in the inferior occipital gyrus (peak $T = -6.27$; coordinates $-36 -88 -2$) but also allowed the significant 120-voxel cluster in the postcentral and inferior parietal gyri (peak $T = 7.47$; coordinates $-51 -22 +49$) to emerge.

Word-related contrasts. Beyond the critical tests regarding localization of the DRC model components, Figure 2 (bottom row) and Table 6 display the results of contrasts related to word processing, in order to facilitate the interpretation of the activation and deactivation patterns. The results of the words vs. fixation contrast are largely similar to those of the pseudowords vs. fixation contrast, including significantly greater signal for words, compared to fixation, in large clusters, posteriorly over both visual streams, from the calcarine sulcus through the lingual, inferior occipital, fusiform, and inferior temporal gyri, as well as through middle and superior occipital and inferior and superior parietal gyri; and anteriorly, over the inferior frontal gyrus, extending to the anterior insula, Rolandic operculum, precentral and postcentral gyri. Significantly less signal for words than fixation was detected in the middle occipital, middle temporal, and angular gyri, as well as in the middle frontal gyrus and in the posterior

insula and superior temporal gyrus.

The parametric modulation analysis revealed significant effects for number of letters in words. Positive effects, that is, greater signal for longer words, were detected in the calcarine sulcus and the lingual, inferior, middle, and superior occipital, and fusiform gyri, as well as in the precentral and postcentral gyri. No negative effects exceeded either the corrected or the uncorrected threshold. Direct comparison between length effects in words vs. pseudowords produced no significant clusters.

Positive effects for syllable frequency in words (i.e., greater signal for words with more frequent syllables) were found in the inferior frontal gyrus. In addition, a cluster in the fusiform gyrus (-45 -58 -14) and another in the inferior parietal lobule (-42 -46 +43) exceeded the uncorrected threshold. There were no negative effects using the corrected threshold; a cluster in the angular gyrus (-45 -73 +31) exceeded the uncorrected threshold.

ROI analyses

Whole-brain analyses failed to identify regions with common significant effects as theoretically predicted. However, the reason may be that variation between individual participants may have obscured patterns of co-localization. For example, the “graphophonemic conversion” region may be in the same gyrus in every person but its exact coordinates may vary between persons beyond the group smoothing window, and therefore it may be missed in the group analysis. To investigate this possibility a set of ROI analyses was performed.

First, group regions of interest were defined by the conjunction of a peak-related sphere and an anatomically defined gyrus. Specifically, three local peaks were selected, in theoretically relevant locations in the left hemisphere cortex: Fusiform gyrus (pseudowords vs. fixation peak at -42 -52 -20), middle temporal gyrus (word vs.

pseudoword peak at $-45 -58 +13$, and inferior/superior parietal gyrus (pseudoword vs. fixation peak at $-27 -58 +46$). Spheres with a 10 mm radius at or near these peaks (slightly displaced, to maximize overlap with the target gyrus) were then defined using MarsBaR r. 0.43 (<http://marsbar.sourceforge.net/>; Brett, Anton, Valabregue, & Poline, 2002) and masked with the corresponding labeled anatomical region in the aal atlas (file ROI_MNI_V4.nii), so that the resulting regions would lie entirely within the respective gyri. Specifically, a sphere centered at $-42 -49 -18$ was masked with the left fusiform gyrus, a sphere centered at $-48 -58 +13$ was masked with the left middle temporal gyrus, and a sphere centered at $-27 -58 +49$ was masked with the conjunction of the left inferior and superior parietal gyri. Figure 3(B) illustrates the three mask images over the inflated cortical surface, for comparison with the whole-brain results.

Within each group region of interest, individual peaks (for each participant) were identified for the main contrasts (pseudoword vs. fixation and word vs. pseudoword) using a custom Matlab script. Individual regions of interest were then defined by $3 \times 3 \times 3$ voxel blocks around the individual peaks, using MarsBaR. In this way we identified each participant's individual contrast peaks within the group regions. The size of the individual ROIs approximated the effective smoothing window in preprocessing, to maximize our power to identify highly localized but individually varying effects. Finally, the critical parametric modulation contrasts (number of letters and word frequency) were evaluated within the individual ROIs using MarsBaR, and the derived contrast values were submitted to t tests against zero. In this way we could examine whether, for example, number of letters had a significant effect in a 27-voxel region around each participant's own pseudoword vs. fixation peak in the parietal cortex.

Although the same functional runs were used for the definition of the ROIs and

the parametric analyses, our approach does not constitute “double dipping” ([Kriegeskorte, Simmons, Bellgowan, & Baker, 2009](#)) because all pseudowords (i.e., of all lengths) and all words (i.e., of all frequencies) were included in the functional contrasts used to define the ROIs (pseudoword vs. fixation and word vs. pseudoword). That is, the factors analyzed on the ROIs were defined *within* the factors contrasted to define the ROIs. Differences in the length or frequency effects concern variation in activation among individual items, neither affecting nor affected by the condition means, because parametric contrasts were centered and between-items variability was ignored in the main contrasts (effectively averaging over items, to permit 2nd-level contrasts with respect to between-participants variance).

Table 7 presents the results of these analyses. Bonferroni adjustment for 24 comparisons sets α to .002; however, any adjustment of this sort will be too conservative as not all 24 comparisons were theoretically important to evaluate and are only shown for completeness. Figure 4 shows the most important comparisons. In the fusiform gyrus, there was no significant effect of word frequency around the individual word vs. pseudoword peaks and no significant effect of pseudoword length around the individual pseudoword vs. fixation peaks. The effect of frequency at the individual middle temporal word vs. pseudoword peak approached significance. This was a *positive* effect, that is, a trend for higher activation associated with higher frequency. The effect of the number of letters at the individual parietal pseudoword vs. fixation peaks was significant and positive. There was also a significant negative effect of word frequency at the same peaks. Therefore, higher parietal activation around the pseudoword vs. fixation peak (which is also around a word-fixation peak, cf. Table 6) was associated with longer pseudowords as well as with lower-frequency words.

Discussion

The primary goal of the analyses was to determine whether word frequency effects co-localize with the words vs. pseudowords contrast and pseudoword length effects co-localize with the pseudowords vs. fixation contrast, as theoretically predicted by the function of the DRC lexical and sublexical route components. Despite a relatively large participant sample and well-controlled stimulus set, whole brain analysis revealed that there was no overlap between words vs. pseudowords and word frequency and that overlap between pseudowords vs. fixation and number of letters in pseudowords was limited to regions associated with visual and articulatory processing. Thus, neither an orthographic lexicon nor a serial mechanism applying grapheme-to-phoneme rules was revealed by the whole-brain analysis.

One potential criticism is that lexical decision does not require computation of a phonological code and therefore might not activate the GPC system. However, naming and lexical decision have been shown to activate the same regions (except sensory and motor components associated with speech production; Carreiras et al., 2007). Moreover, pseudowords activated the precentral gyrus (associated with articulation) more than words, and the same precentral region showed positive pseudoword length effects (greater signal for longer pseudowords). This suggests that a phonological code was indeed computed in our lexical decision task.

