Taxi vs. Taksi: On Orthographic Word Recognition in the Left Ventral Occipitotemporal Cortex

Martin Kronbichler[1,2], Jürgen Bergmann[1], Florian Hutzler[3], Wolfgang Staffen[2], Alois Mair[2], Gunther Ladurner[2], and Heinz Wimmer[1]

Abstract

■ The importance of the left occipitotemporal cortex for visual word processing is highlighted by numerous functional neuroimaging studies, but the precise function of the visual word form area (VWFA) in this brain region is still under debate. The present functional magnetic resonance imaging study varied orthographic familiarity independent from phonological–semantic familiarity by presenting orthographically familiar and orthographically unfamiliar forms (pseudohomophones) of the same words in a phonological lexical decision task. Consistent with orthographic word recognition in the VWFA, we found lower activation for familiar compared with unfamiliar forms, but no difference between pseudohomophones and pseudowords. This orthographic familiarity effect in the VWFA differed from the phonological familiarity effect in left frontal regions, where phonologically unfamiliar pseudowords led to higher activation than phonologically familiar pseudohomophones. We suggest that the VWFA not only computes letter string representations but also hosts word-specific orthographic representations. These representations function as recognition units with the effect that letter strings that readily match with stored representations lead to less activation than letter strings that do not. ■

INTRODUCTION

Based on an impressive set of neuroimaging and electrophysiological findings, Cohen et al. (2000, 2002), Dehaene, Le Clec'H, Poline, Le Bihan, and Cohen (2002), and Dehaene et al. (2001) proposed that a region in the left ventral occipitotemporal cortex, at approximately $x = -43$, $y = -54$, $z = -12$, functions as a visual word form area (VWFA). However, different from its name, which suggests recognition of letter strings as whole visual–orthographic words, the VWFA was originally assumed to be only involved in the computation of strings of abstract letter identities, which serve as input to subsequent language-based processes. This limited prelexical string processing function was mainly based on a finding by Dehaene et al. (2002), showing that the VWFA exhibited about the same activation to letter strings of familiar words (e.g., *street*), which allow whole-word recognition, as to letter strings of pseudowords (e.g., *stroat*), which do not.

Our research group questioned this limited prelexical string processing function attributed to the VWFA based on the finding that an increase in the familiarity of letter strings in five levels from unfamiliar (i.e., pseudowords) to highly familiar (i.e., high-frequency words) was accompanied by a decrease in brain activation in the left occipitotemporal cortex including the VWFA (Kronbichler et al., 2004). This finding corresponded to the majority of imaging studies that contrasted pseudowords with words, and also found lower activation for the familiar letter strings of words in regions close to the VWFA (see the review by Mechelli, Gorno-Tempini, & Price, 2003). Based on this pattern—increased familiarity of whole letter strings accompanied by decreased VWFA activation—we suggested that the VWFA serves as an orthographic input lexicon, where incoming letter strings are matched against orthographic lexicon entries. This match is achieved faster and with less neuronal activity for high-frequency than for low-frequency words. The high activation in response to pseudowords was assumed to result from partial and possibly prolonged activation of more than one orthographic word entry (e.g., *stroat* may activate *street* and *stroll* and *boat* and *float*). Alternatively, the VWFA may not only host entries for whole visual words but also entries for recurring letter sequences and single letters, and a pseudoword may activate several such sublexical recognition units, compared to the activation of a single whole-word recognition unit by a familiar visual word.

The interpretation advanced by Kronbichler et al. (2004) of the inverse relation between letter string familiarity and VWFA activation was inspired by studies that found that experimentally increased familiarity of

[1]University of Salzburg, Austria, [2]Paracelsus Private Medical University, Austria, [3]University of Vienna, Austria

pictured objects (by repeated presentations) led to decreased activation in bilateral occipitotemporal regions (e.g., van Turennout, Bielamowicz, & Martin, 2003; Chao, Weisberg, & Martin, 2002; van Turennout, Ellmore, & Martin, 2000). Similarly, experimentally increased familiarity of faces led to decreased activation in a right fusiform gyrus region (Rossion, Schiltz, & Crommelinck, 2003). The reduced occipitotemporal activation to familiarized objects or faces—similar to the reduced activation to familiar letter strings—is suggestive of recognition, that is, instantiation of a single stored representation with little neuronal activity. In contrast, encoding of a new object or a new face as a new configuration of relevant features will be accompanied by increased neuronal activity.

A problem of the Kronbichler et al. (2004) study and of studies contrasting words and pseudowords is the confound between visual–orthographic familiarity of the letter strings and familiarity of the corresponding phonology. This confound is problematic because a most critical step in visual word processing consists in accessing phonology and meaning, and these processes are engaging semantic and phonological language areas of the brain. In this perspective, the familiarity-related VWFA activation may reflect little more than a side effect of semantic and phonological processes. Consistent with this possibility is the finding of Kronbichler et al. that not only the VWFA but also a left middle temporal and several left frontal brain regions exhibited decreasing activation in response to increasing word familiarity. One may reason that the more demanding phonological processing of an unfamiliar word may lead to prolonged activation of the letter string representation computed by the VWFA.

The present study avoided the confound between familiarity of the visual word and familiarity of the phonological form by presenting orthographically familiar and unfamiliar forms (i.e., pseudohomophones) of the very same words and also kept constant the overt response to these forms by asking for a phonological lexical decision (''Does this letter string sound like an existing word?''). Obviously, letter strings of pseudowords had to be included, and the contrast between the two types of unfamiliar forms (i.e., pseudohomophones vs. pseudowords) is of additional interest. Our focus was on differences in VWFA activation in response to these three item types. If VWFA activation predominantly reflects search of the orthographic word lexicon, then the orthographically familiar forms should result in lower VWFA activation than the pseudohomophones because the former find matching orthographic word entries and the latter do not. Orthographic lexicon search further predicts little difference in VWFA activation between the pseudohomophones and pseudowords, as neither of these letter string types finds a matching orthographic word entry. A different pattern is expected when differences in VWFA activation reflect only letter string support for the critical differences in

accessing the phonological lexicon in response to the task requirement. Obviously, only words and pseudohomophones can find a matching phonological word entry, whereas pseudowords should lead to an extensive and unsuccessful search of the phonological lexicon. These phonological processing differences should be reflected in lower activation for words and pseudohomophones than for pseudowords in left frontal phonological language areas (Jobard, Crivello, & Tzourio-Mazoyer, 2003; Mechelli et al., 2003; Fiez, Balota, Raichle, & Petersen, 1999). If the VWFA provides letter string support for these differing phonological processes, then one would expect the left frontal activation pattern mirrored in the VWFA.

