

Left ventral occipitotemporal activation during orthographic and semantic processing of auditory words



Philipp Ludersdorfer^{a,*}, Heinz Wimmer^a, Fabio Richlan^a, Matthias Schurz^a, Florian Hutzler^a, Martin Kronbichler^{a,b}

^a Centre for Cognitive Neuroscience, University of Salzburg, Austria

^b Neuroscience Institute, Christian-Doppler-Clinic, Paracelsus Medical University, Salzburg, Austria

ARTICLE INFO

Article history:

Received 18 May 2015

Accepted 19 September 2015

Available online 28 September 2015

Keywords:

Neuroimaging

fMRI

Orthography

Semantics

Ventral occipitotemporal cortex

VWFA

ABSTRACT

The present fMRI study investigated the hypothesis that activation of the left ventral occipitotemporal cortex (vOT) in response to auditory words can be attributed to lexical orthographic rather than lexico-semantic processing. To this end, we presented auditory words in both an orthographic (“three or four letter word?”) and a semantic (“living or nonliving?”) task. In addition, a auditory control condition presented tones in a pitch evaluation task. The results showed that the left vOT exhibited higher activation for orthographic relative to semantic processing of auditory words with a peak in the posterior part of vOT. Comparisons to the auditory control condition revealed that orthographic processing of auditory words elicited activation in a large vOT cluster. In contrast, activation for semantic processing was only weak and restricted to the middle part vOT. We interpret our findings as speaking for orthographic processing in left vOT. In particular, we suggest that activation in left middle vOT can be attributed to accessing orthographic whole-word representations. While activation of such representations was experimentally ascertained in the orthographic task, it might have also occurred automatically in the semantic task. Activation in the more posterior vOT region, on the other hand, may reflect the generation of explicit images of word-specific letter sequences required by the orthographic but not the semantic task. In addition, based on cross-modal suppression, the finding of marked deactivations in response to the auditory tones is taken to reflect the visual nature of representations and processes in left vOT.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

Over the last two decades a multitude of neuroscientific studies has shown that the left ventral occipitotemporal cortex (vOT) plays a critical role in reading. In particular, results from lesion studies suggest that damage to this region causes a relatively isolated deficit in visual word recognition (e.g., Leff et al., 2001; Cohen et al., 2003; Gaillard et al., 2006). In addition, functional neuroimaging studies have shown that the left vOT is one of the most consistently activated brain regions during word reading tasks (Jobard et al., 2003; Mechelli et al., 2003; Turkeltaub et al., 2002). One of the most prominent accounts of left vOT functioning is the so-called “Visual Word Form Area” (VWFA) hypothesis by Dehaene, Cohen and colleagues (e.g., Cohen et al., 2000, 2002; Dehaene and Cohen, 2011) which suggests that a circumscribed region in the middle segment of the left vOT [$y = -50$ to -60 in Montreal Neurological Institute (MNI) standard space] becomes specialized for the encoding of orthographic stimuli (i.e., written words) through

reading experience. However, it should be noted that the assumption that the left vOT hosts neuronal representations tuned to written words is controversial. A prominent opposing view is that the region acts as interface area between generic visual input and phonology and semantics (Price and Devlin, 2003, 2004, 2011).

One of the main assumptions of the VWFA hypothesis by Dehaene and Cohen (2011) is that orthographic representations in left vOT are primarily tuned to sublexical units (i.e., frequent letter-sequences within words). This assumption is based on studies that have shown that pseudowords (i.e., non-words that are consistent with the structural constraints of the writing system) elicit activation in left vOT at least as strongly as do words (e.g., Dehaene et al., 2002). In addition, other studies have shown that activation in this region is modulated by the familiarity of sublexical orthographic features (e.g., the frequency of pairs or triplets of letters) with frequent features eliciting higher activation relative to less frequent features (Binder et al., 2006; Vinckier et al., 2007). However, more recent studies that have simultaneously investigated the effects of multiple factors (e.g., sublexical and lexical frequency, length, etc.) on left vOT activation during reading have consistently failed to find sensitivity to sublexical familiarity in the left vOT (Hauk et al., 2008a; Graves et al., 2010; Woollams et al., 2011). Woollams

* Corresponding author at: Centre for Cognitive Neuroscience, University of Salzburg, Hellbrunnerstrasse 34/II, 5020 Salzburg, Austria.

E-mail address: philipp.ludersdorfer@sbg.ac.at (P. Ludersdorfer).

et al. (2011), for example, tried to dissociate the effects of sublexical familiarity (i.e., bigram frequency) and lexicality (i.e., words vs. pseudowords). While an effect of lexicality (i.e., pseudowords > words) was found in left vOT, an effect of sublexical familiarity was only found in a more posterior inferior occipital region.

Our research group questioned the limitation to sublexical letter string computation of the VWFA and proposed that the region also hosts representations of frequently encountered whole words (Kronbichler et al., 2004). In this perspective, the left vOT region would be functionally equivalent to the orthographic word lexicon of cognitive dual-route models of reading (Coltheart et al., 2001; Coltheart, 2004). Initial evidence for this came from Kronbichler et al. (2004), who showed that activation in left vOT was affected by the lexical frequency of written words with reduced activation for high relative to low-frequency words. This was interpreted as reflecting easier access to frequently relative to rarely used whole-word representations. Although other early neuroimaging studies failed to find a sensitivity to lexical frequency in left vOT regions (Fiez et al., 1999; Fiebach et al., 2002), more recent studies consistently replicated this effect (Hauk et al., 2008a, 2008b; Graves et al., 2010). Other studies have also shown that left vOT exhibits reduced activation to familiar compared to unfamiliar orthographic forms of the same phonological words (e.g., TAXI vs. TAKSI; Kronbichler et al., 2007, 2009; Bruno et al., 2008; Van der Mark et al., 2009; Twomey et al., 2011). These “orthographic familiarity” effects are also interpreted as reflecting orthographic whole-word coding in left vOT. In addition, Schurz et al. (2010) found a lexicality (words vs. pseudowords) by length (short vs. long letter strings) interaction effect on activation in left vOT. Specifically, a length effect (= higher activation for long relative to short letter strings) was present for pseudowords but absent for words. The absence of a length effect for familiar words is expected when letter strings are assimilated by whole-word representations. Specific support for orthographic whole-word coding in left vOT was also provided by an fMRI priming study (Glezer et al., 2009), which found that the priming effect on vOT activation present for word repetition (coat-coat) disappeared when the prime differed in just one letter from the target word (boat-coat). In contrast, for pseudowords the priming effect was dependent on the number of shared letters.

In the debate around the VWFA it has also been questioned whether left vOT is restricted to visual (word) processing as assumed by the original VWFA hypothesis (Dehaene et al., 2002). Instead, it has been argued, the region might have a polymodal function (Price and Devlin, 2003). Interestingly, a similar interpretation has recently been presented by Dehaene and Cohen (2011) who have suggested that the VWFA might be a “meta-modal” reading area. This was based on studies showing the left vOT activation in congenitally blind individuals during reading of Braille (e.g., Büchel et al., 1998; Reich et al., 2011). However, in a recent fMRI study our research group has argued that it might be premature to generalize these findings from congenitally blind to sighted individuals (Ludersdorfer et al., 2013). We showed that the left vOT region which exhibited an orthographic familiarity effect (i.e., higher activation for visual pseudowords relative to words) was also activated for unfamiliar visual stimuli (i.e., false-fonts) and strongly deactivated for unfamiliar auditory stimuli (i.e., reversed spoken words) relative to rest. According to the phenomenon of cross-modal suppression (Laurienti et al., 2002) such deactivations during auditory processing should occur in brain regions dedicated to visual processes. Therefore, the findings of Ludersdorfer et al. were taken to speak for a visual rather than meta-modal role of the left vOT.