A second possible criticism is that the relatively transparent Greek orthography may encourage a sublexical mapping for both pseudowords and words. This interpretation seems unlikely for the following reasons. First, as in other languages, there was strong activation for words relative to pseudowords in regions hypothesized to form a semantic network (discussed below). Second, even though Greek is relatively consistent, it does not have a one-to-one mapping between letters and phonemes

(Protopapas & Vlahou, 2009). Instead, there are several ambiguous graphemes and, moreover, there are alternative spellings for the same pronunciation, some of which may constitute distinct words (homophones). Persistent and pervasive spelling difficulties (Protopapas, Fakou, Drakopoulou, Skaloumbakas, & Mouzaki, 2013) suggest that orthographic knowledge is necessary in Greek as in less transparent orthographies. Thus there is no reason to suppose that words are read sublexically in Greek.

In the following sections we examine whether the ROI analyses centered on inferior/superior parietal gyrus (pseudoword vs. fixation peak at -27 -58 +46), middle temporal gyrus (word vs. pseudoword peak at -45 -58 +13), and fusiform gyrus (pseudowords vs. fixation peak at -42 -52 -20) are more sensitive than the whole-brain analysis and support the existence of (a) a sublexical serial processor, (b) a phonological lexicon, and (c) an orthographic lexicon, respectively.

Dorsal parietal cortex

There was a significant effect of pseudoword length in regions of interest formed around individual participant peaks of the pseudowords vs. fixation contrast in the parietal cortex. This effect was located near the cluster identified by Taylor et al. (2013) as a candidate region for the grapheme-phoneme conversion system. It was accompanied by a negative word frequency effect, which would seem incompatible with a dedicated sublexical processing route that has no lexical knowledge. However, this objection can be alleviated by hypothesizing that sublexical conversion from graphemes to phonemes is aborted as soon as the lexical route identifies a word, as this should occur earlier for more frequent words, resulting in reduced activation. Although the model currently lacks backward projections from the phonological output buffer to the sublexical GPC route, such an amendment seems consistent with the conception of the model (cf. Coltheart et al., 2001, section on Future Developments, p. 248).

A further concern with the idea that dorsal parietal cortex plays a role in GPC is that only a few studies in the meta-analysis of [Taylor et al. \(2013\)](#) reported significant clusters in this region for the pseudowords vs. words contrast (Table 3, p. 16). A significant meta-analytic cluster may result from the smoothing operation of activation likelihood estimation over neighboring peaks from non-overlapping clusters in the original studies. It is likely that some of the peaks reported in individual studies did not fall within the meta-analytic cluster but did contribute to its location and size. Thus, on the one hand, more studies than reported in the table may have obtained a significant effect in the inferior/superior parietal cortex for the pseudoword vs. word contrast; but on the other hand, cluster localizations must have been inconsistent between studies for this to occur.

The region in question overlaps with clusters found to be active during reading, in both adults and children, in the meta-analysis of Martin, Schurz, Kronbichler, and Richlan ([2015](#)). However, it was not included in the otherwise very widely distributed set of left-hemisphere regions discussed in a recent comprehensive review and synthesis of PET and fMRI studies related to speech, language, and reading ([Price, 2012](#)), suggesting that it may not be universally associated with reading processes. A “dorsal pathway” is frequently identified in the literature as part of the “reading network” (and discussed in the review of [Price, 2012](#)), associated with graphophonemic or sublexical processing, typically said to be located at the angular and supramarginal gyri (in addition to the posterior superior temporal gyrus; [Pugh et al., 2000, 2001](#)). Our region of interest specifically excludes the angular and supramarginal gyri, lying dorsally to both of them and apparently within the dorsal visual stream. However, anatomical labeling may not always be consistent, and it seems likely that the precise localization of activation clusters in previous studies may extend dorsally beyond the

supramarginal gyrus. This would explain the appearance of this region in meta-analyses ([Martin et al., 2015](#); [Taylor et al., 2013](#)) as well as in our study.

We can attempt to estimate the selectivity of this region by considering the kinds of contrasts in which it is activated ([cf. Poldrack, 2006](#)). For an unbiased—though necessarily limited—indication of processes associated with this region, we performed a search through the BrainMap database for fMRI activation around MNI coordinates – 27 –58 +47 (a rectangular Talairach ROI from –29 –62 +40 to –24 –58 +47, transformed from MNI using an online application from the Yale BioImage Suite Package at <http://noodle.med.yale.edu/~papad/mni2tal/>, based on Lacadie, Fulbright, Rajeevan, Constable, & Papademetris, 2008) using the Sleuth client v. 2.3 ([Fox et al., 2005](#); [Laird, Lancaster, & Fox, 2005](#)) in February 2015, which returned 14 studies including 80 experiments. Contrasts resulting in significant activation of this area concerned a wide variety of experimental conditions, ranging from finger tapping and nociception to working memory and reasoning, most of which included a spatial visual component in carrying out the task. There was no evidence that an operation akin to rule-based transformation might underlie even a subset of the returned results from this search.

[Cohen, Dehaene, Vinckier, Jobert, and Montavont \(2008\)](#) presented to their participants words that were difficult to recognize due to visual degradation (by rotation, increased spacing, or off-center location). They found an effect of degradation bilaterally in the posterior intraparietal cortices, substantially overlapping with our parietal region of interest. Cohen et al. associated this activation with an “attention-based serial reading strategy” (p. 361) and attributed the specific involvement of this region to spatial attentional processes that are “distinct from those subtending normal reading” (p. 363). This interpretation is consistent with a multitude of findings that have associated this region with top-down visual spatial attention ([Gottlieb, 2007](#)). In

the context of reading normal, undegraded stimuli, spatial attention is likely relevant for processing letters (e.g., for efficient processing of multiple letters; cf. Reilhac et al., 2013) prior to their mapping to phonemes. However, the visual operation of spatial attention is distinct from GPC mapping: In the DRC model, all letters are hypothesized to be activated simultaneously (a view supported by recent behavioral and neural findings: [Adelman, Marquis, & Sabatos-DeVito, 2010](#); [Forget, Buiatti, & Dehaene, 2010](#)). The serial operation that causes length effects is posited to originate not in the visual processing of letters but in their conversion to phonemes after they have been fully recognized and passed on to the orthographic lexicon. Interpreted this way, an effect of pseudoword length associated with spatial attention for multi-letter processing is incompatible with the DRC model.

However, an alternative interpretation is possible in the context of the DRC model: Once recognized, letters are presented sequentially to GPC system. Therefore, a sequential operation is implicated, to allocate attention to individual letters, which are then subjected to conversion. This operation can be theoretically distinguished from the core rule-applying operation of the GPC system. If the observed activation in our parietal region of interest reflects this serial allocation of attention, then our finding is compatible with the DRC model and with the proposal of [Taylor et al. \(2013\)](#). The localization of the GPC rule application core must await future investigations.

Middle temporal gyrus

Examining ROIs formed around individual participant peaks of the word vs. pseudoword contrast revealed a marginally significant effect of frequency at the middle temporal gyrus, near the cluster identified by [Taylor et al. \(2013\)](#) as a candidate region for the phonological lexicon or the semantic system. However, the word frequency effect was positive, that is, higher signal was associated with more frequent words. This

seems to go against the predictions of the engagement-effort framework proposed by Taylor et al (2013), which suggested that all words should engage the phonological lexicon, but that low frequency words should be more effortful to process, and thus should result in increased activity, relative to high frequency words.