The findings of the present study may also be of interest in relation to recent imaging studies, which relied on priming for studying left occipitotemporal activation in response to visual words. Devlin, Jamison, Gonnerman, and Matthews (2006), in this journal, provided priming data that were interpreted as speaking against orthographic whole-word recognition sensu Kronbichler et al. (2004), but also against prelexical letter string processing in the VWFA as proposed by Dehaene et al. (2002). Devlin et al. interpreted their findings as support for an interface function implying that the VWFA generates transient representations that link sensory information with conceptual information (see Discussion). Also based on a priming study, Dehaene, Cohen, Sigman, and Vinckier (2005) and Dehaene et al. (2004) proposed a functional differentiation of the left occipitotemporal cortex in terms of a hierarchical sequence of processing stages. Of specific interest here is that a region, slightly anterior to the original VWFA, was assumed to be involved in orthographic whole-word recognition. However, as will be pointed out in the Discussion, an inherent problem of priming studies is to realize orthographic whole-word priming without abstract letter priming. The present contrast between orthographically familiar and unfamiliar forms of the very same words may supply complementary information to the priming studies on the issue of visual–orthographic whole-word recognition in the VWFA.

METHODS

Participants

Participants were 24 German-speaking persons (21 males), all right-handed and with no history of neurological disease or reading difficulty. We started with a group of 12 normally reading adolescents (age, 15–16 years) from a longitudinal study on reading development, for whom the functional magnetic resonance imaging (fMRI) study was part of a larger assessment. An additional group of competent adult readers (university students, 20–34 years old) was included. These were expected to exhibit a more marked behavioral difference between familiar forms and pseudohomophones than

do adolescent readers. However, different from expectation, the two groups performed rather similarly on the phonological lexical decision task (see below). Informed consent was provided by each participant and, in the case of the adolescents, by a parent as well. All participants received payment for their participation.

Stimuli and Task

The 180 stimuli consisted of 60 orthographically familiar forms of German nouns, 60 orthographically unfamiliar forms of the same words (i.e., pseudohomophones), and 60 pseudowords. Examples for the three item types are *Taxi–Taksi–Tazi* or *Chaos–Kaos–Kuse*, and characteristics of these item types are shown in Table 1. The familiar forms consisted of 4–9 letters and began with a consonant (in upper case, following German spelling convention for nouns). Table 1 shows that the familiar forms and the pseudohomophones did not differ in number of letters, syllables, bigram frequency, and number of orthographic neighbors (i.e., words of the same length differing by one letter). The mean frequency of 86 occurrences per million according to the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993) indicates that the majority of words was of moderate to high frequency. Pseudowords were generated in such a way that they could not be distinguished from pseudohomophones by superficial characteristics such as absence of vowel letters or length. Because the main focus of the study was on the contrast between familiar and unfamiliar forms of the same words, matching of the pseudowords with the pseudohomophones was not perfect. As evident from Table 1, pseudowords were of slightly lower mean bigram frequency, and the mean number of neighbors was smaller. The latter difference can be expected to result in reduced response latencies because pseudowords with a low number of neighbors led to fewer errors and shorter latencies in lexical decision tasks than did pseudowords with a high number of

neighbors (e.g., Binder et al., 2003; Forster & Shen, 1996). However, despite this potential bias toward easier judgments of the present pseudowords, we found massively increased response latencies for pseudowords compared to pseudohomophones (see Results section).

A separate orthographic judgment task (i.e., ''Is this a correctly spelled word?'') ascertained that our participants had little difficulty in distinguishing between the familiar and the unfamiliar forms of the same words. This task was administered several weeks after the fMRI study and included only the adolescent participants. Nearly all (99%) of the familiar forms were judged as orthographically correct, whereas only 7% of the unfamiliar forms received such a judgment.

To examine differences in the brain response to the three stimulus types, a fast event-related design was used. Each item was displayed for 1600 msec with an interstimulus interval of 2100 msec during which a fixation cross was shown. This stimulus onset asynchrony of 3700 msec is not an integer of the repetition time of 2000 msec (see below), this enhances the efficiency of the design by sampling the hemodynamic response at different time points. The 180 stimuli were presented in two pseudorandomized lists, and each list was divided into two runs of 90 items, each composed of 30 items per stimulus type. In addition, 10 null events of 3700-msec duration with a fixation cross were included in each run. The two runs were separated by a short (1–2 min) break. The null events were included to improve evaluation of stimulus-related activation relative to baseline. The order of the 90 stimuli and of the 10 null events within each run was determined by a genetic algorithm (Wager & Nichols, 2003) that selects the most efficient sequence for testing stimulus contrasts. A critical feature for creating the two pseudorandomized lists was the sequencing of the familiar and the unfamiliar form of the same phonological word form. When in List 1 the familiar form was presented in the first run, which was the case for half of the words, then this order was reversed

Table 1. Means (Standard Deviations) for Item Characteristics and Task Performance

Measure	Words	Pseudohomophones	Pseudowords
Item characteristics			
Letters	5.6 (1.2)	5.6 (1.2)	5.6 (1.2)
Syllables	2.0 (0.4)	2.0 (0.4)	2.0 (0.4)
Neighbors	2.3 (2.8)	2.3 (2.5)	1.3 (1.7)
Bigram frequency	9965 (6679)	9767 (6903)	9171 (7447)
Word frequency	86 (148)	–	–
Phonological lexical decision task (fMRI)			
Reaction time (msec)	930 (352)	1075 (330)	1362 (338)
Error (%)	1 (2)	6 (5)	2 (2)

in List 2. Change of the list from participant to participant ascertained that the familiar form was equally often presented before and after the unfamiliar form of the same word. It also was avoided that three stimuli of the same type occurred in immediate succession.

