

# Adding Words to the Brain's Visual Dictionary: Novel Word Learning Selectively Sharpens Orthographic Representations in the VWFA

Laurie S. Glezer, Judy Kim, Josh Rule, Xiong Jiang, and Maximilian Riesenhuber

Department of Neuroscience, Georgetown University Medical Center, Washington, District of Columbia 20007

The nature of orthographic representations in the human brain is still subject of much debate. Recent reports have claimed that the visual word form area (VWFA) in left occipitotemporal cortex contains an orthographic lexicon based on neuronal representations highly selective for individual written real words (RWs). This theory predicts that learning novel words should selectively increase neural specificity for these words in the VWFA. We trained subjects to recognize novel pseudowords (PWs) and used fMRI rapid adaptation to compare neural selectivity with RWs, untrained PWs (UTPWs), and trained PWs (TPWs). Before training, PWs elicited broadly tuned responses, whereas responses to RWs indicated tight tuning. After training, TPW responses resembled those of RWs, whereas UTPWs continued to show broad tuning. This change in selectivity was specific to the VWFA. Therefore, word learning appears to selectively increase neuronal specificity for the new words in the VWFA, thereby adding these words to the brain's visual dictionary.

**Key words:** language; learning; object recognition; plasticity; reading; VWFA

## Introduction

Numerous neuroimaging studies have provided evidence that reading engages the left ventral occipitotemporal cortex (Price and Devlin, 2003; Cohen and Dehaene, 2004; Baker et al., 2007; Ben-Shachar et al., 2007; Szwed et al., 2011) and that learning to read changes the activity in this region (Ben-Shachar et al., 2007, 2011; Brem et al., 2010; Dehaene et al., 2010), leading to the development of a “visual word form area” (VWFA). To better understand how new words are incorporated into existing neural representations, several studies have explored plasticity in the brain's reading system (Sandak et al., 2004; Callan et al., 2005; Xue et al., 2006; Fisher et al., 2012). Results have been mixed—whereas some studies show an increase in activation in the VWFA with learning (Xue et al., 2006), others report a decrease (Sandak et al., 2004; Xue and Poldrack, 2007) and others find no training effects in the VWFA (Callan et al., 2005; Fisher et al., 2012). However, inferring neuronal selectivity based on average BOLD-contrast signal change is complicated by the fact that both the density of selective neurons and the broadness of their tuning contribute to the average activity level in a voxel (Jiang et al., 2006): A given BOLD-contrast signal change in a voxel could

arise from a small number of unselective neurons that all respond broadly to a large set of stimuli or from a larger population of more selective neurons, of which only subsets (those with preferred stimuli that most resemble the current stimulus) respond to any given stimulus. In contrast, it has been suggested that fMRI rapid adaptation (fMRI-RA) techniques can probe neuronal tuning more directly and selectively (for review, see Grill-Spector et al., 2006). Given two sequentially presented stimuli, the BOLD-contrast response to the pair is taken to reflect similarity of the neuronal activation patterns corresponding to the two individual stimuli. Therefore, the lowest response reflects activation of identical neuronal populations and maximum signal indicates activation of disjoint groups of neurons for the two stimuli (Jiang et al., 2006).

Recently, we used fMRI-RA to examine the nature of the representation in the VWFA by systematically altering the visual word form and lexicality between pairs of stimuli to examine the effect of similarity on the hemodynamic response (Glezer et al., 2009). Our results demonstrated that the VWFA showed high selectivity for whole real words (RWs), yet broader tuning to orthographically legal nonwords (pseudowords, PWs). These results are compatible with the notion that extensive visual experience with RWs (and the need to discriminate between these words), but not unfamiliar PWs, results in experience-driven refinement of neuronal tuning. Although this interpretation is plausible, a more direct test of this visual dictionary theory would be to demonstrate that learning to visually recognize novel words induces RW-like selectivity profiles specific to those words after learning. We tested this hypothesis here by training subjects to recognize novel PWs and scanned before and after learning to probe selectivity in the VWFA to RWs, untrained PWs (UTPWs), and trained PWs (TPWs).