The posterior middle temporal cluster formed part of an extensive temporoparietal area that was consistently and significantly deactivated, relative to fixation, for both word and pseudoword stimuli. The region was more deactivated for pseudowords than words and for low than high frequency words (cf. Figure 2). This pattern of deactivation at and around the angular gyrus has been previously reported (e.g., Binder et al., 2005; Mechelli et al., 2003; Taylor et al., 2014) and has been interpreted as consistent with the idea that this network is involved in semantic processing. Specifically, semantic activation is said to predominate during the interstimulus periods, which are likely “highly active ... involving ongoing retrieval of conceptual and autobiographical knowledge, problem-solving, and planning” (Binder et al., 2005, p. 686; see also Binder & Desai, 2011; Binder, Desai, Graves, & Conant, 2009). This activity is interrupted by presentation of a stimulus. If the stimulus is a single word, the associated semantic activation is relatively shallow in comparison to the preceding rumination, hence the pattern of deactivation for words vs. fixation. If the stimulus is a pseudoword, there is no associated semantic activation, hence greater deactivation. Thus, the observed deactivation is a measure of the extent of detraction from ongoing semantic activity that is caused by processing of the stimulus.

One could also argue that such a pattern of deactivation is compatible with the function of a phonological lexicon, as described by Taylor et al. (2013), by hypothesizing that rumination during fixation, being linguistic in nature, involves retrieval of multiple spoken word representations, and thus activation of the phonological lexicon. This

could lead to deactivation during single word processing due to the reduction in linguistic content. Extending this interpretation to encompass effects of word frequency in terms of the engagement-effort framework, one might note that more frequent words would require less processing than less frequent words. Therefore, more frequent words would detract less from ongoing rumination, causing less deactivation, relative to the fixation baseline, than less frequent words. This pattern of differential deactivation amounts to a positive effect of word frequency, consistent with our observation. The relative weakness of the frequency effect in this region (only marginally significant in the ROI analysis) could be attributed to the frequency measure used, which is based on printed text corpora. Hence it reflects printed rather than spoken word frequency, and is not as well aligned with phonological lexicon activation.

Fusiform gyrus

In ROIs formed around individual participant peaks from the pseudoword vs. fixation contrast in the left fusiform gyrus we observed a negative effect of word frequency but no effect of word length. In addition, this region was not differentially active for words vs. pseudowords and our findings are therefore consistent with the role of the left fusiform gyrus as a distributed orthographic processor rather than as the locus for an orthographic lexicon. Our study joins previous reports failing to reveal evidence for a word-specific area of orthographic processing (e.g., [Binder et al., 2005](#); [Jobard et al., 2003](#); cf. [Fiebach, Ricker, Friederici, & Jacobs, 2007](#), for a nonlexical account of orthographic neighborhood size effects implicating executive control functions).

The visual word form area has long been known to be involved in orthographic processing of both word and nonword letter strings as well as other types of visual stimuli ([Price, 2012](#)). It has emerged in previous studies as significantly more active for

pseudoword than word stimuli (Binder et al., 2005; but cf. Fiebach et al., 2002, 2007), for pseudohomophones than for words (Schurz et al., 2014), for less frequent words (Graves et al., 2010; Hauk et al., 2008; Kronbichler et al., 2004; Yarkoni et al., 2008), for more familiar pseudoword strings (compared to less familiar strings; Binder, Medler, Westbury, Liebenthal, & Buchanan, 2006), or for case-deviant and letter-deviant stimuli (compared to non-deviant; Kronbichler et al., 2009).

The lack of a difference between words and pseudowords in this region may be attributed to some extent to the transparency of the orthography (cf. Cherodath & Singh, 2015). An additional and perhaps more important reason is that the pseudowords were well matched to the words in orthographic and phonological properties. Other studies have also attempted to match word and pseudoword stimuli, but not as extensively as in our study and usually taking into account only a few variables, including bigram frequencies, which have limited if any effects on visual word recognition latencies (Adelman, 2012; Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004), and excluding syllable frequencies, which have emerged as a significant variable in several languages (discussed below). In addition to matching on several variables, we have allowed the pseudowords to have lower orthographic and phonological neighborhood sizes than the words, in order to minimize lexical activation from pseudowords and thereby to facilitate detection of the orthographic lexicon. This strengthens our conclusion that the lack of evidence for an orthographic lexicon cannot be ascribed to superficial aspects of the study, such as material selection.

Price and Devlin (2011; see also Dehaene & Cohen, 2011) have proposed that the ventral occipitotemporal cortex functions as an interactive integration hub for visuospatial, speech sound, action, and meaning information. Under this view, orthographic processing, at least further down the processing hierarchy, may be

phonologically structured in the sense that graphemes mapping onto the same phonemes are orthographically equivalent. This account seems sensible from a language-oriented receptive-field point of view, whereby initially visual representations proceed through the fusiform gyrus gradually turning into graphemic representations abstracting away from visual details and increasingly conforming to categorical equivalence conditions subsuming size, case, font, and, eventually, alternative combinations mapping onto the same phoneme. If high-level orthographic representations already reflect phonological structure ([Pylkkänen & Okano, 2010](#)) then no GPC system or structured spelling–sound mappings would be necessary. Phonologically structured visual representations could simply activate their corresponding speech-sound representations.

Other regions sensitive to variables investigated in the current study

In the inferior frontal gyrus pars opercularis and pars triangularis we observed a negative word frequency effect (higher activation for less frequent words) in addition to a positive effect of syllable frequency (higher activation for words with more frequent syllables). Negative effects of word frequency in the left inferior frontal gyrus have consistently been reported in previous studies across several languages such as English, German, and French ([Carreiras et al. 2006; Fiebach et al., 2002; Graves et al., 2010;](#) [Hauk et al. 2008; Joubert et al. 2004; Kronbichler et al. 2004; Yarkoni et al. 2008](#)).

Syllable frequency effects have rarely been investigated in neuroimaging studies. One study failed to find such effects within the reading network but it used a block design in which no behavioral effect of syllable frequency was observed ([Carreiras, Mechelli, & Price, 2006](#)). In contrast, as in our study, slower behavioral responses to words with high frequency phonological syllables have been observed across languages ([Conrad & Jacobs, 2004; Conrad, Grainger, & Jacobs, 2007](#); see review in [Conrad, Tamm, Carreiras,](#)

& Jacobs, 2010) and have been shown to be independent of bigram frequency (Conrad, Carreiras, Tamm, & Jacobs, 2009).

One possible interpretation of our findings is that the left inferior frontal gyrus corresponds to a phonological output lexicon: Within such a system, resolving a phonological output representation should be more effortful for low than high frequency words, and lexical competition should be greater for words with high frequency syllables (Perea & Carreiras, 1998). The argument that the left inferior frontal gyrus subserves a lexical function is also consistent with the fact that we did not observe a syllable frequency effect for pseudowords. However, we did not observe greater BOLD responses to words than pseudowords in this region, which should be the case in the phonological lexicon. Thus, a conservative explanation is that this region subserves some function relating to phonological output, in line with Jacquemot and Scott (2006) who proposed that the left inferior frontal gyrus is involved in phonological output, whereas the left supramarginal gyrus and posterior superior temporal gyrus process incoming phonological information. More research is necessary to determine how these regions process and represent phonological forms.