Participants were familiarized with the phonological lexical decision task (i.e., "Does it sound like an existing word?") and with the response mode outside the scanner. Participants responded with the index finger ("yes") and middle finger ("no") of their right hands. Stimulus delivery and response registration was controlled by Presentation (Neurobehavioral Systems Inc., Albany, CA).

fMRI Data Acquisition and Analysis

During each of the two runs, 190 functional images sensitive to blood oxygenation level dependent (BOLD) contrast were acquired with a T2*-weighted echo-planar imaging sequence (echo time = 40 msec, TR = 2000 msec, flip angle = 86°, 21 slices with a thickness of 6 mm, 220-mm field of view with a 64×64 matrix resulting in 3.44×3.44-mm in-plane resolution). In addition, a low- ($3.5 \times 3.5 \times 6$ mm) and a high-resolution ($1 \times 1 \times 1.3$ mm) structural scan were acquired from each participant with T1-weighted MPRAGE sequences. A 1.5-T Intera Scanner (Philips Medical System, Best, the Netherlands) was used for magnetic resonance imaging.

Data analysis used SPM2 (www.fil.ion.ucl.ac.uk/spm). Functional images were realigned and unwarped, slice time corrected, and coregistered to the high-resolution structural image. The structural image was normalized to the Montreal Neurological Institute T1 template image, and the resulting parameters were used for normalization of the functional images, which were resampled to isotropic 3-mm³ voxels and smoothed with a 9-mm full width at half maximum Gaussian kernel. Statistical analysis was performed in a two-stage mixed-effects model. In the subject-specific first-level model, each stimulus type was modeled by a canonical hemodynamic response function and its temporal derivative. The few incorrectly answered and missed items were modeled as covariates of no interest. The functional data in these first-level models were high-pass filtered with a cutoff of 128 sec, and corrected for autocorrelation by an AR(1) model (Friston et al., 2002). In these first-level models, the parameter estimates reflecting signal change for each stimulus type versus fixation baseline (which consisted of the interstimulus interval and the null events) were calculated in the context of a general linear model (see Henson, 2004). Contrasts for the effect of interest were obtained for each participant from these first-level parameter estimates and used in a second-level random effects model. Contrasts of interest were examined by F tests. Regions with reliable activation differences had to consist of at least five voxels surviving a voxelwise threshold of $p < .01$, corrected for multiple comparisons using the false discovery rate (FDR; Genovese, Lazar, &

Nichols, 2002). The FDR accepts a specific proportion of false positives and is relatively liberal compared to other correction methods. The mean percentage signal change for specific regions and the event-related time course of the signal change were extracted with the MARSBAR toolbox (Brett, Anton, Valabregue, & Poline, 2002). To obtain event-related time courses, a deconvolution analysis using a finite impulse response (FIR) model was performed (see Henson, 2004). One should note, that such an FIR model makes no a priori assumption on the shape of the hemodynamic response.

RESULTS

Behavioral Results

Table 1 (bottom rows) shows performance on the phonological lexical decision task ("Does it sound like an existing word?") used during fMRI scanning. Of importance is the near absence of erroneous responses to words and pseudowords. Only the pseudohomophones led to a small increase in misjudgments. The nearly perfect rejection of pseudowords speaks against the possibility that the 2:1 ratio between yes responses (words and pseudohomophones) and no responses (pseudowords) led to a bias to respond with yes. An analysis of variance (ANOVA) of reaction times (correct responses only) found only item type reliable, $F(2,44) = 187.86$, $p < .001$, whereas neither age nor the age by item type interaction was reliable ($Fs < 1$). Post hoc t tests found the difference between all three item types reliable ($ps < .001$), but, as shown in Table 1, the reaction time difference between pseudowords and pseudohomophones (about 300 msec) was more than doubled compared to the reaction difference between words and pseudohomophones (about 150 msec). This pattern of reaction time results corresponds closely to the results of studies that used a phonological lexical decision task with a balanced number of items requiring yes and no responses (Martensen, Dijkstra, & Maris, 2005; Grainger, Spinelli, & Ferrand, 2000). It also corresponds to the results of studies that measured reading onset time for words, pseudohomophones, and pseudowords in mixed lists (e.g., Borowsky, Owen, & Masson, 2002; Grainger et al., 2000; Marmurek & Kwantes, 1996). This correspondence suggests that the present reaction pattern is not heavily affected by the unbalanced number of yes and no responses.