Complementary to reading research, an interesting approach to investigate orthographic processing in left vOT has been presented by studies presenting auditory words in the context of spelling or writing tasks. The relatively few studies in this field have consistently reported left vOT activation (for reviews see Purcell et al., 2011; Planton et al., 2013). A recent spelling study from our lab (Ludersdorfer et al., 2015) investigated whether left vOT activation during spelling can be attributed to accessing orthographic whole-word representations. We

presented a spelling task in which participants had to indicate whether a visually presented letter was present in the spelling of an auditorily presented word. In the critical condition correct spelling decisions had to be based on known word spellings (i.e., orthographic whole-word representations) since they could not be based on sublexical phoneme–letter conversions. No such reliance on orthographic word representations was possible in an additional spelling condition which presented unfamiliar pseudowords. Here, decisions had to be based on sublexical-phonological information (i.e., phoneme-letter correspondences). Consistent with the assumption of orthographic whole-word representations in left vOT we found that the decisions based on known word spellings led to higher activation in left vOT compared to sublexical-phonological decisions.

The present study attempted to strengthen the evidence for visual-orthographic whole-word codes in left vOT. In reading studies, a major concern has been that the left vOT activation differences, which are taken to reflect lexical (= whole word) orthographic processes (e.g., Kronbichler et al., 2004) could, in fact, be driven by lexico-semantic processes (Devlin et al., 2006; Duncan et al., 2010). Evidence for this comes from fMRI priming studies showing reduced left vOT activation for words preceded by a semantically related prime relative to an unrelated prime (Wheatley et al., 2005; Devlin et al., 2006). In addition, some studies have found a positive relation between activation in left vOT regions and semantic variables such as word imageability (Wise et al., 2000; Sabsevitz et al., 2005; Hauk et al., 2008b). Critically, the argument that the left vOT activation could be driven by semantics also applies to findings from spelling studies. For example, the higher vOT activation observed by Ludersdorfer et al. (2015) for word compared to pseudoword spelling could be taken to be affected by the obvious difference in the availability of word meaning.

To disentangle lexical orthographic and lexico-semantic contributions to left vOT activation, the present study presented the same auditory words in both an orthographic and a semantic decision task. In the orthographic task, participants had to indicate whether the written form of the presented word consisted of three or four letters. The fact that all words (German nouns) had three phonemes ascertained that participants had to rely on known word spellings (i.e., orthographic whole-word representations) since sublexical-phonological strategies (e.g., counting the number of heard phonemes) would have led to erroneous responses. In the semantic task, participants had to decide whether the presented word referred to a living or a nonliving entity. If left vOT activation is indeed mainly driven by lexico-semantic processes, then higher activation for semantic relative to orthographic processing of auditory words was expected. The opposite pattern was expected when left vOT activation can be attributed to accessing orthographic whole-word codes (Ludersdorfer et al., 2015). One may also note that the latter finding would be generally consistent with an orthography-specific function of left vOT as assumed by the VWFA hypothesis (Dehaene and Cohen, 2011). In contrast, no differentiation or even higher activation for semantic relative to orthographic processing would speak for alternative accounts of left vOT function (Price and Devlin, 2011).

In addition to the direct comparison, activation for orthographic and semantic processing of auditory words in left vOT was evaluated in relation to a auditory control task in which participants had to evaluate the pitch of presented tones. The response of the left vOT to the tones was of interest for the hypothesis that processes and representations in this region are visual (Dehaene et al., 2002) rather than meta-modal as has been recently suggested (Dehaene and Cohen, 2011). As mentioned above, Ludersdorfer et al. (2013) interpreted a marked negative response of left vOT to unfamiliar reversed speech stimuli together with a positive response to visual stimuli as support for a visual role of left vOT. The tone stimuli of the present pitch evaluation task are much simpler than reversed spoken words. Therefore, it was of interest whether the present tones would also result in a negative vOT response as expected from cross-modal suppression of a visual region.

Methods

Participants

Twenty-nine German-speaking participants (14 females) aged 18 to 35 years ($M = 26$ years) were recruited for the present fMRI study. All participants were right-handed, had normal or corrected-to-normal vision and reported no history of neurological disease or reading/spelling difficulties. All gave written informed consent and were paid for participation.

Tasks and procedure

In the scanner, participants performed three tasks: in the orthographic task, participants were presented with auditory words and had to indicate with a button press whether the written form corresponding to the presented word consisted of three or four letters. In the semantic task, participants were also presented with auditory words and had to decide whether the presented word referred to a living or a nonliving entity. In the pitch evaluation task, participants were presented with tones and had to decide whether the presented tone was high or low in pitch.

Each of the three tasks was presented in six blocks of four trials. Each block started with an instruction screen for 1600 ms indicating the task (i.e., orthographic, semantic, or pitch evaluation). Each of the following trials started with the presentation of a fixation cross centrally on the screen for 300 ms. Next, in the orthographic and the semantic task an auditory word (the average audio length was 2024 ms) was presented preceded by its definite article. In the pitch evaluation task a square wave tone (varying in length from 1600 to 1900 ms) was presented instead of a word. During the presentation of the word or tone the fixation cross remained visible on the screen. Following the auditory presentation, visual cues (i.e., single letters) appeared to the left and the right of the fixation cross. These cues represented the two response alternatives of the respective task. In the orthographic task, for example, the letters D (the beginning letter of DREI [German for three]) and V (the beginning letter of VIER [four]) were presented. The visual cues were visible for the entire response interval which was jittered between 3900 and 4700 ms. The total length of each block was 23.3 s. Across participants, the order of block presentation was pseudo-randomized. Between the task blocks fixation periods of 14 s were inserted. The total length of the experiment amounted to approximately 11 min.

Participants were familiarized with all tasks outside the scanner. During recording of the functional brain images auditory stimuli were presented via MR-compatible headphones. Visual stimuli were projected on a semi-transparent screen by a video projector outside the scanner room. An MR-compatible response box was used for the participants to respond. Projection and timing of the stimuli as well as the recording of responses were controlled by Presentation software (Neurobehavioral Systems Inc., Albany, CA, USA).

Stimuli

For the orthographic and the semantic task of the present study 48 mono-syllabic German nouns were selected. While all words consisted of three phonemes, half of the words had three letters (e.g., HUT [hat]) and the remainder four (e.g., KNIE [knee]). In addition, half of the words referred to living entities (e.g., BUB [boy]) and half to nonliving entities (e.g., ZUG [train]). The average lexical frequency of the words was 88 per million ($SD = 188$) and the average summated bigram frequency was 18,983 ($SD = 16,492$). The large standard deviations of these measures resulted from some extremely high-frequency words [e.g., FRAU (woman) with 857 occurrences per million] and some words consisting of very frequent bigrams [e.g., TIER (animal) with a summated bigram frequency of 94,848]. The words had on

average 4.4 ($SD = 2.2$) orthographic and 12.7 ($SD = 6.3$) phonological neighbors. Both neighborhood size measures were calculated using CLEARPOND (Marian et al., 2012) using only substitutions (i.e., number of words of the same letter or phoneme length that differ in only one letter or phoneme). All item characteristics are based on the SUBTLEX database for German (Brysbaert et al., 2011). Word items were divided into 2 subsets that, across participants, were used about equally often in the orthographic and in the semantic task. The subsets were matched on all mentioned item characteristics. For the pitch evaluation task square-wave tones (100 or 300 Hz) were used.