Summary of results

The current study improved on previous research by using large and representative ranges of all variables examined (Table 1; see also Appendix B in Protopapas & Kapnoula, in press) and testing more participants. To guard against the possibility that the effects of the variables under study are nonlinear with respect to response latency we constructed the stimulus set based on nonparametric indices of correlation and documented the empirical shape of variable effects on response time (Protopapas & Kapnoula, 2013), since it should not be assumed that effects can be statistically “removed” by linear regression. In combination with a stringent data

quality procedure, we believe our imaging results are reasonably reliable.

Using these methods we tested predictions arising from the DRC model of reading aloud; specifically, that word frequency effects should co-localize with significant clusters from the word vs. pseudoword contrast and that pseudoword length effects should co-localize with significant clusters from the pseudoword vs. fixation contrast. Results from the whole-brain analysis failed to support these hypotheses. However, looking into regions of interest defined individually for each participant, a different picture emerged that can be interpreted as partially consistent with the DRC model. Specifically, in dorsal parietal cortex, pseudoword length effects were obtained as hypothesized for a region corresponding to the GPC system. Accompanying word frequency effects were interpreted as arising from a backward abort signal from the lexical system upon successful completion. In the middle temporal gyrus a marginally significant effect of frequency was obtained, consistent with involvement in semantic or phonological processing.

However, further research is necessary before we can be confident in these interpretations. It is unclear whether activation in dorsal parietal cortex primarily reflects visual attention or internal serial allocation of attention, and whether the serial rule application of the GPC system can be associated with this region. As for the deactivation observed in the left middle temporal gyrus for reading relative to fixation, and the associated reverse frequency effect, it remains to be fully understood. In addition, there was no evidence consistent with the existence of an orthographic lexicon in the fusiform gyrus. Overall, it seems that the results of the more sensitive ROI analysis constitute, at best, weak evidence for the DRC model components.

Evaluating cognitive models with neuroimaging

That the DRC permits such a well-specified evaluation is a strength of the model.

In comparison, the distributed representations in the connectionist network of the triangle model do not seem to support the derivation of similarly specific predictions beyond a localization of major model components (cf. Cox, Seidenberg, & Rogers, in press). Because activation always flows through all layers in a recurrent connectionist network, activating nodes in every layer, regardless of the type of input, and because all processes and representations in the model are of the same type, it is unclear whether specific functional hypotheses based on the triangle model can be disconfirmed by mass-univariate BOLD signal contrasts. In this respect our study cannot be construed as a fair comparison between two models. In particular, because no specific predictions from the connectionist triangle model were tested, our study cannot be construed as evidence in support of this model even if it is possible to identify correspondences between model components and regions activated in our contrasts.

A frequent criticism against the application of neuroimaging findings toward cognitive theory (and models) is that theories are meant to be functional. That is, we should not expect to be able to localize components of cognitive models to particular brain regions, because cognitive operations may be performed by distributed networks in ways that do not result in BOLD signal differences detectable in fMRI. Furthermore, regions revealed to be “activated” by specific contrasts may be related to generic or supplementary aspects of performance and not to the hypothesized cognitive operations (cf. [Binder et al., 2005](#); but see [Taylor et al. 2014](#) for an opposing view). In addition, the usual model of the hemodynamic response function (the “canonical” double-gamma hrf) used in most fMRI studies does not constitute an equally good fit to the empirical hrf in every brain region ([Gonzalez-Castillo et al., 2012](#)). Thus, in general, it may not be possible to disconfirm a cognitive model on the basis of neuroimaging localizations ([Uttal, 2001](#); cf. [Coltheart, 2006, 2010](#)). However, the prospects may not

always be so dim. In the present study, specific predictions were based on unquestionable properties of the DRC model. We can reasonably expect model components to localize to brain regions because the model is not an entirely functional and ungrounded hypothesis but, like the triangle model, it was originally created to address findings from neuropsychological patients presenting with dissociable deficits following damage to different brain regions.

Finally, it should not pass without comment that acceptance of the individual ROI findings in the face of the absence of such effects in the whole-brain analysis would effectively undermine the entire research paradigm, as most of the findings in the literature have been based on whole-brain analyses. If our study, with a participant sample more than twice the usual size, has failed to reveal significant effects in the whole-brain comparisons, yet such effects are discernible in the individual ROI analyses, this means that the results reported by lower-powered studies are simply not reliable enough to lead to any robust theoretical conclusions. In other words, individual variability in precise localization of activation corresponding to some purported cognitive operations may simply be too high to allow detection by standard fMRI analysis procedures (cf. Nieto-Castañon & Fedorenko, 2012). Since meta-analyses typically use reported significant group-level peaks, this means that a host of potential effects are unlikely to be detected even in large and comprehensive reviews. Before embracing the devastating implications of this suggestion, it may be constructive to conduct further intensive studies of individual variability in the localization of contrast peaks with theoretical import for cognitive models. Although general skepticism may not be warranted just yet, our results suggest that whole-brain analysis perhaps cannot reveal all the detail that is necessary for model evaluation.

In conclusion, our study contributes data from a less studied language that are

consistent with our understanding of the components of the reading network as previously documented and interpreted (Price, 2012). In particular, due to our well-controlled stimulus set with decorrelated parameters, our findings on the parametric effects of frequency, length, and syllable frequency, may contribute toward future elaborations of the cognitive mechanisms of reading as they are implemented in the brain across languages.

Acknowledgements

We are grateful to Paul Mazaika for the Alphascript preprocessing code and help; to Fotis Fotiadis for help setting up the experimental procedures; to the personnel of the Radiology and Medical Imaging Research Unit for facilitating and supporting our research; and to Sofia Loui for help in testing participants.

References

- Adelman, J. S. (2012). Methodological issues with words. In J. S. Adelman (Ed.), *Visual word recognition* (Vol. 1, pp. 116–138). Hove, UK: Psychology Press.
- Adelman, J. S., Marquis, S. J., & Sabatos-DeVito, M. G. (2010). Letters in words are read simultaneously, not in left-to-right sequence. *Psychological Science*, 21, 1799–1801.
- Baayen, R. H. (2008). *Analyzing linguistic data*. Cambridge University Press.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412.
- Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, 3, 12–28.
- Balota, D., Cortese, M., Sergent-Marshall, S., Spieler, D., & Yap, M. (2004). Visual word recognition for single-syllable words. *Journal of Experimental Psychology: General*, 133, 283–316.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4. R package v. 1.0-5. <http://CRAN.R-project.org/package=lme4>.
- Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, 15, 527–536.
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic

system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19, 2767–2796.

Binder, J. R., Medler, D. A., Desai, R., Conant, L. L., & Liebenthal, E. (2005). Some neurophysiological constraints on models of word naming. *NeuroImage*, 27, 677–693.

Binder, J. R., Medler, D. A., Westbury, C. F., Liebenthal, E., & Buchanan, L. (2006). Tuning of the human left fusiform gyrus to sublexical orthographic structure. *NeuroImage*, 33, 739–748.

Boukrina, O., & Graves, W. W. (2013). Neural networks underlying contributions from semantics in reading aloud. *Frontiers in Human Neuroscience*, 7, 518.

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). Region of interest analysis using an SPM toolbox. Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2–6, Sendai, Japan.