fMRI Results

Each item type was independently contrasted with the fixation baseline (which consisted of the null events and the interstimulus interval). Because of space limitations and because all three item types activated largely the same regions, an activation map is shown only for the orthographically familiar forms. Figure 1A shows that
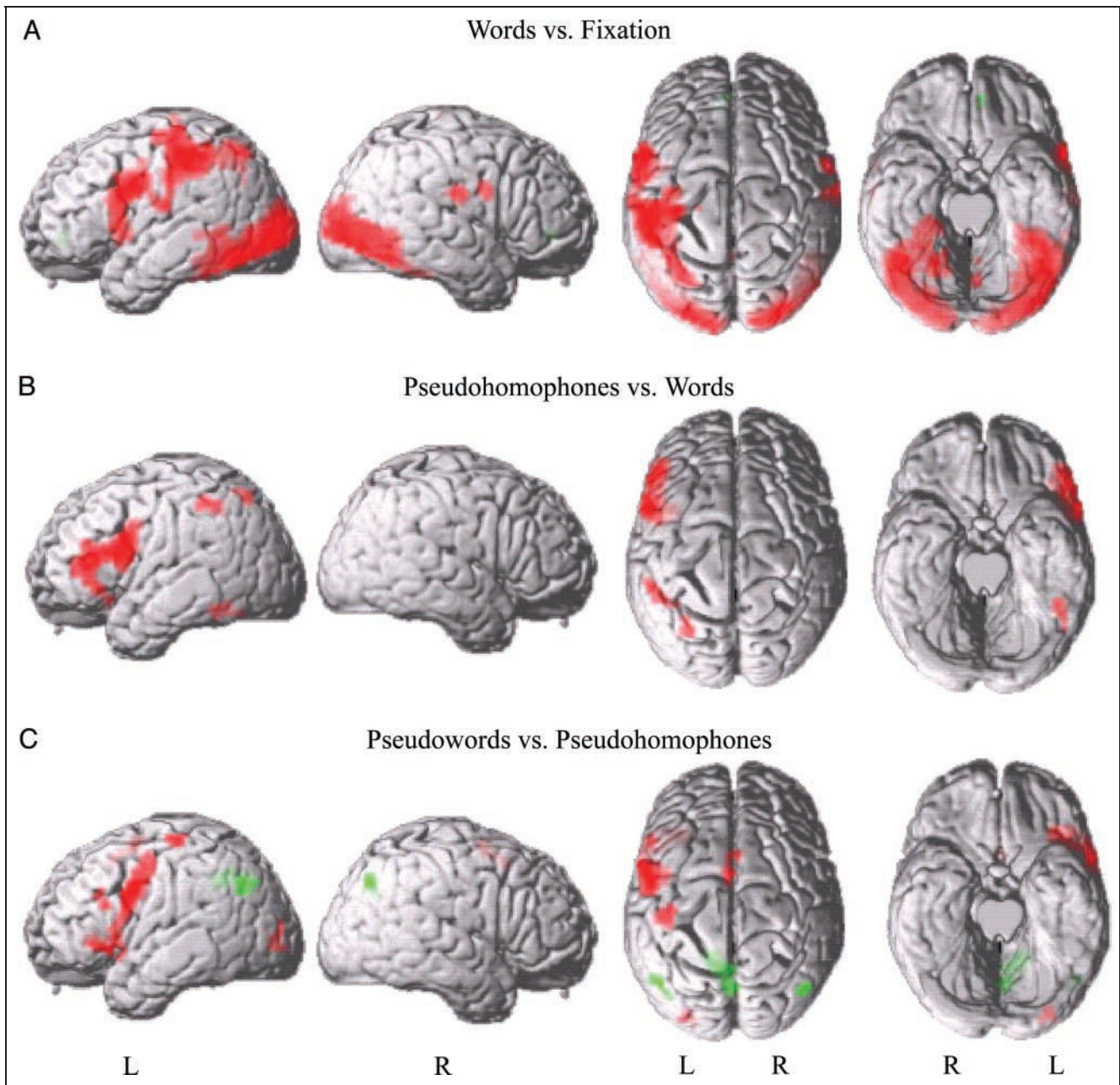
Figure 1. Differences in brain activity for (A) words compared to the fixation baseline, (B) pseudohomophones compared to words, and (C) pseudowords compared to pseudohomophones rendered on a standard canonical brain surface. Regions in red show higher activity for the item type mentioned first in each contrast; regions in green show higher activity for the item type mentioned second.

familiar forms activated bilateral occipital and occipito-temporal regions with a larger anterior extension in the left hemisphere. Additional activations in the left hemisphere were mainly found in superior temporal, temporoparietal, parietal, inferior frontal, and precentral regions. In the right hemisphere, only inferior frontal activation was found in addition to the occipital and occipitotemporal activations. A region in the medial orbitofrontal cortex showed reliable deactivation (in green) compared to baseline. The contrasts of pseudowords and pseudohomophones with baseline identified mainly the same areas as the contrast of familiar forms

with baseline. However, the extent of the activated regions was larger for pseudohomophones than for words and, again, larger for pseudowords than for pseudohomophones. In addition, pseudowords resulted in more areas with deactivation compared to baseline.

As the main interest of the study was on the expected activation difference between familiar forms and pseudohomophones in the left occipitotemporal cortex, a region-of-interest (ROI) analysis would have been justified. However, to gain information on additional brain regions exhibiting item type differences, all voxels with reliable activity differences ($p < .01$, FDR-corrected) com-

pared to fixation for at least one of the three item types were included in the analysis. This approach reduces the multiple comparison problem, but nevertheless includes all brain regions that are relevant for visual word processing as evident from Figure 1A. Due to the large number of voxels included, this is a relatively conservative approach compared to a ROI analysis. In a number of recent fMRI studies on visual word processing, a comparable approach was used (e.g., Carreiras, Mechelli, & Price, 2006; Dehaene et al., 2004).

Figure 1B and Table 2 (top rows) show the results of the contrast between orthographically familiar and unfamiliar forms of the same words. Higher activation for unfamiliar forms was found in the left fusiform and in the left inferior temporal gyrus along the occipitotemporal sulcus. Higher activation for unfamiliar forms was also found in the left inferior and superior parietal lobule and in a large left inferior frontal region, including the triangular and opercular parts of the inferior frontal gyrus and a part of the precentral gyrus. No region showed higher activation for familiar than for unfamiliar forms of the same word. The contrast between pseudo-

words and pseudohomophones shown in Figure 1C and Table 2 (bottom rows) identified a region in the left middle occipital gyrus and a number of left frontal regions (inferior frontal, supplementary motor area [SMA], precentral and postcentral gyrus) with higher activation for pseudowords. The opposite, that is, higher activity for pseudohomophones than pseudowords (shown in green), was found in the left and right angular gyri and in the precuneus. Of specific interest is that no left occipitotemporal region showed a difference between pseudowords and pseudohomophones. Even when lowering the threshold to $p < .01$ (uncorrected), no voxel of the left occipitotemporal cluster, which was identified by the pseudohomophone versus word contrast, exhibited higher activation for pseudowords compared to pseudohomophones.