Image acquisition and analysis

During each of two functional runs, 148 images sensitive to blood oxygenation level dependent (BOLD) contrast were acquired with a T2*-weighted echo-planar imaging sequence (echo time = 40 ms, TR = 2000 ms, 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 × 3.5 × 6 mm) and a high-resolution (1 × 1 × 1.3 mm) structural scan were acquired from each participant with T1-weighted MPRAGE sequences. A 1.5-T Intera Scanner (Philips Medical System Inc., Maasricht, The Netherlands) was used for magnetic resonance imaging.

For preprocessing and statistical analysis, SPM8 software was used (<http://www.fil.ion.ucl.ac.uk/spm>) running in a MATLAB 7.6 environment (Mathworks Inc., Natick MA, USA). Functional images were realigned, unwarped, and slice-time corrected. The high-resolution structural image was preprocessed and normalized using the VBM8 toolbox (<http://dbm.neuro.uni-jena.de/vbm8>). The image was segmented into gray matter, white matter and CSF, denoised, and warped into MNI space by registering it to the DARTEL template of the VBM8 toolbox using the high-dimensional DARTEL registration algorithm (Ashburner, 2007). Based on these steps, a skull-stripped version of the structural image was created in native space. The functional images were co-registered to the skull-stripped structural image and then the parameters from the DARTEL registration were used to normalize the functional images to the MNI space. The functional images were further resampled to isotropic 3 × 3 × 3 mm voxels and smoothed with an 8 mm FWHM (full width half maximum) Gaussian kernel.

Statistical analysis of the fMRI data was performed within a two-stage mixed effects model. For the participant-specific first-level models, the onsets of each task's trials were modeled by a canonical hemodynamic response function (HRF) resulting in three regressors (orthographic, semantic, pitch evaluation). Additionally, the onsets of block instructions and visual cues were modeled as regressors of no interest. This trial-based (event-related) analysis allowed us to best characterize positive and negative activations for auditory words and tones relative to rest (i.e., fixation baseline). An analysis based on the task blocks would have obscured these comparisons since the task blocks also contained visual input (instruction and cue screens) as well as motor responses. The first-level models also included six covariates corresponding to the movement parameters (rotations and translations). The functional data were high-pass filtered with a cut-off of 128 s and corrected for autocorrelation by an AR(1) model (Friston et al., 2002). In the first level models, the parameter estimates reflecting signal change for all conditions compared to rest were calculated in the context of a GLM (Henson, 2004). These participant-specific images were then used for the second-level random effects analysis. For statistical whole-brain comparisons we used a voxelwise threshold of $p < 0.05$, corrected for multiple comparisons using the family-wise error (FWE) rate, and a cluster extent threshold of 5 voxels. Anatomical descriptions for activation peaks are based on the probabilistic Harvard–Oxford Atlas (Desikan et al., 2006) as implemented in FSL (<http://www.fmrib.ox.ac.uk/fsl>), thresholded at 25%.

Results

Behavioral results

Due to technical problems during data acquisition five participants had to be excluded from the behavioral analyses. Table 1 shows the mean response times (measured from the onset of the auditory stimuli) and accuracies (percentage of correct trials) for the remaining 24 participants. Outliers, more than two standard deviations from the group mean, were excluded (in all tasks less than three outliers were identified). Between-task comparisons showed that there was no difference in response times and accuracy between the orthographic and the semantic task ($t_{(23)} < 1.49$, $p > .15$). However, both the orthographic and the semantic task led to prolonged response times relative to the pitch evaluation task ($t_{(23)} > 9.81$, $p < .001$). In addition, the semantic task led to more errors than the pitch evaluation task ($t_{(23)} = 3.80$, $p < .001$).

Additional within-task analyses showed that there was no difference between three and four letter responses in the orthographic task in neither response times ($t_{(23)} = 1.95$, $p = 0.15$) nor accuracy ($t_{(23)} = -1.57$, $p = 0.13$). In the semantic task, although there was no difference in response times ($t_{(23)} < 1$), accuracy was lower for living compared to nonliving entities ($t_{(23)} = -3.6$, $p < 0.005$). The latter difference, which was also responsible for the overall difference in accuracy between the semantic and the pitch evaluation task, resulted from some living items being frequently misclassified as nonliving (e.g., ABT [abbot]).

fMRI results

First, via whole-brain analyses we investigated activation differences between orthographic and semantic processing of auditory words, activations for auditory words in the two tasks relative to the tones, and activations for auditory words and tones relative to rest (positive and negative). In addition, we performed region-of-interest (ROI) analyses to visualize activation levels for all conditions relative to rest in left posterior ventral regions.

Orthographic versus semantic processing of auditory words

In a first step, we directly compared activation for orthographic and semantic processing of auditory words. As can be seen from the left column of Fig. 1, only few activation differences were observed. More activation for semantic processing was only observed in the precuneus with a peak at $[-3 -70 31]$ in MNI standard space (cluster size: 14 voxels; peak t value: 5.30). In the opposite direction, also only one cluster was identified. More activation for orthographic relative to semantic processing of auditory words was observed in the left vOT with a peak at $[-45 -64 -11]$ (cluster size: 9 voxels; peak t value: 6.31). Even with a more lenient threshold ($p < .001$, uncorrected) we did not identify any left vOT region exhibiting more activation for the semantic relative to the orthographic task.

Table 2 and the middle column of Fig. 1 further show activation for the auditory words in the orthographic and semantic tasks relative to

the tones in the pitch evaluation task. In line with the previous analysis, we primarily identified activation common to both orthographic and semantic processing. Regions activated for both tasks included bilateral superior and middle temporal gyri as well as the left inferior frontal gyrus, precentral gyrus, and parieto-occipital regions. Of specific interest were activations for the auditory words in left vOT. Here, we only identified activation for the orthographic task with a peak at $[-45 -64 -11]$. However, to more thoroughly investigate left vOT activations, we repeated the analyses with a more lenient threshold ($p < .001$, uncorrected) and restricted the search space to left vOT regions (i.e., left fusiform and inferior temporal gyrus). As can be seen in Fig. 1-B, this analysis revealed a small semantic activation cluster in middle vOT at $[-39 -58 -17]$ (cluster size: 5 voxels, peak t value: 3.86) overlapping with the large orthographic activation cluster (with the lower statistical threshold the size of the orthographic activation cluster increased to 66 voxels).

Auditory stimuli versus rest

The right column of Fig. 1 shows positive and negative activations for the auditory words and tones versus rest. Positive activations for the auditory stimuli were found in bilateral superior/middle temporal regions and left precentral gyrus. Interestingly, even with a more lenient threshold ($p < .001$) no positive activations were observed in left vOT (only in the ROI-based analysis, we did identify positive activations for the auditory words in these regions – see below). Marked deactivations in response to all auditory stimuli were observed in dorsal and ventral occipital regions. Of particular interest were deactivations in left ventral posterior regions (see Fig. 1-B). While all auditory stimuli led to deactivations in posterior parts, only the auditory tones did so in anterior parts including the left vOT (see Appendix A for statistics).