Carreiras, M., Armstrong, B. C., Perea, M., & Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences*, 18, 90–98.

Carreiras, M., Mechelli, A., Estévez, A., & Price, C. J. (2007). Brain activation for lexical decision and reading aloud: two sides of the same coin? *Journal of Cognitive Neuroscience*, 19, 433–444.

Carreiras, M., Mechelli, A., & Price, C. J. (2006). Effect of word and syllable frequency on activation during lexical decision and reading aloud. *Human Brain Mapping*, 27, 963–972.

Chang, Y. N., Furber, S., & Welbourne, S. (2012). “Serial” effects in parallel models of reading. *Cognitive Psychology*, 64, 267–291.

Cherodath, S., & Singh, N. C. (2015). The influence of orthographic depth on reading networks in simultaneous biliterate children. *Brain and Language*, 143, 42–51.

- Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. *NeuroImage*, 40, 353–366.
- Coltheart, M. (2006). What has functional neuroimaging told us about the mind (so far)? *Cortex*, 42, 323–331.
- Coltheart, M. (2010). What is functional neuroimaging for? In S. J. Hanson & M. Bunzl (Eds.), *Foundational issues in human brain mapping* (pp. 263–272). Cambridge, MA: MIT Press.
- Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, 100, 589–608.
- Coltheart, M., Davelaar, E., Jonasson, T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention & performance IV* (pp. 535–555). Hillsdale, NJ: Erlbaum.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204–256.
- Conrad, M., Carreiras, M., & Jacobs, A. M. (2008). Contrasting effects of token and type syllable frequency in lexical decision. *Language and Cognitive Processes*, 23, 296–326.
- Conrad, M., Carreiras, M., Tamm, S., & Jacobs, A. M. (2009). Syllables and bigrams: orthographic redundancy and syllabic units affect visual word recognition at different processing levels. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 461–479.
- Conrad, M., Grainger, J., & Jacobs, A. M. (2007). Phonology as the source of syllable

frequency effects in visual word recognition: Evidence from French. *Memory & Cognition*, 35, 974–983.

Conrad, M., & Jacobs, A. (2004). Replicating syllable frequency effects in Spanish in German: One more challenge to computational models of visual word recognition. *Language and Cognitive Processes*, 19, 369–390.

Conrad, M., Tamm, S., Carreiras, M., & Jacobs, A. M. (2010). Simulating syllable frequency effects within an interactive activation framework. *European Journal of Cognitive Psychology*, 22, 861–893.

Cox, C. R., Seidenberg, M. S., & Rogers, T. T. (in press). Connecting functional brain imaging and parallel distributed processing. *Language, Cognition and Neuroscience*. doi:10.1080/23273798.2014.994010

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, 15, 254–262.

Fiebach, C., Friederici, A., Müller, K., & von Cramon, D. V. (2002). fMRI evidence for dual routes to the mental lexicon in visual word recognition. *Journal of Cognitive Neuroscience*, 14, 11–23.

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

Forget, J., Buiatti, M., & Dehaene, S. (2010). Temporal integration in visual word recognition. *Journal of Cognitive Neuroscience*, 22, 1054–1068.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments & Computers*, 35, 116–124.

Fox, P. T., Laird, A. R., Fox, S. P., Fox, P. M., Uecker, A. M., Crank, M., Koenig, S. F., &

- Lancaster, J. L. (2005). BrainMap taxonomy of experimental design: Description and evaluation. *Human Brain Mapping*, 25, 185–198.
- Gonzalez-Castillo, J., Saad, Z. S., Handwerker, D. A., Inati, S. J., Brenowitz, N., & Bandettini, P. A. (2012). Whole-brain, time-locked activation with simple tasks revealed using massive averaging and model-free analysis. *Proceedings of the National Academy of Sciences*, 109, 5487–5492.
- Gottlieb, J. (2007). From thought to action: The parietal cortex as a bridge between perception, action, and cognition. *Neuron*, 53, 9–16.
- Graves, W. W., Binder, J. R., Desai, R. H., Humphries, C., Stengel, B. C., & Seidenberg, M. S. (2014). Anatomy is strategy: skilled reading differences associated with structural connectivity differences in the reading network. *Brain and Language*, 133, 1–13.
- Graves, W. W., Desai, R., Humphries, C., Seidenberg, M. S., & Binder, J. R. (2010). Neural systems for reading aloud: A multiparametric approach. *Cerebral Cortex*, 20, 1799–1815.
- Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: insights from connectionist models. *Psychological Review*, 106, 491–528.
- Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, 111, 662–720.
- Harwell, J., Drury, H. A., Hanlon, D., & Van Essen, D. C. (2012). Computerized anatomical reconstruction and editing toolkit (v. 5.65). Washington University School of Medicine, St Louis, MO. <http://brainvis.wustl.edu/wiki/index.php/Caret:About>
- Hauk, O., Davis, M. H., & 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.
- Jacquemot, C., & Scott, S. K. (2006). What is the relationship between phonological short-term memory and speech processing? *Trends in Cognitive Sciences*, 10, 480–486.
- Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: a metaanalysis of 35 neuroimaging studies. *NeuroImage*, 20, 693–712.
- Joubert, S., Beauregard, M., Walter, N., Bourgouin, P., Beaudoin, G., Leroux, J. M., ... & Lecours, A. R. (2004). Neural correlates of lexical and sublexical processes in reading. *Brain and Language*, 89, 9–20.
- Keuleers, E. (2013). vwr: Useful functions for visual word recognition research. R package version 0.3.0. <http://CRAN.R-project.org/package=vwr>
- Keuleers, E., & Brysbaert, M. (2011). Detecting inherent bias in lexical decision experiments with the LD1NN algorithm. *Mental Lexicon*, 6, 34–52.
- Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience*, 12, 535–540.
- 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. *Journal of Cognitive Neuroscience*, 21, 222–229.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2013). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 2.0-3. <http://CRAN.R-project.org/package=lmerTest>

Lacadie, C. M., Fulbright, R. K., Rajeevan, N., Constable, R. T., & Papademetris, X. (2008).

More accurate Talairach coordinates for neuroimaging using non-linear registration. *NeuroImage*, 42, 717–725.

Laird, A. R., Lancaster, J. L., & Fox, P. T. (2005). BrainMap: The social evolution of a functional neuroimaging database. *Neuroinformatics* 3, 65–78.

Levy, J., Pernet, C., Treserras, S., Boulanouar, K., Aubry, F., Démonet, J. F., & Celsis, P. (2009). Testing for the dual-route cascade reading model in the brain: An fMRI effective connectivity account of an efficient reading style. *PLoS One*, 4, e6675.

Levy, J., Pernet, C., Treserras, S., Boulanouar, K., Berry, I., Aubry, F., ... & Celsis, P. (2008). Piecemeal recruitment of left-lateralized brain areas during reading: A spatio-functional account. *NeuroImage*, 43, 581–591.

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. *Human Brain Mapping*, 36, 1963–1981.

Marzelli, M., Hoeft, F., Mazaika, P., & Sheau, K. (2009) Alphascript for fMRI preprocessing at CIBSR. Unpublished manuscript.

Mazaika, P., Hoeft, F., Glover, G. H., & Reiss, A. L. (2009). Methods and software for fMRI analysis for clinical subjects. Presented at the 15th Annual Meeting of the Organization for Human Brain Mapping, San Francisco, CA, June 18–23.