For illustration of the item type effect on left occipitotemporal activation, we present in Figure 2 mean percentages of signal change relative to fixation baseline, and also the event-related signal change time courses. This was done for voxels within spheres of 5-mm radius around the two left occipitotemporal peak voxels, which

Table 2. Brain Regions with Reliable Differences in Brain Activity ($p < .01$, FDR-corrected) for Pseudohomophones Compared to Words and Pseudowords Compared to Pseudohomophones

Regions	Coordinates			Z	Extent (mm^3)	Approximate BA
	x	y	z			
Pseudohomophones > words						
L occipitotemporal (inferior temporal/fusiform)	−45	−48	−15	4.80	783	20/37
	−48	−60	−18	3.80		37
L superior parietal	−30	−66	57	4.25	513	7
L inferior parietal	−51	−42	51	4.18	972	40
L inferior frontal, opercular/precentral	−54	9	18	5.40	10,152	6/44
L inferior frontal, triangular	−51	30	15	5.14		45
Pseudowords > pseudohomophones						
L middle occipital	−27	−87	0	3.74	621	18
L postcentral	−39	−24	69	4.13	648	4
SMA	3	15	54	4.02	999	6
L inferior frontal, triangular	−51	24	30	3.64	324	44/45
L inferior frontal, triangular	−33	27	0	4.58	6345	47
L precentral	−48	3	33	4.53		6
Pseudowords < pseudohomophones						
L angular	−45	−69	39	5.00	459	39
R angular	45	−72	42	4.41	270	39
L precuneus	−15	−54	39	4.33	1485	–

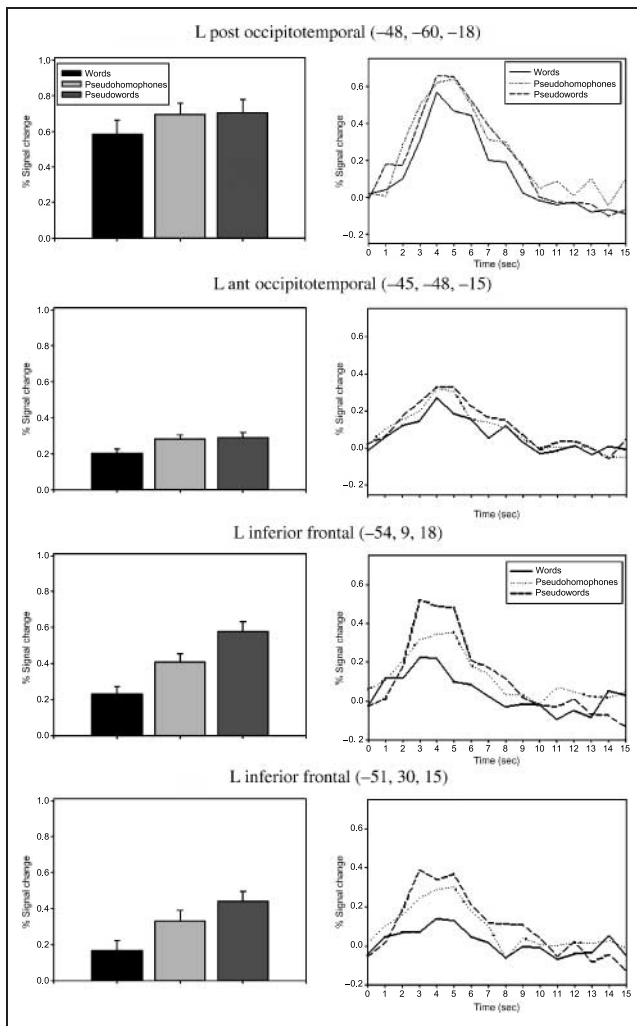BA = Brodmann's area; L = left; R = right.

Figure 2. Means and standard errors of percentage signal change and event-related time courses of signal change for words, pseudohomophones and pseudowords compared to fixation in four ROIs (see text). Note that 0 on the timescale corresponds to the onset of the stimuli that were presented for 1600 msec.

were identified in the contrast between words and pseudohomophones. For comparison, the same information is presented for the two left inferior frontal peak voxels of the word–pseudohomophone contrast. We preferred to extract signal changes for these spheres over average signal changes for the whole clusters, because the clusters in the inferior frontal and in the occipitotemporal cortex were relatively large, and we wanted to avoid averaging over potentially heterogeneous regions.

The overall shape of the time courses in Figure 2 corresponds to the expected shape of the hemodynamic response (e.g., Miezin, Maccotta, Ollinger, Petersen, & Buckner, 2000) and speaks for the validity of the data. Of interest is that there was no difference between the item types in the onset and the peak latency of the hemodynamic response, despite the differences in behavioral reaction time. For interpretation of the mean signal change patterns in Figure 2, it has to be remem-

bered that the regions were identified by the word–pseudohomophone contrast, so that the reduced mean signal change of the familiar forms is expected. Therefore, of main interest is the relation between the pseudohomophones and pseudowords. Figure 2 shows hardly any difference between pseudowords and pseudohomophones in mean signal change percentages for the two occipitotemporal ROIs. In contrast, markedly increased activation in response to pseudowords is evident for the left inferior frontal ROIs. This impression was supported by a reliable interaction, $F(1,23) = 10.89$, $p = .003$, between item type and brain region (occipitotemporal vs. inferior frontal) with averaged mean signal change for the two occipitotemporal and the two inferior frontal ROIs as dependent measure. Post hoc t tests showed that, for the occipitotemporal ROIs, words differed from both pseudohomophones and pseudowords (ps $< .003$), with no difference between pseudohomophones and pseudowords ($p = .785$). In contrast, in the inferior frontal ROIs, all three item types differed from each other (ps $< .001$).

A possible concern is that the lower activation for words may be due to a larger priming effect for words than for pseudohomophones. For example, *Taksi* presented in the first half of the item sequence (i.e., in the first run) may act as a more effective prime for *Taxi* in the second run than vice versa. This possibility would imply that the activation difference between words and pseudohomophones is only present in the second run. Such differential priming should be reflected in a reliable interaction between item type and run (first vs. second). However, an ANOVA with item type and run as factors and signal change (combined for the two left occipitotemporal ROIs) as dependent measure found the interaction not to be reliable, $F(2,46) < 1$, different from the reliable main effects for item type, $F(2,46) = 8.68, p = .001$, and run, $F(1,23) = 5.86, p = .024$. The absent interaction reflects the fact that the lower activation for words shown in Figure 2 was similarly present in the first and in the second run. The main effect of run resulted from the lower activation in the second run.