ROI-based analyses

We further investigated activation levels for auditory words and tones by means of ROI-based analyses. These analyses mainly served to supplement the whole-brain analysis by providing a clear visualization of activation levels relative to rest in left ventral posterior regions of interest. Three non-overlapping spherical ROIs ($r = 4$ mm) were defined along the posterior-to-anterior dimension and centered on maximum intensity voxels of an “effects of interest” contrast (i.e., all conditions versus rest) which was anatomically restricted to the left fusiform and inferior temporal gyrus. The anterior ROI was centered on the maximum intensity voxel between $y = -50$ and -59 , the middle ROI between $y = -60$ and -69 , and the posterior ROI between $y = -70$ and -90 . Fig. 2 depicts the approximate locations of ROIs. For all regions, mean contrast estimates were extracted for the three tasks versus rest. Outliers, more than two standard deviations from the group mean, were excluded (in all ROIs, less than three outliers were identified per task).

An initial TASK (orthographic, semantic, pitch evaluation) \times ROI (posterior, middle, anterior) ANOVA revealed a significant interaction between the factors ($F_{(4,112)} = 11.56$, $p < .001$) reflecting differentiated activation gradients for the three tasks across the ROIs. To investigate regional task differences, we additionally carried out one-way ANOVAs with the factor TASK for each ROI. These analyses showed that there was no difference between the tasks in the posterior ROI ($F_{(2,56)} < 1$). In contrast, in both the middle and the anterior ROI significant task differences were identified [middle ROI: $F_{(2,56)} = 28.81$, $p < .001$; anterior ROI: $F_{(2,56)} = 16.61$, $p < .001$]. Paired t -tests showed that higher activation for the orthographic compared to the pitch evaluation task was present in both ROIs ($p < .05$). In contrast, higher activation for the semantic compared to the pitch evaluation task was found in the anterior ($p < .05$) but not the middle ROI ($p = .21$). Direct comparisons of the auditory word tasks further revealed higher activation for the orthographic relative to the semantic task in both ROIs ($p < .05$).

We additionally performed comparisons of regional task activation levels relative to rest (see Table 3). These comparisons revealed that

Table 1

Behavioral results. Mean response times (measured from stimulus onset) and accuracy (percentage of correct trials) for the orthographic, the semantic, and the pitch evaluation task. Standard deviations are presented in parentheses.

Task	Response times (s)	Accuracy (% correct trials)
Orthographic (all responses)	3.06 (0.05)	93.9 (1.5)
Three letters	3.09 (0.05)	92.4 (2.0)
Four letters	3.03 (0.06)	95.5 (1.6)
Semantic (all responses)	3.06 (0.05)	91.3 (1.0)
Living	3.06 (0.05)	87.5 (1.9)
Nonliving	3.08 (0.06)	95.1 (1.1)
Pitch evaluation	2.26 (0.09)	96.4 (0.7)

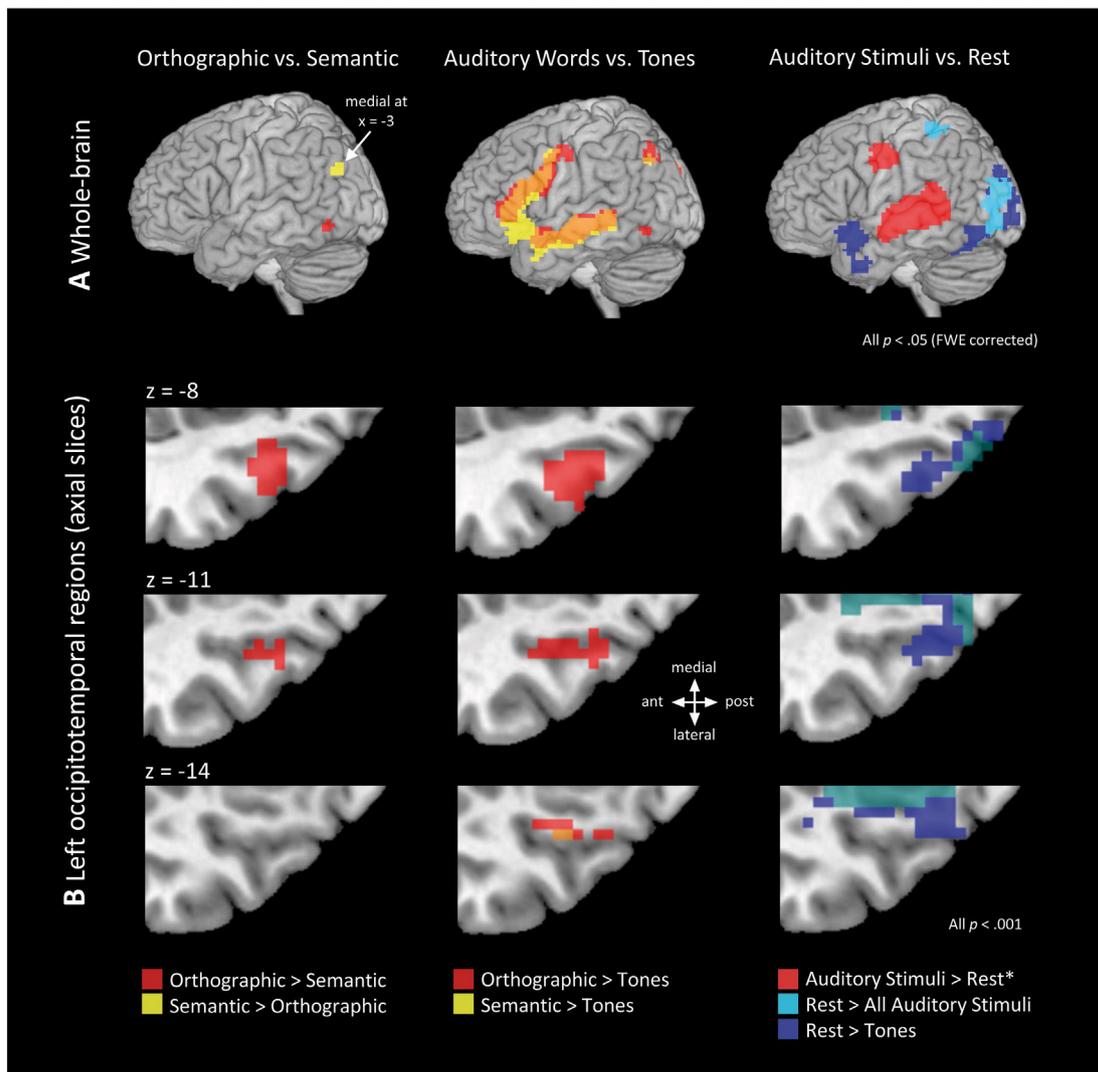


Fig. 1. fMRI results. Panel A shows the results of the whole-brain comparisons ($p < .05$, FWE corrected). Panel B shows the results for the same comparisons ($p < .001$, uncorrected) restricted to left ventral posterior regions (i.e., left fusiform and inferior temporal gyrus) superimposed on axial slices. The left column presents activation differences between orthographic and semantic processing of auditory words, the middle column presents activations for auditory words in the two tasks relative to the tones in the pitch evaluation task, and the right column presents activations and deactivations for the auditory stimuli relative to rest. *Brain regions activated for auditory words or tones relative to rest.

while the pitch evaluation task resulted in significant deactivations throughout all ROIs, more differentiated gradients were found for the orthographic and the semantic task. Activation for the orthographic task was negative in the posterior but positive in both the middle and the anterior ROIs. Activation for the semantic task was negative in both the posterior and the middle ROIs but did not differ from rest in the anterior ROI.