Mazaika, P., Whitfield-Gabrieli, S., & Reiss, A. L. (2007). Artifact repair for fMRI data from high motion clinical subjects. Presented at the 13th Annual Meeting of the Organization for Human Brain Mapping, Chicago, IL, June 10–14.

Mechelli, A., Gorno-Tempini, M., & Price, C. (2003). Neuroimaging studies of word and pseudoword reading: Consistencies, inconsistencies, and limitations. *Journal of Cognitive Neuroscience*, 15, 260–271.

- Mechelli, A., Crinion, J., Long, S., Friston, K., Ralph, M., Patterson, K., ... & Price, C. (2005). Dissociating reading processes on the basis of neuronal interactions. *Journal of Cognitive Neuroscience*, 17, 1753–1765.
- Mei, L., Xue, G., Lu, Z.-L., He, Q., Zhang, M., Wei, M., ... & Dong, Q. (2014) Artificial language training reveals the neural substrates underlying addressed and assembled phonologies. *PLoS ONE*, 9, e93548.
- Nieto-Castañón, A., & Fedorenko, E. (2012). Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. *NeuroImage*, 63, 1646–1669.
- Paulesu, E., Danelli, L., & Berlingeri, M. (2014). Reading the dyslexic brain: Multiple dysfunctional routes revealed by a new meta-analysis of PET and fMRI activation studies. *Frontiers in Human Neuroscience*, 8, 830.
- Perea, M., & Carreiras, M. (1998). Effects of syllable frequency and syllable neighborhood frequency in visual word recognition. *Journal of Experimental Psychology Human Perception and Performance*, 24, 134–144.
- Plaut, D. C., McClelland, J. L., Seidenberg, M., & Patterson, K. E. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103, 56–115.
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, 10, 59–63.
- Price, C. J. (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, C. J., & Devlin, J. T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, 15, 246–253.
- Protopapas, A., Fakou, A., Drakopoulou, S., Skaloumbakas, C., & Mouzaki, A. (2013). What

do spelling errors tell us? Classification and analysis of spelling errors of Greek schoolchildren with and without dyslexia. *Reading & Writing: An Interdisciplinary Journal*, 26, 615–646.

Protopapas, A., Tzakosta, M., Chalamandaris, A., & Tsakoulis, P. (2012). IPLR: An online resource for Greek word-level and sublexical information. *Language Resources & Evaluation*, 46, 449–459.

Protopapas, A., & Kapnoula, E. C. (in press). Short-term and long-term effects on visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. doi: 10.1037/xlm0000191

Protopapas, A., & Kapnoula, E. C. (2013). Exploring word recognition with selected stimuli: The case for decorrelated parameters. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 1169–1174). Austin, TX: Cognitive Science Society.

Protopapas, A., & Vlahou, E. L. (2009). A comparative quantitative analysis of Greek orthographic transparency. *Behavior Research Methods*, 41, 991–1008.

Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., ... & Shaywitz, B. A. (2000). Functional neuroimaging studies of reading and reading disability (developmental dyslexia). *Mental Retardation and Developmental Disabilities Research Reviews*, 6, 207–213.

Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., ... & Shaywitz, B. A. (2001). Neurobiological studies of reading and reading disability. *Journal of Communication Disorders*, 34, 479–492.

Pylkkänen, L., & Okano, K. (2010). The nature of abstract orthographic codes: evidence from masked priming and magnetoencephalography. *PloS one*, 5, e10793.

R Core Team (2013). R: A language and environment for statistical computing. R

Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>

Reilhac, C., Peyrin, C., Démonet, J. F., & Valdois, S. (2013). Role of the superior parietal lobules in letter-identity processing within strings: FMRI evidence from skilled and dyslexic readers. *Neuropsychologia*, 51, 601–612.

Richardson, F. M., Seghier, M. L., Leff, A. P., Thomas, M. S., & Price, C. J. (2011). Multiple routes from occipital to temporal cortices during reading. *The Journal of Neuroscience*, 31, 8239–8247.

Roux, F. E., Durand, J. B., Jucla, M., Réhault, E., Reddy, M., & Démonet, J. F. (2012). Segregation of lexical and sub-lexical reading processes in the left perisylvian cortex. *PloS one*, 7, e50665.

Scholz, F., & Zhu, A. (2015). kSamples: K-sample rank tests and their combinations. R package version 1.0.1. <http://CRAN.R-project.org/package=kSamples>

Schurz, M., Kronbichler, M., Crone, J., Richlan, F., Klackl, J., & Wimmer, H. (2014). Top-down and bottom-up influences on the left ventral occipito-temporal cortex during visual word recognition: An analysis of effective connectivity. *Human Brain Mapping*, 35, 1668–1680.

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.

Seghier, M. L., Lee, H. L., Schofield, T., Ellis, C. L., & Price, C. J. (2008). Inter-subject variability in the use of two different neuronal networks for reading aloud familiar words. *NeuroImage*, 42, 1226–1236.

Simos, P. G., Rezaie, R., Fletcher, J. M., & Papanicolaou, A. C. (2013). Time-constrained

functional connectivity analysis of cortical networks underlying phonological decoding in typically developing school-aged children: A magnetoencephalography study. *Brain and Language*, 125, 156–164.

Spencer, K. A. (2009). Feedforward, -backward, and neutral transparency measures for British English. *Behavior Research Methods*, 41, 220–227.

Taylor, J. S. H., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin*, 139, 766–791.

Taylor, J. S. H., Rastle, K., & Davis, M. H. (2014). Interpreting response time effects in functional imaging studies. *NeuroImage*, 99, 419–433.

Thyreau, B., Schwartz, Y., Thirion, B., Frouin, V., Loth, E., Vollstädt-Klein, S., ... & IMAGEN Consortium. (2012). Very large fMRI study using the IMAGEN database: Sensitivity–specificity and population effect modeling in relation to the underlying anatomy. *NeuroImage*, 61, 295–303.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., & Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, 15, 273–289.

Uttal, W. R. (2001). *The new phrenology: The limits of localizing cognitive processes in the brain*. Cambridge, MA: MIT Press.

Van Essen, D. C. (2005) A population-average, landmark- and surface-based (PALS) atlas of human cerebral cortex. *Neuroimage*, 28, 635–662

Van Essen, D. C., Dickson, J., Harwell, J., Hanlon, D., Anderson, C. H. & Drury, H. A. (2001). An integrated software system for surface-based analyses of cerebral cortex. *Journal of the American Medical Informatics Association*, 8, 443–459.

- Van Essen, D. C., & Dierker, D. (2007) Surface-based and probabilistic atlases of primate cerebral cortex. *Neuron*, 56, 209–225.
- 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.
- Woollams, A. M., Silani, G., Okada, K., Patterson, K., & Price, C. J. (2011). Word or word-like? Dissociating orthographic typicality from lexicality in the left occipito-temporal cortex. *Journal of Cognitive Neuroscience*, 23, 992–1002.