In addition to this ROI-based analysis, we limited the voxel-based analysis of the word–pseudohomophone contrast to the first run only. This contrast—similar to the one of the original analysis—again identified a left occipitotemporal region as well as left inferior parietal and frontal regions with higher activation for pseudohomophones. As expected from the fact that only half of the data was used, the Z and p values of the peak voxels were reduced.

Although the adolescent and the adult participants showed similar behavioral performance on the phonological lexical decision task, we examined whether the item type effect on percentage signal change (combined for the two left occipitotemporal ROIs) differed for the two groups. This was not the case, as no reliable interaction between age group (adolescents vs. adults) and

item type was found, $F(2,44) < 1$. The age effect was also not reliable, $F(1,22) < 1$.

DISCUSSION

With respect to localization of brain activation, it is of importance that the present contrast between orthographically familiar and unfamiliar forms of the very same words identified a left occipitotemporal region with reduced activation for the familiar forms. Reduced activation was also found in left inferior parietal regions and in a large left inferior frontal region. As the present study was focused on the function of the VWFA as specified by Cohen et al. (2002), the correspondence between the original VWFA and the presently identified left occipitotemporal region is of interest. Such spatial correspondence is given as the present two peak voxels with the most reliable difference between familiar and unfamiliar forms at −45, −48, −15 and −48, −60, −18 are close to the coordinates of the original VWFA at −43, −54, −12. The posterior one of the present occipitotemporal peak activations is in close proximity to the mean coordinates of −44, −58, −15 found in a meta-analysis by Jobard et al. (2003). Of specific interest for the issue of orthographic whole-word recognition is the spatial correspondence between the present anterior occipitotemporal peak voxel at −45, −48, −15 and the peak voxel at −48, −48, −8 that was identified by Dehaene et al. (2004) as being involved in orthographic word recognition (at least for short words). Given the spatial correspondence between the left occipitotemporal region identified here and the original VWFA, we refer to the present left occipitotemporal region as VWFA in the following.

The present finding of lower VWFA activation for orthographically familiar compared to unfamiliar forms of the same words, stands in marked contrast to Ischebeck et al. (2004) who, similar to the present study, contrasted familiar and unfamiliar forms of Japanese words. Neither the VWFA nor any other left occipitotemporal region was found to exhibit reduced activation for familiar compared to unfamiliar forms. Specifically, neither the contrast between familiar and unfamiliar writings of existing words, nor the contrast between familiar forms of words and pseudowords, identified the VWFA or any other left occipitotemporal region. Lower activation for familiar compared to unfamiliar forms was found in left and right temporoparietal and left and right frontal regions. The only posterior region exhibiting reduced activation for familiar forms was in the right precuneus. These localizations of the orthographic familiarity effect, specifically the right-hemisphere involvement, seems to indicate that the neuronal reading network established for the processing of the Japanese syllabic writing system may differ from the network established for alphabetic writing systems, for which the large majority of imaging studies showed predominant left-hemisphere

involvement (e.g., Jobard et al., 2003). For a comprehensive review on neuronal differences between alphabetic and nonalphabetic scripts, see Bolger, Perfetti, and Schneider (2005).

The present study attempted to control for the natural confound between visual–orthographic familiarity and phonological–semantic familiarity by presenting orthographically familiar and unfamiliar forms of the very same words together with pseudowords in a phonological lexical decision task. As pointed out in the Introduction, the confound between orthographic familiarity and phonological familiarity poses a problem for interpreting the Kronbichler et al. (2004) finding of decreasing VWFA activation in response to parametrically increasing visual word frequency, and also for findings showing lower VWFA activation to words compared to pseudowords (reviewed by Mechelli et al., 2003). Kronbichler et al. interpreted the decreased VWFA activation in response to increased visual word familiarity as reflecting differences in accessing the orthographic word lexicon in the VWFA. However, the confound between visual–orthographic and phonological familiarity also allows the interpretation that the familiarity-related differences in VWFA activation reflect little more than letter string support for familiarity-related differences in phonological processing. Based on the hypothesized orthographic lexicon function of the VWFA, we expected that orthographically familiar forms will result in lower activation than orthographically unfamiliar forms, with little difference between the unfamiliar forms of words (i.e., pseudohomophones) and pseudowords. Based on the letter string support function of the VWFA, we expected lower activation for both words and pseudohomophones compared to pseudowords. This was expected because both words and pseudohomophones find a matching phonological entry, and pseudowords do not. The observed VWFA activation pattern obviously corresponds to the prediction derived from the orthographic lexicon function.

The pattern expected from phonological processing differences, posed by judging the three item types as sounding like existing words, was observed in left inferior frontal regions. Here the pseudohomophones led to higher activation than words in a large left inferior frontal cluster, and pseudowords led to a further increase. An interpretation of the left inferior frontal activation differences in terms of phonological processing demands is in general correspondence with a number of studies that found higher left inferior frontal activation in response to pseudowords compared to words (e.g., Jobard et al., 2003; Mechelli et al., 2003; Fiez et al., 1999). The pseudoword versus pseudohomophone contrast of the present study not only identified a large left inferior frontal cluster, but also a cluster in the SMA with higher activation for pseudowords compared to pseudohomophones. The SMA is assumed to be involved in speech production (e.g., Bohland & Guenther, 2006; Ziegler,

Kilian, & Deger, 1997). The higher activation for pseudo-words may reflect higher demands on subvocal pronunciation assembly, because pseudowords require the assembly of a new pronunciation, whereas for pseudohomophones, an existing phonological form can be accessed. In relation to the marked difference between pseudohomophones and pseudowords in left frontal regions, the absence of such a difference in the VWFA speaks against the letter string support function of the VWFA, which would predict correspondence between activation patterns in left frontal language areas and in the VWFA.