Discussion

The aim of the present fMRI study was to investigate whether the left vOT activation in response to auditory words can be attributed to lexical orthographic processing or instead to lexico-semantic processing. To this end, we presented auditory words in both an orthographic (“three or four letter word?”) and a semantic (“living or non-living?”) task. For orthographic decisions, an opaque phoneme–grapheme relation (all words had three phonemes but could have three or four letters) assured that participants had to access whole-word representations and did not solve the task sublexically (e.g., serially converting each phoneme into a grapheme). Our main result was that we found higher left vOT activation for orthographic relative to semantic processing of auditory words. Furthermore, comparing the auditory words to tones

presented in a pitch evaluation task revealed activation for orthographic processing in a large left vOT cluster. In contrast, activation for semantic processing was only small and restricted to middle vOT. In addition, the auditory tones elicited marked deactivations throughout left vOT regions.

The present result of higher left vOT activation for orthographic relative to semantic processing of auditory words strongly supports the assumption that activation in response to auditory words in this region reflects orthographic rather than lexico-semantic processing. In particular, in line with the conclusion of our previous spelling study (Ludersdorfer et al., 2015) we argue that left vOT activation can be attributed to the access to orthographic whole-word codes as suggested by previous reading-based findings (Kronbichler et al., 2004, 2007; Glezer et al., 2009). Critically, both the present study as well as Ludersdorfer et al. experimentally ascertained that participants had to rely on known word spellings (i.e., orthographic whole-word representations) to respond correctly in the orthographic spelling task. In general, the present results are also consistent with the assumption of an orthography-specific function of the left vOT by proponents of the VWFA hypothesis (Dehaene and Cohen, 2011) and speak against alternative accounts (Price and Devlin, 2011).

Table 2

Brain regions activated for auditory words in the orthographic and the semantic task relative to the tones in the pitch evaluation task ($p < .05$, FWE corrected). H = hemisphere, k = cluster extent in voxels.

Region	H	k	MNI coordinates			t
			x	y	z	
<i>Auditory words (orthographic task) > auditory tones</i>						
Inferior frontal gyrus (pTri)	L	360	-48	32	7	8.31
Precentral gyrus			-42	5	25	7.72
Inferior frontal gyrus (pOp)			-48	20	22	7.18
Middle/superior temporal gyrus	L	231	-60	-28	-5	7.80
Ventral occipitotemporal cortex	L	16	-45	-64	-11	6.75
Lateral occipital cortex	L	76	-27	-67	40	7.50
Cuneus/lateral occipital cortex	L	22	-12	-85	34	6.08
Middle/superior temporal gyrus	R	116	60	-22	-5	6.99
<i>Auditory words (semantic task) > auditory tones</i>						
Superior temporal gyrus	L	686	-60	-34	5	8.87
Inferior frontal gyrus (pOp)			-48	20	19	8.21
Lateral occipital cortex	L	20	-33	-67	40	5.51
Cuneus	L	42	-12	-82	28	6.16
Orbitofrontal cortex	R	29	36	26	-11	5.82
Superior temporal gyrus	R	132	60	-22	5	6.81
Calcarine cortex	R	50	9	-76	7	5.41

The identified activation cluster for orthographic processing (in the orthographic > semantic and the orthographic > tones contrast) with a peak at MNI coordinates $[-45 -64 -11]$ closely corresponds to the activation cluster found in our previous spelling study (Ludersdorfer et al., 2015) identified by contrasting orthographic spelling decisions on auditory words with a non-spelling control condition (i.e., gender decision) presenting the same words. A similar vOT region has also been consistently identified in previous spelling and writing studies (Rapp & Lipka, 2011; Rapp & Dufor, 2011). In a recent meta-analysis of such studies, Planton et al. (2013) identified a left vOT cluster with a peak at $[-46 -62 -12]$.

Interestingly, both the present orthographic activation peak as well as the left vOT peaks of previous spelling-based fMRI studies (see Planton et al., 2013) are located in a more posterior left vOT segment compared to what reading studies have identified as orthography-selective region or VWFA. The latter is classically located in middle vOT between $y = -50$ and -60 (e.g., Cohen et al., 2000, 2002; Glezer et al., 2009). However, one should note that although spelling activation peaks are usually located more posterior, activation clusters also reach in the more anterior parts of left vOT identified by reading studies. A possible explanation for the difference in peak location is that the more posterior peaks identified by spelling studies may not reflect the proper localization of abstract orthographic whole-word codes but the localization of explicit visual images of word-specific letter sequences. These letter sequences are derived from the more anteriorly located

orthographic whole-word codes and are maintained for the demands of orthographic tasks such as deciding whether the written form of an auditorily presented word consists of three or four letters. In reading studies, no such potentially effortful derivation and maintenance of letter sequences is required, because activation of the letter string is stimulus-driven and constitutes an only transient process on the way to whole-word recognition. This reasoning finds support in the results of our previous spelling study (Ludersdorfer et al., 2015) in which the left vOT activation for orthographic word spelling had a peak at $y = -64$ compared to a non-spelling control condition and a more anterior peak at $y = -55$ compared to pseudoword spelling. While the former contrast isolated spelling processes more generally, the latter specifically isolated the process of accessing orthographic word representations during spelling.

In this perspective it is of interest that in addition to the orthographic task also the semantic task elicited (weak) activation relative to the tones in left middle vOT. Activation of left vOT regions in response to auditory words presented in non-orthographic tasks (such as the present semantic task) is not an isolated finding. Several previous studies have identified activation in this region when participants listened to spoken words and had to make rhyming judgments (e.g., Booth et al., 2002; Yoncheva et al., 2010), repeat and think about the meaning of words (Price et al., 2003) or simply evaluate whether a word did or did not follow an identical one (Ludersdorfer et al., 2013). Initially, these findings were taken to challenge an orthography-specific function of left vOT (Price and Devlin, 2003). The present findings, however, do not support this view. In contrast to the strong orthographic activation, semantic activation (relative to the tones) observed in the whole-brain analysis was only present at a very lenient statistical threshold ($p < .001$, uncorrected). Furthermore, the ROI-based analyses showed that in the anterior ROI in which semantic activation was found, orthographic processing still elicited significantly higher activation relative to semantic processing. It is possible that the weak activation reflects that auditory words presented in the semantic task also automatically accessed their corresponding orthographic representations. Although speculative, this interpretation is in line with Dehaene et al. (2010, 2015) who argue that left vOT activation in response to spoken words in non-orthographic tasks could reflect automatic top-down recruitment of orthographic codes during demanding tasks in which all available information is gathered to facilitate speech processing. It is, however, also possible that orthographic and semantic activations in this middle vOT region stem from distinct sub-regions that are difficult to distinguish in group-based analyses. It may be possible that while the orthographic task engaged an orthographic region (i.e., the VWFA), the semantic task might have engaged a functionally distinct but overlapping region involved in multimodal word processing. Cohen et al. (2004), for example, identified a lateral inferior multimodal area (LIMA), which was indistinguishable from the VWFA in group comparisons but was reliably identified at the individual level. In future studies

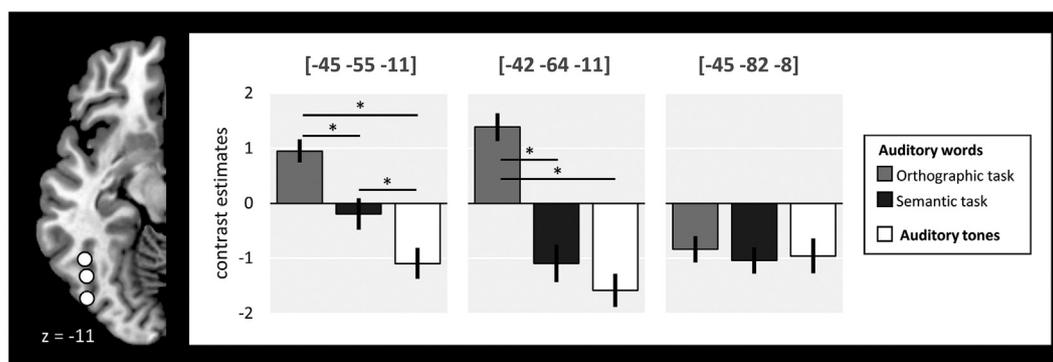


Fig. 2. Left ventral posterior regions of interest. On the left approximate locations of ROIs are shown. Activation plots on the right depict brain activation estimates (in arbitrary units) for auditory words in the orthographic and the semantic task as well as for the tones in the pitch evaluation task. Error bars denote ± 1 SEM. Asterisks denote significant differences ($p < .05$).