Table 1

Descriptive statistics for words, pseudowords, and corpus types and tokens, for each variable, and results from the Anderson-Darling nonparametric test comparing word and pseudoword distributions

	Words			Pseudowords		A-D test ^b		Corpus Types		Corpus Tokens	
	M	SD	Range	M	SD	AD	p	M	SD	M	SD
Log frequency ^a	0.88	1.89	-3.38-4.22					-1.49	1.81	6.22	3.27
Number of letters ^a	7.24	1.93	4-10	7.31	1.84	0.289	.949	10.07	2.71	5.43	3.17
Number of phonemes	7.09	1.93	4-11	7.00	1.82	0.434	.817	9.45	2.61	5.02	3.03
Number of syllables	3.11	0.95	2-5	3.10	0.92	0.158	.998	4.35	1.29	2.38	1.45
Orth. neighbors ^a	2.17	1.50	0-7	0.39	1.21	96.545	< .001	1.38	1.58	5.88	4.83
Phon. neighbors	3.25	1.88	0-10	1.31	3.31	73.883	< .001	2.36	3.27	10.72	8.42
Orth. bigram frequency ^a	0.76	0.26	0.22-1.31	0.79	0.32	1.057	.328	1.02	0.42	1.91	1.46
Phon. bigram frequency	1.03	0.37	0.23-2.03	0.98	0.44	1.844	.111	1.23	0.59	2.10	1.63
Orth. syllable frequency	7.86	5.70	0.85-20.1	6.36	4.73	2.635	.042	8.93	6.08	11.64	8.23
Phon. syllable frequency ^a	11.07	5.78	0.92-21.7	9.82	5.40	2.220	.069	12.61	6.66	15.89	9.40
G-P consistency ^a	32.67	8.03	15.2-49.5	32.75	7.68	0.388	.863	32.13	9.41	38.84	15.66

Note: Orth=Orthographic; Phon=Phonological; G-P=Graphophonemic; A-D=Anderson-Darling

^a Variable used in parametric modulation analysis; ^b Version 1 of the Anderson-Darling test, with associated asymptotic probability

Table 2

Nonparametric correlation coefficients (Spearman's ρ) between variables for words (above the diagonal) and pseudowords (below the diagonal)

	2	3	4	5	6	7	8	9	10	11
1 Log frequency ^a	-.049	-.084	-.084	-.002	.069	.104	.091	-.081	-.024	.002
2 Number of letters ^a		.965*	.860*	-.007	.039	.008	.034	-.080	-.045	-.106
3 Number of phonemes	.947*		.872*	.057	.023	-.067	.089	-.096	-.100	-.057
4 Number of syllables	.870*	.904*		.044	.086	-.096	.042	.017	.082	-.050
5 Orth. neighbors ^a	-.504*	-.505*	-.486*		.640*	.071	.084	-.043	-.081	-.018
6 Phon. neighbors	-.612*	-.659*	-.649*	.685*		.133	.095	-.046	-.017	-.052
7 Orth. bigram frequency ^a	.083	-.060	-.056	-.037	.106		.326*	.060	.086	.056
8 Phon. bigram frequency	-.055	.039	.000	.100	.177*	.214*		-.253*	-.068	.110
9 Orth. syllable frequency	-.065	-.010	.124	.096	.019	.092	.017		.786*	.064
10 Phon. syllable frequency ^a	-.077	-.081	.106	.065	.056	.065	.101	.697*		-.093
11 G-P consistency ^a	.003	.122	.036	-.032	-.121	.046	.122	.044	-.123	

Note: Orth=Orthographic; Phon=Phonological; G-P=Graphophonemic

^aVariable used in parametric modulation analysis; * $p < .05$

Table 3

Speed and accuracy of responding to the stimuli

Stimulus type	Response time (ms)		Proportion correct	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Words	985.8	145.7	95.1	3.8
Pseudowords	1141.7	236.5	97.5	3.6
Fixation	725.7	93.0	99.6	0.8

Table 4

Whole-brain inference information

Contrast	Resels	FWEp	Smoothness		
			x	y	z
Pseudowords vs. fixation	337.8	5.13	14.5	14.6	14.3
Words vs. pseudowords	293.7	5.08	15.4	15.4	14.7
Number of letters (pseudo)	245.6	5.02	16.3	16.4	15.6
Word frequency	268.8	5.05	15.8	16.0	15.2
Words vs. fixation	316.3	5.11	14.8	15.0	14.6
Number of letters (words)	317.8	5.11	15.0	15.2	14.1
Syllable frequency (words)	280.9	5.07	15.6	15.7	14.8

Note. FWEp = familywise error-corrected ($p < 0.05$) threshold for T; Smoothness, in mm FWHM (full-width half-maximum).

Table 5

Clusters of at least 4 voxels with T exceeding 5.1, and associated peak coordinates

cluster size	peak T	x	y	z	Anatomical labels
<i>Pseudowords > fixation</i>					
1067	17.64	-18	-94	-11	L IO, L MO, L fus,
	12.90	-30	-82	-14	L calc, L ling, L SP,
	12.68	-45	-70	-14	L SO, L IT, L IP, L cer
697	14.23	21	-91	-2	R calc, R MO, R IO,
	8.28	39	-82	4	R SO, R cer, R cun
	8.24	39	-82	-5	
1095	13.97	-45	8	22	L IFop/tri/orb,
	13.86	-51	-1	40	L preC, L ins,
	11.67	-30	26	1	L STpole, L postC,
					L RolOp
409	12.70	-3	14	49	L/R SMA,
	12.40	-3	5	58	L/R MCing, L/R mSF
175	10.37	36	23	-2	R ins, R IForb/tri/op
10	6.28	-18	8	4	L put, L pall
5	6.25	-24	-4	49	L MF, L SF
<i>Fixation > pseudowords</i>					
1490	12.93	48	-55	25	R ang, R MO, R MT,
	11.12	42	-67	34	R supMar, R ST,
	10.62	48	-70	25	R RolOp, R postC,
					R IP, R IT
1378	12.57	-3	-67	40	L/R precun, R cun,
	10.80	6	-58	37	L/R MCing,
	10.45	6	-61	28	L/R pCing, L/R calc
668	11.98	-42	-67	40	L ang, L MT, L MO,
	11.63	-48	-58	19	L supMar, L IP
	9.55	-42	-76	28	
202	9.15	33	26	40	R MF, R SF
267	8.18	-24	23	49	L MF, L SF
425	7.61	3	38	-8	R medSF, R/L medF
	7.58	12	59	4	orb, L aCing, R aCing,
	6.94	12	50	-2	R SF
120	6.83	42	-7	-2	R ins, R ST, R RolOp
	5.64	54	-10	7	R H
54	6.42	-45	-10	1	L ins, L ST, L RolOp
64	6.33	-36	-25	52	L postC, L preC
	5.96	-48	-22	52	
17	6.14	30	-40	-14	R fus
16	6.08	-48	-25	16	L RolOp
7	5.88	-30	-40	-14	L fus
4	5.79	54	-4	-23	R MT
<i>Words > pseudowords</i>					
993	11.41	-54	-43	-11	L MTG, L supMar,
	11.07	-51	-49	37	