Of methodological importance is that the pattern of left frontal activations, reflecting differing phonological processing demands, corresponds to the pattern of reaction times (words < pseudohomophones < pseudowords) in the phonological lexical decision task. In contrast, the VWFA activation pattern does not correspond to this reaction time pattern, because pseudowords elicited about the same activation as pseudohomophones. This absence of an activation increase from pseudohomophones to pseudowords is remarkable because the reaction time increase from pseudohomophones to pseudowords was more than doubled compared to the reaction time increase from words to pseudohomophones. More importantly, this large reaction time increase from pseudohomophones to pseudowords found no correspondence in VWFA activation, whereas the smaller reaction time increase from the orthographically familiar to the orthographically unfamiliar forms of the same words did find a correspondence in VWFA activation. This dissociation between reaction times and brain activation in the VWFA speaks against the concern that the activation differences in the VWFA may simply reflect time-on-task, as suggested by Binder, Medler, Desai, Conant, and Liebenthal (2005). One may be concerned that the absence of an increase in VWFA activation from pseudohomophones to pseudowords is due to a saturation of the BOLD response to the pseudohomophones, which would prevent a further increase. Inspection of the signal change values in Figure 2 does not favor such an interpretation, because for the anterior occipitotemporal ROI, the signal change values to pseudohomophones were rather low, and a further increase for pseudowords should have been possible. Furthermore, in inferior frontal regions, pseudohomophones elicited slightly higher activation than in the anterior occipitotemporal cortex, but in inferior frontal regions, there was a further activation increase from pseudohomophones to pseudowords.

The present main finding, that is, reduced VWFA activation in response to orthographically familiar compared to orthographically unfamiliar forms of the very same words, corresponds to the findings of priming studies, which found reduced VWFA activation when target words were subliminally primed by the same words in different letter case (Devlin et al., 2006; Dehaene et al., 2001, 2004). The idea here is that such a prime preactivates the corresponding orthographic word entry, so that it responds with reduced activation to the following target. Dehaene et al. (2004) contrasted the effect of two types of primes on target word activation. One type was the same French word in different letter case, resulting in prime–target pairs such as *REFLET–reflet*. The other type of prime was a circular anagram of the target, that is, a different French word, which consisted of the same letters as the target, resulting in prime–target pairs such as *TREFLE–reflet*. Participants had to judge the syllable length of the targets. An anterior segment (−48, −48, −8) of the original VWFA exhibited lower activation when targets were primed by same words as compared to anagrams. This effect was interpreted as orthographic whole-word priming. One may note that the same word primes share with the target the same letters in the same positions (relative to each other), whereas the anagram primes present the letters of the target in positions different from the target. Therefore, one may be concerned that the increased priming by the same word primes may not result from orthographic whole-word priming, but from the higher letter string similarity between prime and target offered by the same word primes (i.e., same letters in relative same positions). Another potential problem is that the same word primes share phonology and meaning with the target words, and the anagrams do not. The present finding of lower VWFA activation for familiar compared to unfamiliar orthographic forms of the same words speak against these alternative interpretations of the occipitotemporal priming results. Obviously, the inherent problem of priming studies to differentiate between letter string priming and orthographic word priming does not exist for the present study. Furthermore, by presenting familiar and unfamiliar forms of the very same words, the natural confound between familiarity of orthographic word forms and familiarity of the corresponding phonology and meaning was avoided.

The present finding of reduced VWFA activation in response to familiar words also corresponds to findings of the subliminal priming study of Devlin et al. (2006). This study, using a standard lexical decision task, presented prime–target pairs that consisted of the same words such as *cabin–CABIN*, or of the same pseudowords such as *solst–SOLST*. The critical finding was that word repetition led to reduced activation in the left fusiform gyrus corresponding to the VWFA, whereas pseudoword repetition did not. This pattern suggests that the word primes preactivated specific orthographic word representations, whereas such preactivation was not possible for pseudowords. This finding is consistent with our position that the VWFA hosts orthographic word representations that are activated by appropriate letter input. However, Devlin et al. question this interpretation based on the additional finding, showing that prime–target pairs of the type *corner–CORN* resulted in

about the same reduction of VWFA activation as prime target pairs such as *cabin–CABIN*. This critique of the orthographic lexicon function of the VWFA assumes that primes such as *corner* cannot lead to a preactivation of those orthographic word entries that then are activated by targets such as *CORN*. However, this assumption can be questioned, as primes such as *corner* may not only preactivate their fully corresponding orthographic word entry, but also orthographic word entries corresponding to the first letters. Furthermore, there is the possibility of letter string priming, although the absence of priming for pseudowords speaks against this possibility. Nevertheless, the finding of Devlin et al., that prime–target pairs such as *corner–CORN* result in about the same VWVA reduction as prime–target pairs such as *cabin–CABIN*, poses a challenge to the orthographic word lexicon function of the VWFA and deserves further study.

We should also note that Devlin et al. (2006), based on additional semantic priming data, proposed an interesting alternative to both the orthographic lexicon function and the prelexical letter string function of the VWFA. They argue that the VWFA acts as a general interface area between sensory information from different modalities and higher order conceptual information. In the case of visual words, this area computes transient intermediate representations that integrate low-level visual information (e.g., the features distinguishing between *acre* and *acne*) and phonological and semantic information. A role of orthographic word representations in visual word processing is not specified. The predictions, following from the Devlin et al. proposal for VWFA activation in the present study, correspond to the already mentioned letter string support function of the VWFA, that is, pseudowords should have elicited higher activation than pseudohomophones, due to prolonged unsuccessful search for meaning. Clearly, the present finding of similar VWFA activation for pseudohomophones and pseudowords is inconsistent with this prediction. However, this negative finding may not affect the general idea of Devlin et al. that the left occipitotemporal cortex is involved in linking visual or other sensory information with higher order conceptual information. We only would like to insist that in the case of visual words, not low-level visual features (like the ones differentiating *acne* and *acre*) but orthographic word representations (like *hare* and *hair*) are linked with phonology and meaning. We also note that the present study, which presented only visual words, cannot provide evidence on a further main aspect of the theorizing of Devlin et al., namely, that the left occipitotemporal cortex is not specialized for visual word processing but functions as a multimodal interface to higher order information (for discussions, see Cohen & Dehaene, 2004; Price & Devlin, 2003, 2004).