Table 3
Statistical comparisons (paired *t*-tests) against rest for auditory words and tones in each of the left ventral posterior ROIs.

Task	ROI		
	<i>y</i> = −55	<i>y</i> = −64	<i>y</i> = −82
Auditory words (orthographic task)	<i>t</i> (28) = 4.37, <i>p</i> < .001	<i>t</i> (28) = 5.35, <i>p</i> < .001	<i>t</i> (28) = −3.36, <i>p</i> < .005
Auditory words (semantic task)	<i>t</i> (28) = −0.66, <i>p</i> = .52 ns	<i>t</i> (28) = −3.20, <i>p</i> < .005	<i>t</i> (28) = −4.12, <i>p</i> < .001
Auditory tones	<i>t</i> (28) = −3.88, <i>p</i> < .001	<i>t</i> (28) = −5.27, <i>p</i> < .001	<i>t</i> (28) = −2.92, <i>p</i> < .01

it will be important to employ high-resolution MRI in combination with individual analyses in order to delineate fine-grained subdivisions of left vOT.

A further finding of interest was that the tones elicited marked deactivations in left vOT regions. According to the phenomenon of cross-modal suppression (Kawashima et al., 1995; Laurienti et al., 2002) deactivations in response to auditory processing are expected in visual brain regions. Therefore, the presently found deactivations in large dorsal and ventral occipital areas (see Fig. 1) are not surprising. In contrast, as mentioned in the Introduction, the assumption that the left vOT is a visual region and constitutes the anterior end of the left ventral visual stream (Dehaene et al., 2002) has been criticized. Instead, it has been suggested that the region might have a polymodal (Price and Devlin, 2003) or meta-modal function (Dehaene and Cohen, 2011). The latter was based on studies showing left vOT activation in congenitally blind individuals during reading of Braille (e.g., Reich et al., 2011). However, the presently found deactivations in left vOT in response to tones rather support a visual role of left vOT – at least in sighted individuals. More direct support for this comes from a previous study from our lab (Ludersdorfer et al., 2013) in which we showed that the left vOT exhibited marked deactivation to unfamiliar auditory stimuli together with a strong positive activation to unfamiliar visual stimuli. The characterization of the left vOT as dedicated to the visual domain has implications for our interpretation of the activations found in response to the auditory words. As mentioned, the higher activation for auditory words relative to the tones is taken to reflect access to orthographic representations. The deactivation to the tones suggests that these representations are shaped by visual experience (i.e., seeing visual words).

Generally unexpected was that activation differences between orthographic and semantic processing of auditory words were rather sparse with both eliciting largely common activation most prominently in left middle temporal and inferior frontal brain regions. A possible explanation for this is that the present study might have not cleanly separated semantic from orthographic processes so that word meaning may have not only been accessed during the semantic but also during the orthographic task. The present blocked task design, however, speaks against the possibility that this is merely an artifact of the experimental setup. Instead, this might point to a more general difficulty in separating semantic from orthographic processing of auditory words. The latter assumption finds support in some cognitive models of spelling which suggest that spellings for familiar words are not accessed directly from phonology but via the semantic system (e.g., Tainturier and Rapp, 2001).

Only one brain region, the left precuneus, was identified with higher activation for semantic compared to orthographic processing of auditory words. Interestingly, the precuneus is generally not associated with semantic processing per se. However, the region has been linked to mental imagery (Cavanna and Trimble, 2006), which poses a plausible strategy for the living/nonliving decision of the present semantic task. More surprising, however, was the failure to identify classical semantic regions such as the angular gyrus as well as the medial and anterior temporal lobes (Binder et al., 2009) in the comparison of the semantic with the auditory control condition. With respect to the angular gyrus,

recent evidence suggests that the region is primarily engaged by the integration of complex semantic information (Binder et al., 2009; Binder and Desai, 2011; Seghier, 2013). It might be the case that the present living/nonliving decisions did not require such high-level semantic integration processing. The anterior temporal lobe, which has also been ascribed a pivotal role for amodal semantic memory by neuropsychological and PET studies (Lambon Ralph et al., 2010) is in general rarely identified in fMRI studies of semantic processing (Patterson et al., 2007). This probably results from a diminished fMRI signal in brain regions close to the air-filled sinuses (Devlin et al., 2000). With respect to the medial temporal lobe, there is also mixed evidence for activation during semantic processing task in fMRI studies (Otten et al., 2001; Chee et al., 1999). It has been suggested that when semantic processing is compared to very simple control tasks with relatively long inter-stimulus intervals (such as the pitch evaluation task of the present study) activation in the medial temporal cortex might be missed due to “mind-wandering” in the control task which also activates these regions (Tielemans et al., 2005). However, of specific interest for semantic activations of the present study is a previous study by Booth et al. (2002), which also contrasted semantic processing of auditory words (i.e., a semantic association task) to processing of tones. Similar to the present findings, Booth and colleagues did not identify activations in the mentioned classic semantic regions but mainly in middle temporal and inferior frontal regions.

Conclusion

The present study provided evidence for the hypothesis that activation of the left vOT in response to auditory words reflects the access to visual-orthographic representations of whole words. We found higher left vOT activation for orthographic processing (“three or four letter word?”) relative to semantic processing (“living or nonliving?”) of auditory words. Comparisons to tones presented in a pitch evaluation task showed that orthographic processing of auditory words elicited activation throughout left vOT. In contrast, semantic processing elicited only weak activation in middle vOT. We interpret our findings as speaking for orthographic processing in left vOT. In particular, we suggest that activation in the middle vOT, the classic localization of the Visual Word Form Area, can be attributed to the access to orthographic whole-word representations. While activation of such representations was experimentally ascertained in the orthographic task, it may have also occurred automatically in the semantic task. Activation in the posterior vOT on the other hand may reflect the generation of explicit images of letter sequences required by the orthographic but not the semantic task. Based on the phenomenon of cross-modal suppression, the finding of marked deactivation in response to the auditory tones also supports the view that orthographic word representations in left vOT are of visual nature.

Acknowledgments

P. Ludersdorfer was supported by the Doctoral College “Imaging the Mind” of the Austrian Science Foundation (FWF-W1233), F. Richlan was supported by the Austrian Agency for International Cooperation in

Education and Research (OeAD PL 11/2015), and M. Kronbichler was supported by the Austrian Science Foundation (FWF P-23916-B18). We want to thank Carys Deeley for proofreading the manuscript.

Appendix A

Table A1

Brain regions exhibiting activation or deactivation in response to auditory words and tones relative to rest (voxelwise threshold: $p < .001$; cluster extent threshold: $p < .05$, FWE corrected). Abbreviations: H = hemisphere of cluster, k = cluster extent in voxels.