Received Sept. 29, 2014; revised Feb. 7, 2015; accepted Feb. 13, 2015.

Author contributions: L.S.G., J.K., J.R., X.J., and M.R. designed research; L.S.G., J.K., J.R., and X.J. performed research; L.S.G., J.K., and X.J. analyzed data; L.S.G., X.J., and M.R. wrote the paper.

This work was supported by the National Science Foundation (Grant 1026934 to M.R.) and the Intellectual and Developmental Disorders Research Center (Grant 5P30HD040677-13).

The authors declare no competing financial interests.

Correspondence should be addressed to Maximilian Riesenhuber, Department of Neuroscience, Georgetown University Medical Center, Research Building Room WP-12, 3970 Reservoir Rd. NW, Washington, DC 20007A. E-mail: mr287@georgetown.edu.

DOI:10.1523/JNEUROSCI.4031-14.2015

Copyright © 2015 the authors 0270-6474/15/354965-08\$15.00/0

## Materials and Methods

**Participants.** A total of 25 right-handed healthy adults who were native English speakers (aged 18–35) were enrolled in the experiment. Subjects were excluded from further analysis if a VWFA could not be identified ( $n = 3$ ), if they had excessive head motion or fell asleep during the scan ( $n = 2$ ), if they withdrew from the study before it was completed ( $n = 5$ ), if they scored <80% accuracy on the posttraining recognition test ( $n = 1$ ), or never reached the 80% training criterion on the discrimination test after 8 sessions ( $n = 1$ ). In addition, one subject needed to be excluded due to image acquisition problems during the postscan. After exclusion of subjects for the aforementioned reasons, a total of 12 subjects (9 female, age 18–27 years) took part in the experiment. Experimental procedures were approved by Georgetown University's Institutional Review Board and written informed consent was obtained from all subjects before the experiment.

**Cognitive testing.** Two subtests from the Woodcock-Johnson III Diagnostic Reading Battery, the Word ID, and Word Attack tests were used to assess reading ability and subjects were screened to have at least average scores (0.5 SD below the mean or 92) on both tests. In addition, all subjects scored 80 or above on the Edinburgh Handedness Inventory.

**Stimuli.** For the first scan, the RW stimuli from Experiments 1 and 3 from Glezer et al. (2009) were used. RW stimuli were chosen using the CELEX database (Baayen et al., 1995). Forty-seven high-frequency (>50 per million) 3–6 letter target words were chosen. Words were presented in pairs and we examined three conditions: (1) SAME, in which the same stimulus was presented twice (as first and second stimulus) in each trial; (2) 1L, in which the first and second stimulus differed by one letter and; (3) DIFF, in which the second stimulus shared no letters with the first. RWs were matched for length, bigram and trigram (where applicable) frequency, and neighborhood size. For scans 2 and 3, we used MCWord (Medler and Binder, 2005) to generate three lists of 150 PWs, 3–6 letters in length. One set of PWs was used for the training set and two sets of PWs were untrained and used to probe for PW selectivity in the pretraining and posttraining scans, respectively, resulting in three sets of 50 PW triplets (SAME, 1L and DIFF, as in Glezer et al., 2009) matched for length, bigram and trigram frequency, and orthographic neighborhood.

**Training.** Building on behavioral studies showing that people can learn novel words rapidly and in the absence of semantic information (Salasoo et al., 1985; Gaskell and Dumay, 2003; Bowers et al., 2005; Chalmers and Burt, 2008) and retain this learning for up to 8 months (Tamminen and Gaskell, 2008), we trained subjects implicitly on novel words to minimize the potential for semantic confounds. Each subject participated in approximately 1 h of implicit word training per session in which they were exposed to each TPW 7 times during training and 1 time during testing. On average, subjects participated in 9 training sessions over 19 d. Because subjects trained until they reached criterion (see below), there was a range in the number of sessions that subjects participated in (range: 3–15 sessions, 24–120 exposures, 7–36 d). During training, each subject performed a two-back task with the TPWs. Stimuli were presented for 500 ms with an intertrial interval of 2500 ms.