The main contribution of the present study to the ongoing debate on the function of the VWFA in visual word processing is the finding that the familiar, well-known letter sequences of visual words result in lower VWFA activation than unfamiliar homophonic letter sequences of the same words. This orthographic familiarity effect supports the position advanced by Kronbichler et al. (2004) that the VWFA is not only engaged in letter string computation, but also hosts an orthographic lexicon in the sense of a memory system containing word-specific orthographic representations (see Coltheart, 2004). These representations function as recognition units for incoming letter string information. In this perspective, processing of familiar words depends in the same way on stored memory representations as processing of familiar faces or familiar visual objects.

Acknowledgments

REFERENCES

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

Binder, J. R., McKiernan, K. A., Parsons, M. E., Westbury, C. F., Possing, E. T., Kaufman, J. N., et al. (2003). Neural correlates of lexical access during visual word recognition. *Journal of Cognitive Neuroscience, 15,* 372–393.

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.

Bohland, J. W., & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *Neuroimage, 32,* 821–841.

Bolger, D. J., Perfetti, C. A., & Schneider, W. (2005). Cross-cultural effect on the brain revisited: Universal structures plus writing system variation. *Human Brain Mapping, 25,* 92–104.

Borowsky, R., Owen, W. J., & Masson, M. E. J. (2002). Diagnostics of phonological lexical processing: Pseudohomophone naming advantages, disadvantages, and base-word frequency effects. *Memory & Cognition, 30,* 969–987.

Brett, M., Anton, J. L., Valabregue, R., & Poline, J. B. (2002). *Region of interest analysis using an SPM toolbox* [Abstract]. Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2–6, 2002, Sendai, Japan. Available on CD-ROM in *Neuroimage, 16(2)*.

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.

Chao, L. L., Weisberg, J., & Martin, A. (2002). Experience-dependent modulation of category-related cortical activity. *Cerebral Cortex, 12,* 545–551.

Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: The case for the visual word form area. *Neuroimage, 22,* 466–476.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., et al. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123,* 291–307.

Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain, 125,* 1054–1069.

Coltheart, M. (2004). Are there lexicons? *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 57,* 1153–1171.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences, 9,* 335–341.

Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J. B., Le Bihan, D., et al. (2004). Letter binding and invariant recognition of masked words—Behavioral and neuroimaging evidence. *Psychological Science, 15,* 307–313.

Dehaene, S., Le Clec'H, G., Poline, J. B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: A prelexical representation of visual words in the fusiform gyrus. *NeuroReport, 13,* 321–325.

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J. F., Poline, J. B., et al. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience, 4,* 752–758.

Devlin, J. T., Jamison, H. L., Gonnerman, L. M., & Matthews, P. M. (2006): The role of the left posterior fusiform gyrus in reading. *Journal of Cognitive Neuroscience, 18,* 911–922.

Fiez, J. A., Balota, D. A., Raichle, M. E., & Petersen, S. E. (1999). Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron, 24,* 205–218.

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning Memory and Cognition, 22,* 696–713.

Friston, K. J., Glaser, D. E., Henson, R. N. A., Kiebel, S., Phillips, C., & Ashburner, J. (2002). Classical and Bayesian inference in neuroimaging: Applications. *Neuroimage, 16,* 484–512.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage, 15,* 870–878.

Grainger, J., Spinelli, E., & Ferrand, L. (2000). Effects of baseword frequency and orthographic neighborhood size in pseudohomophone naming. *Journal of Memory and Language, 42,* 88–102.

Henson, R. N. A. (2004). Analysis of fMRI time series: Linear time-invariant models, event-related fMRI and optimal experimental design. In R. S. J. Frackowiak, K. J. Friston, C.

Frith, R. Dolan, C. J. Price, S. Zeki, et al. (Eds.), *Human brain function* (2nd ed., pp. 793–822). London: Academic Press.

Ischebeck, A., Indefrey, P., Usui, N., Nose, I., Hellwig, F., & Taira, M. (2004). Reading in a regular orthography: An fMRI study investigating the role of visual familiarity. *Journal of Cognitive Neuroscience, 16,* 727–741.

Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: A metanalysis of 35 neuroimaging studies. *Neuroimage, 20,* 693–712.

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.

Marmurek, H. H. C., & Kwantes, P. J. (1996). Reading words and wirds: Phonology and lexical access. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 49,* 696–714.

Martensen, H., Dijkstra, T., & Maris, E. (2005). A werd is not quite a word: On the role of sublexical phonological information in visual lexical decision. *Language and Cognitive Processes, 20,* 513–552.

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

Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E., & Buckner, R. L. (2000). Characterizing the hemodynamic response: Effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage, 11,* 735–759.

Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *Neuroimage, 19,* 473–481.

Price, C. J., & Devlin, J. T. (2004). The pros and cons of labelling a left occipitotemporal region: ''The visual word form area.'' *Neuroimage, 22,* 477–479.

Rossion, B., Schiltz, C., & Crommelinck, M. (2003). The functionally defined right occipital and fusiform ''face areas'' discriminate novel from visually familiar faces. *Neuroimage, 19,* 877–883.

van Turennout, M., Bielamowicz, L., & Martin, A. (2003). Modulation of neural activity during object naming: Effects of time and practice. *Cerebral Cortex, 13,* 381–391.

van Turennout, M., Ellmore, T., & Martin, A. (2000). Long-lasting cortical plasticity in the object naming system. *Nature Neuroscience, 3,* 1329–1334.

Wager, T. D., & Nichols, T. E. (2003). Optimization of experimental design in fMRI: A general framework using a genetic algorithm. *Neuroimage, 18,* 293–309.

Ziegler, W., Kilian, B., & Deger, K. (1997). The role of the left mesial frontal cortex in fluent speech: Evidence from a case of left supplementary motor area hemorrhage. *Neuropsychologia, 35,* 1197–1208.