Region	H	k	MNI coordinates			t
			x	y	z	
Auditory words (orthographic task) > rest						
Superior/middle temporal gyrus	L	872	-51	-22	1	11.66
Superior/middle temporal gyrus	R	736	54	-22	4	11.51
Precentral gyrus	L	118	-51	-10	-43	6.18
Auditory words (semantic task) > rest						
Superior/middle temporal gyrus	L	875	-45	-37	7	11.89
Superior/middle temporal gyrus	R	767	54	-22	4	11.29
Precentral gyrus	L	72	-51	-10	-43	5.69
Auditory tones > rest						
Superior/middle temporal gyrus	R	656	54	-22	4	12.76
Superior/middle temporal gyrus	L	722	-45	-37	7	11.93
Rest > auditory words (orthographic task)						
Fusiform gyrus (extending to lingual gyrus)	R	263	21	-70	-11	6.13
Fusiform gyrus (extending to lateral occipital cortex, lingual gyrus)	L	504	-24	-58	-14	5.91
Lateral occipital cortex (extending to calcarine cortex)	L/R	634	36	-82	19	5.89
Superior parietal	L	61	-33	-43	58	4.5
Rest > auditory words (semantic task)						
Parahippocampal gyrus (extending to fusiform gyrus)	R	288	30	-34	-14	5.85
Fusiform gyrus	L	272	-27	-64	-14	5.53
Lateral occipital cortex	L	180	-43	-82	10	5.38
Supramarginal gyrus (extending to postcentral gyrus)	R	118	54	-25	46	5.31
Lateral occipital cortex	R	175	36	-82	19	5.21
Temporal pole	R	80	36	2	-35	4.9
Superior parietal lobule	L	93	-33	-43	58	4.89
Lateral occipital	R	73	12	-61	61	4.09
Rest > auditory tones						
Calcarine cortex (extending to lingual gyrus, fusiform gyrus, lateral occipital cortex)	L/R	3341	21	-70	-11	8.03
Temporal pole	R	131	36	5	-35	5.53
Orbitofrontal cortex	L	194	-12	14	-20	5.26
Anterior cingulate gyrus	L/R	160	0	26	28	5.22
Superior parietal lobule (extending to postcentral gyrus)	L	120	-27	-43	55	4.94
Thalamus	L/R	368	3	-16	10	4.87

References

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. *NeuroImage* 38 (1), 95–113.

Binder, J.R., Desai, R.H., 2011. The neurobiology of semantic memory. *Trends Cogn. Sci.* 15 (11), 527–536.

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 (2), 739–748.

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. *Cereb. Cortex* 19 (12), 2767–2796.

Booth, J.R., Burman, D.D., Meyer, J.R., Gitelman, D.R., Parrish, T.B., Mesulam, M.M., 2002. Modality independence of word comprehension. *Hum. Brain Mapp.* 16 (4), 251–261.

Bruno, J.L., Zumberge, A., Manis, F.R., Lu, Z.-L., Goldman, J.G., 2008. Sensitivity to orthographic familiarity in the occipito-temporal region. *NeuroImage* 39 (4), 1988–2001.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A.M., Bölte, J., Böhl, A., 2011. The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. *Exp. Psychol.* 58 (5), 412–424.

Büchel, C., Price, C., Friston, K., 1998. A multimodal language region in the ventral visual pathway. *Nature* 394, 14–17.

Cavanna, A.E., Trimble, M.R., 2006. The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* 129 (3), 564–583.

Chee, M.W., O'Craven, K.M., Bergida, R., Rosen, B.R., Savoy, R.L., 1999. Auditory and visual word processing studied with fMRI. *Hum. Brain Mapp.* 7 (1), 15–28.

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.A., Michel, F., 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 (2), 291–307.

Cohen, L., Lehéricy, 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 (5), 1054–1069.

Cohen, L., Martinaud, O., Lemer, C., Lehericy, S., Samson, Y., Obadia, M., Slachevski, A., Dehaene, S., 2003. Visual word recognition in the left and right hemispheres: anatomical and functional correlates of peripheral alexia. *Cereb. Cortex* 13 (12), 1313–1333.

Cohen, L., Jobert, A., Le Bihan, D., Dehaene, S., 2004. Distinct unimodal and multimodal regions for word processing in the left temporal cortex. *NeuroImage* 23 (4), 1256–1270.

Coltheart, M., 2004. Are there lexicons? *Q. J. Exp. Psychol. A Hum. Exp. Psychol.* 57 (7), 1153–1171.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., Ziegler, J., 2001. DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol. Rev.* 108 (1), 204–256.

Dehaene, S., Cohen, L., 2011. The unique role of the visual word form area in reading. *Trends Cogn. Sci.* 15 (6), 254–262.

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 (3), 321–325.

Dehaene, S., Pegado, F., Braga, L.W., Ventura, P., Nunes, Filho G., Jobert, A., ... Cohen, L., 2010. How learning to read changes the cortical networks for vision and language. *Science* 330 (6009), 1359–1364.

Dehaene, S., Cohen, L., Morais, J., Kolinsky, R., 2015. Illiterate to literate: behavioural and cerebral changes induced by reading acquisition. *Nat. Rev. Neurosci.* 16 (4), 234–244.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* 31 (3), 968–980.

Devlin, J.T., Russell, R.P., Davis, M.H., Price, C.J., Wilson, J., Moss, H.E., Matthews, P.M., Tyler, L.K., 2000. Susceptibility-induced loss of signal: comparing PET and fMRI on a semantic task. *NeuroImage* 11 (6), 589–600.

Devlin, J., Jamison, H., Gonnerman, L., Matthews, P., 2006. The role of the posterior fusiform gyrus in reading. *J. Cogn. Neurosci.* 18 (6), 911–922.

Duncan, K.J., Pattamadilok, C., Devlin, J.T., 2010. Investigating occipito-temporal contributions to reading with TMS. *J. Cogn. Neurosci.* 22 (4), 739–750.

Fiebach, C.J., Friederici, A.D., Müller, K., von Cramon, D.Y., 2002. fMRI evidence for dual routes to the mental lexicon in visual word recognition. *J. Cogn. Neurosci.* 14 (1), 11–23.

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 (1), 205–218.

Friston, K.J., Glaser, D.E., Henson, R.N., Kiebel, S., Phillips, C., Ashburner, J., 2002. Classical and Bayesian inference in neuroimaging: applications. *NeuroImage* 16 (2), 484–512.

Gaillard, R., Naccache, L., Pinel, P., Clémenceau, S., Volle, E., Hasboun, D., Dupont, S., Baulac, M., Dehaene, S., Adam, C., Cohen, L., 2006. Direct intracranial, fMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron* 50 (2), 191–204.

Glezer, L.S., Jiang, X., Riesenhuber, M., 2009. Evidence for highly selective neuronal tuning to whole words in the "visual word form area". *Neuron* 62 (2), 199–204.

Graves, W.W., Desai, R., Humphries, C., Seidenberg, M.S., Binder, J.R., 2010. Neural systems for reading aloud: a multiparametric approach. *Cereb. Cortex* 20 (8), 1799–1815.

Hauk, O., Davis, M.H., Pulvermüller, F., 2008a. Modulation of brain activity by multiple lexical and word form variables in visual word recognition: a parametric fMRI study. *NeuroImage* 42 (3), 1185–1195.

Hauk, O., Davis, M.H., Kherif, F., Pulvermüller, F., 2008b. Imagery or meaning? Evidence for a semantic origin of category-specific brain activity in metabolic imaging. *Eur. J. Neurosci.* 27 (7), 1856–1866.