**Word discrimination testing.** To measure progress in learning words during training, after each training session, subjects participated in a discrimination test. Subjects were presented with all of the TPWs and one letter different “foils” and asked to determine which word was familiar. Three sets of 1L “foils” were created, in which the letter change occurred in the same letter position as the letter change from SAME (e.g., the SAME and 1L TPW “soat” and “poat” differ in the initial position, the foils then would include the final 3 letters—oat with a different letter in the initial position, e.g., loat and joat). Before participating in the postscan fMRI, subjects needed to perform this task with at least 80% accuracy over two sessions.

**Testing of recognition performance pretraining and posttraining.** To obtain a behavioral measure of subjects' overall learning of novel words, we conducted a pretraining and posttraining recognition test. Before training, subjects were presented with the 150 TPWs and 150 RWs; after training, they were presented with 150 RWs, 150 TPWs, and 300 UTPWs. To determine whether subjects had learned any words during the initial scanning session, a portion of the subjects ( $n = 7$ ) were also tested on

novel PWs before training. For these tests, the subjects were asked to determine whether the stimulus was a “familiar” or “unfamiliar” letter string.

**Pretraining and posttraining rapid event-related scans 1, 2, and 3.** To probe the effects of varying orthographic similarity and visual experience on BOLD-contrast response in the VWFA, MRI images from three (Scan 1) and six (Scans 2 and 3) event-related (ER) runs were collected. For Scan 1, each run lasted 530.4 s and had two 10.2 s fixation periods, one at the beginning and the other at the end. Between the two fixation periods, a total of 130 trials were presented to participants at a rate of one every 4080 ms. For Scan 2 and 3, each run lasted 477.36 s and had two 10.2 s fixation periods one at the beginning and the other at the end. Between the two fixation periods, a total of 117 trials were presented to participants at a rate of one every 4080 ms. During each trial, for all three scans, the two stimuli in the pair were displayed sequentially (timing: first stimulus for 300 ms, blank for 400 ms, second stimulus for 300 ms, and blank for 3080 ms). For all three scans, the number of repetitions of each word stimulus was counterbalanced across all conditions to control for long-lag priming effects (Henson et al., 2000). Trial order and timing was adjusted using M-sequences (Buras and Boynton, 2002). To engage subjects' attention yet avoid potential task-related confounding modulations of the BOLD-contrast response to the conditions of interest (Grady et al., 1996; Sunaert et al., 2000), subjects were asked to perform an “oddball” detection task (Jiang et al., 2006; Glezer et al., 2009) in the scanner. Subjects were asked to press a button (with their right hand) every time they saw the sequential letters “xyz” or “abc.” These “oddball” stimuli were created by randomly replacing three sequential letters at the beginning, middle, or end of either the first or second stimulus in both the TPW and UTPW pairs. All stimuli were rendered in Courier font (36 point size, 100 ppi), average letter size  $\frac{1}{4} \times \frac{1}{4}$  inch ( $25 \times 25$  pixels), for an approximate size of 0.5 degrees of visual angle per letter in the scanner.

**Functional localizer scans.** At the end of Scan 2, separate localizer scans were conducted to identify the VWFA in each subject individually as described previously (Glezer et al., 2009; Glezer and Riesenhuber, 2013). Using a block design, echoplanar images (EPIs) from two functional localizer scans were collected. Subjects passively viewed blocks of images of written words (high-frequency nouns, >50 per million), faces, objects, and scrambled words and objects. Each block lasted 20400 ms (stimuli were displayed for 500 ms and were separated by a 100 ms blank interval) and stimulus blocks were separated by a 10200 ms fixation block. Each run consisted of two blocks of each stimulus group and 10 fixation blocks. The face and object images used in the localizer scans were obtained from a commercial database. They were postprocessed using programs written in MATLAB (The Mathworks) to eliminate background variations and to adjust image size, luminance, and contrast. The final size of all images was scaled to  $200 \times 200$  pixels. Word stimuli were chosen using the CELEX database (Baayen, et al. 1995). Scrambled images of words were generated by scrambling the word images with a tile size of 4 pixels. Scrambled images of the objects were generated by scrambling the objects with a tile size of 10 pixels.