	8.64	-57	-49	28	L IP, L ang, L MO, L IT, L ST
724	8.74	51	-43	49	R IP, R supMar, R ST,
	8.44	60	-40	28	R ang, R MT,
	7.23	63	-19	25	R postC, R SP
493	8.33	-6	50	34	L/R mSF, L MF, L SF,
	6.88	-30	14	46	L aCing
	6.78	-30	26	46	
120	7.84	57	-37	-11	R MT, R IT
	5.95	48	-31	-8	
	5.78	60	-55	-2	
118	7.47	-51	-22	49	L postC, L IP
	6.86	-45	-22	43	
	5.66	-42	-34	55	
169	7.36	-6	11	-5	L/R caud
	7.03	9	11	-5	
	5.86	12	11	7	
29	6.83	-39	2	-2	L ins
35	6.24	-30	20	-14	L ins, L IForb
10	5.81	-6	-97	1	L calc
	5.80	0	-94	7	
4	5.71	-21	-49	64	L SP
8	5.67	48	-67	22	R MT, R MO
18	5.64	-3	-34	31	L pCing, L MCing
19	5.52	0	-88	-8	L calc
6	5.50	36	17	-14	R ins
<i>Pseudowords > words</i>					
22	7.40	-51	-1	43	L preC, L postC
4	5.70	-3	2	58	L SMA

<i>N letters (pseudowords), positive effect</i>					
233	12.25	-12	-88	-8	L ling, L calc, L fus,
	8.09	-24	-79	-14	L IO, L MO, L cer
	10.94	15	-85	-5	R ling, R calc, R SO, R cun, R fus, R cer, R MO
38	8.22	-51	-4	43	L preC, L postC
<i>N letters (pseudowords), negative effect</i>					
4	5.34	54	-55	-2	R MT, R IT

<i>Word frequency, negative effect</i>					
24	6.43	33	23	-8	R IForb, R ins
67	6.37	-39	29	-2	L IFtri, L IForb
	6.04	-36	23	-8	
33	5.92	-51	14	4	L IFop, L IFtri
	5.74	-42	8	22	
33	5.88	0	17	49	L/R SMA

Note. L = left, R = right, I = inferior, M = middle, S = superior, F = frontal, P = parietal, T = temporal, O = occipital, C = central, m = medial, a = anterior, p = posterior, ang = angular, calc = calcarine, caud = caudate, cer = cerebellum, cing = cingulum, cun = cuneus, fus = fusiform, hip = hippocampus, ins = insula, ling = lingual, pall = pallidum, precun = precuneus, put = putamen, SupMar = supramarginal, RolOp = Rolandic operculum, orb = pars orbitalis, tri = pars triangularis, op = pars opercularis, SMA = supplementary motor area.

Table 6

Clusters of at least 4 voxels with T exceeding 5.1, and associated peak coordinates, for additional contrasts involving words.

cluster size	peak T	x	y	z	
<i>Words > fixation</i>					
2365	18.42	-18	-91	-11	L/R MO, L/R IO, L/R
	14.11	18	-88	-8	SO, L/R calc, L/R fus,
					L ling, L SP, L IP, L IT,
	12.97	-45	-67	-14	L/R cereb, L/R cun
1455	14.28	-45	11	22	L IFop/tri/orb, L ins,
	12.23	-45	2	34	L STpole, L postC, L
	12.14	-33	23	-2	RolOp, L MF, L SF
352	11.31	36	20	-2	R IFop/tri/orb, R MF
477	10.86	-3	17	46	L/R SMA, L/R mCing,
	10.44	3	26	40	L/R medSF
	9.29	-3	5	58	
39	7.36	-24	-25	-8	L hip
55	7.10	9	11	-5	R caud
49	6.58	0	-19	13	L tha
26	6.45	-12	11	-5	L caud, L put
19	5.93	-57	-31	-2	L MT
12	5.92	-42	-40	43	L IP
6	5.31	42	5	31	R preC
<i>Fixation > words</i>					
750	9.75	-3	-64	40	L/R precun, L/R cun,
	9.51	9	-55	34	L/R mCing, L/R
	7.09	-15	-61	22	pCing
642	9.45	51	-55	25	R MT, R ang, R
	8.28	54	-58	13	supMar, R ST, R MO,
	8.07	45	-73	31	R RolOp
155	8.63	-45	-67	40	L ang, L MT, L MO
	7.30	-45	-58	22	
107	6.87	30	29	40	R MF, R SF
	6.84	21	23	49	
58	6.01	57	-10	4	R ST, R ins
	5.92	45	-10	-5	
31	5.91	-39	-19	1	L ST, L ins
	5.85	-45	-13	-2	
6	5.51	6	38	-8	R aCing, R MF
5	5.39	-24	29	43	L MF
<i>N letters (words), positive effect</i>					
456	11.75	15	-85	-8	R ling R calc, R SO,
	8.12	21	-94	7	R fus, R MO, R cun
	6.99	24	-88	22	

376	9.90	-12	-88	-8	L ling, L calc, L fus,
	8.89	-21	-82	-14	L IO, L cereb, L MO
	7.10	-21	-73	-8	
37	6.69	-48	-7	43	L preC, L postC
64	6.16	-24	-88	22	L MO, L SO
	6.01	-27	-88	13	

N letters (words), negative effect
(no significant clusters)

Syllable frequency (words), positive effect

19	6.07	-45	38	4	L IFtri
16	5.66	-42	5	28	L IFop
4	5.27	-3	23	43	L msF

Syllable frequency (words), negative effect
(no significant clusters)

Note. Abbreviations as in Table 5.

Table 7

Results of the region-of-interest analyses

<i>Group mask in:</i>	Fusiform				Middle temporal				Inferior/superior parietal			
<i>Individual ROI from:</i>	pseudo-fix		word-pseudo		pseudo-fix		word-pseudo		pseudo-fix		word-pseudo	
Evaluated contrast	<i>t</i>	<i>p</i>	<i>t</i>	<i>p</i>	<i>t</i>	<i>p</i>	<i>t</i>	<i>p</i>	<i>t</i>	<i>p</i>	<i>t</i>	<i>p</i>
Number of letters	-0.19	.848	-1.26	.214	0.51	.614	-0.58	.562	4.08	<.001	-1.14	.259
Word frequency	-2.10	.042	-0.78	.439	0.59	.557	2.76	.008	-3.23	.002	0.51	.613

Note. pseudo = pseudowords; fix = fixation.

Figure Captions

Figure 1. Graphic illustration of the trial sequence.

Figure 2. Significant left hemisphere activations for each contrast. Left hemisphere only, lateral and ventral view. Yellow and blue, thresholded at $|T| > 5.1$ (approximately corresponding to FWE-corrected $p < .05$); orange and light blue, thresholded at $|T| > 3.2$ (approximately corresponding to uncorrected $p < .001$). Orange/yellow, positive effects; blue, negative effects.

Figure 3. A (left): Comparison of significant clusters with and without duration modeling. Red and dark blue, significant only without duration modeling; orange and medium blue, significant only with duration modeling; yellow and light blue, significant in both analyses. Red-orange-yellow, positive effects; blue, negative effects. B (right): The three masks used in the region-of-interest analysis.

Figure 4. Results of the individual ROI analyses, showing the distributions of parametric contrast values (pseudoword length and word frequency effects) evaluated at individual contrast peaks (pseudoword vs. fixation and word vs. pseudoword, respectively) within the three global masks (in the fusiform, middle temporal, and inferior/superior parietal gyri). Each boxplot encompasses 44 points (one for each participant). Boxes enclose the middle half of the data, notched at the median.