Henson, R.N.A., 2004. Analysis of fMRI time series: linear time-invariant models, event-related fMRI and optimal experimental design. In: Frackowiak, R.S.J., Friston, K.J., Frith, C., Dolan, R., Price, C.J., Zeki, S., Ashburner, J.T., Penny, W.D. (Eds.), *Human Brain Function*. Academic Press, London, UK, pp. 793–822.

Jobard, G., Crivello, F., Tzourio-Mazoyer, N., 2003. Evaluation of the dual route theory of reading: a metanalysis of 35 neuroimaging studies. *NeuroImage* 20 (2), 693–712.

Kawashima, R., O'Sullivan, B.T., Roland, P.E., 1995. Positron-emission tomography studies of cross-modality inhibition in selective attentional tasks: closing the "mind's eye". *Proc. Natl. Acad. Sci.* 92 (13), 5969–5972.

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 (3), 946–953.

Kronbichler, M., Bergmann, J., Hutzler, F., Staffen, W., Mair, A., Ladurner, G., Wimmer, H., 2007. Taxi vs. taks: on orthographic word recognition in the left ventral occipitotemporal cortex. *J. Cogn. Neurosci.* 19 (10), 1584–1594.

Kronbichler, M., Klackl, J., Richlan, F., Schurz, M., Staffen, W., Ladurner, G., Wimmer, H., 2009. On the functional neuroanatomy of visual word processing: effects of case and letter deviance. *J. Cogn. Neurosci.* 21 (2), 222–229.

Lambert Ralph, M.A., Sage, K., Jones, R.W., Mayberry, E.J., 2010. Coherent concepts are computed in the anterior temporal lobes. *Proc. Natl. Acad. Sci. U. S. A.* 107 (6), 2717–2722.

Laurienti, P.J., Burdette, J.H., Wallace, M.T., Yen, Y.-F., Field, A.S., Stein, B.E., 2002. Deactivation of sensory-specific cortex by cross-modal stimuli. *J. Cogn. Neurosci.* 14 (3), 420–429.

- Leff, A.P., Crewes, H., Plant, G.T., Scott, S.K., Kennard, C., Wise, R.J.S., 2001. The functional anatomy of single-word reading in patients with hemianopic and pure alexia. *Brain* 124 (3), 510–521.
- Ludersdorfer, P., Schurz, M., Richlan, F., Kronbichler, M., Wimmer, H., 2013. Opposite effects of visual and auditory word-likeness on activity in the visual word form area. *Front. Hum. Neurosci.* 7, 491.
- Ludersdorfer, P., Kronbichler, M., Wimmer, H., 2015. Accessing orthographic representations from speech: the role of left ventral occipitotemporal cortex in spelling. *Hum. Brain Mapp.* 36, 1393–1406.
- Marian, V., Bartolotti, J., Chabal, S., Shook, A., 2012. CLEARPOND: cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PLoS One* 7 (8), e43230.
- Mechelli, A., Gorno-Tempini, M., Price, C., 2003. Neuroimaging studies of word and pseudoword reading: consistencies, inconsistencies, and limitations. *J. Cogn. Neurosci.* 15 (2), 260–271.
- Otten, L.J., Henson, R.N., Rugg, M.D., 2001. Depth of processing effects on neural correlates of memory encoding. *Brain* 124 (2), 399–412.
- Patterson, K., Nestor, P.J., Rogers, T.T., 2007. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat. Rev. Neurosci.* 8 (12), 976–987.
- Planton, S., Jucla, M., Roux, F.E., Démonet, J.F., 2013. The “handwriting brain”: a meta-analysis of neuroimaging studies of motor versus orthographic processes. *Cortex* 49 (10), 2772–2787.
- Price, C.J., Devlin, J.T., 2003. The myth of the visual word form area. *NeuroImage* 19 (3), 473–481.
- Price, C.J., Devlin, J.T., 2004. The pro and cons of labelling a left occipitotemporal region: “the visual word form area”. *NeuroImage* 22 (1), 477–479.
- Price, C.J., Devlin, J.T., 2011. The interactive account of ventral occipitotemporal contributions to reading. *Trends Cogn. Sci.* 15 (6), 246–253.
- Price, C.J., Winterburn, D., Giraud, A.L., Moore, C.J., Noppeney, U., 2003. Cortical localisation of the visual and auditory word form areas: a reconsideration of the evidence. *Brain Lang.* 86, 272–286.
- Purcell, J.J., Turkeltaub, P.E., Eden, G.F., Rapp, B., 2011. Examining the central and peripheral processes of written word production through meta-analysis. *Front. Psychol.* 2, 239.
- Rapp, B., Dufor, O., 2011. The neurotopography of written word production: an fMRI investigation of the distribution of sensitivity to length and frequency. *J. Cogn. Neurosci.* 23 (12), 4067–4081.
- Rapp, B., Lipka, K., 2011. The literate brain: the relationship between spelling and reading. *J. Cogn. Neurosci.* 23 (5), 1180–1197.
- Reich, L., Szwed, M., Cohen, L., Amedi, A., 2011. A ventral visual stream reading center independent of visual experience. *Curr. Biol.* 21 (5), 363–368.
- Sabsevitz, D.S., Medler, D.A., Seidenberg, M., Binder, J.R., 2005. Modulation of the semantic system by word imageability. *NeuroImage* 27 (1), 188–200.
- 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 (3), 2649–2661.
- Seghier, M.L., 2013. The angular gyrus: multiple functions and multiple subdivisions. *Neuroscientist* 19 (1), 43–61.
- Tainturier, M.J., Rapp, B., 2001. The spelling process. In: Rapp, B. (Ed.), *The Handbook of Cognitive Neuropsychology: What Deficits Reveal About the Human Mind*. Psychology Press, Philadelphia, PA, pp. 263–289.
- Tieleman, A., Seurinck, R., Deblaere, K., Vandemaele, P., Vingerhoets, G., Achten, E., 2005. Stimulus pacing affects the activation of the medial temporal lobe during a semantic classification task: an fMRI study. *NeuroImage* 26 (2), 565–572.
- Turkeltaub, P.E., Eden, G.F., Jones, K.M., Zeffiro, T.A., 2002. Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *NeuroImage* 16 (3), 765–780.
- Twomey, T., Kawabata, K.J., Price, C.J., Devlin, J.T., 2011. Top-down modulation of ventral occipito-temporal responses during visual word recognition. *NeuroImage* 55 (3), 1242–1251.
- Van der Mark, S., Bucher, K., Maurer, U., Schulz, E., Brem, S., Buckelmüller, J., ... Brandeis, D., 2009. Children with dyslexia lack multiple specializations along the visual word-form (VWF) system. *NeuroImage* 47 (4), 1940–1949.
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J.P., Sigman, M., Cohen, L., 2007. Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron* 55 (1), 143–156.
- Wheatley, T., Weisberg, J., Beauchamp, M.S., Martin, A., 2005. Automatic priming of semantically related words reduces activity in the fusiform gyrus. *J. Cogn. Neurosci.* 17 (12), 1871–1885.
- Wise, R.J.S., Howard, D., Mummery, C.J., Fletcher, P., Leff, A., Büchel, C., Scott, S.K., 2000. Noun imageability and the temporal lobes. *Neuropsychologia* 38 (7), 985–994.
- 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. *J. Cogn. Neurosci.* 23 (4), 992–1002.
- Yoncheva, Y.N., Zevin, J.D., Maurer, U., McCandliss, B.D., 2010. Auditory selective attention to speech modulates activity in the visual word form area. *Cereb. Cortex* 20 (3), 622–632.