The stimuli of both localizer and ER scans were presented using E-Prime software (<http://www.pstnet.com/products/e-prime/>), back-projected on a translucent screen located at the rear of the scanner, and viewed by participants through a mirror mounted on the head coil.

**MRI acquisition.** All MRI data were acquired at Georgetown University's Center for Functional and Molecular Imaging using an EPI sequence on a 3 tesla Siemens Trio scanner. A 12-channel head coil was used (flip angle = 90°, TR = 2040 ms, TE = 29 ms, FOV = 205 mm,  $64 \times 64$  matrix). Thirty-five interleaved axial slices (thickness = 4.0 mm, no gap; in-plane resolution =  $3.2 \times 3.2$  mm<sup>2</sup>) were acquired. At the end of Scan 1, a 3D T1-weighted MPRAGE images (resolution  $1 \times 1 \times 1$  mm<sup>3</sup>) were acquired from each subject.

**MRI data analysis.** All preprocessing and most statistical analyses were done using the SPM2 software package (<http://www.fil.ion.ucl.ac.uk/spm/software/spm2/>). After discarding the first five acquisitions of each run, the EPI images were temporally corrected to the middle slice (for event-related scans only), spatially realigned, resliced to  $2 \times 2 \times 2$  mm<sup>3</sup>, and normalized to a standard MNI reference brain in Talairach space. For the localizer and the group analyses, images were then

smoothed with an isotropic 6 mm Gaussian kernel. The VWFA regions were identified for each individual subject independently with the data from the localizer scans as described previously (Glezer et al., 2009; Glezer and Riesenhuber, 2013). We first modeled the hemodynamic activity for each condition (word, scrambled word, face, scrambled object, object, and fixation) in the localizer scans with the standard canonical hemodynamic response function and then identified a word-selective ROI with the contrast of word versus fixation ( $p < 0.00001$ , uncorrected) masked by the contrast of word versus scrambled word ( $p < 0.05$ , uncorrected). This contrast typically resulted in only 1–2 foci in the left ventral occipitotemporal cortex ( $p < 0.05$ , corrected). ROIs were selected by identifying in each subject the most anterior cluster that was significant at the corrected cluster level of at least  $p < 0.05$  in the ventral occipitotemporal cortex (specifically, the occipitotemporal sulcus/fusiform gyrus region) in a location closest to the published location of the VWFA, approximate Talairach coordinates  $-43 -54 -12 \pm 5$  (Cohen and Dehaene, 2004). Based on reports showing that the VWFA is quite small (average size 45 voxels; Baker et al., 2007) and to select ROIs that were of equivalent size (Murray and Wojciulik, 2004; Jiang et al., 2006, 2007), the thresholds were adjusted beyond this point to obtain a cluster that was between 20 and 50 voxels.

In the ER scans, after removing low-frequency temporal noise from the EPI runs with a high-pass filter (1/128 Hz), fMRI responses were modeled with a design matrix comprising the onset of trial types and movement parameters as regressors using a standard canonical hemodynamic response function. Proportional scaling was then applied to remove the effects of global variations (Aguirre et al., 1998) because, based on previous findings (Glezer et al., 2009), we expected differences between the conditions of interest to be small and limited to local VWFA regions. We then extracted the mean percentage signal change of the VWFA ROI for each subject with the MarsBar toolbox (Brett et al., 2002) and conducted statistical analyses within-subject repeated-measures ANOVA with Greenhouse–Geisser correction, followed by planned  $t$ -tests,  $\alpha = 0.05$ , two-tailed) on the percentage signal change. Given the variability in overall activation levels across subjects, we present results using within-subject SEM error bars (Loftus and Masson, 1994; also see our previous studies: Jiang et al., 2006, 2007; Glezer et al., 2009). These numbers were derived by first normalizing the data from each individual subject ( $f$ ) to the mean of subjects and conditions as follows:

$$f' = f_j - \frac{1}{N} \sum_{i=1}^N f_j^i + \frac{1}{N \times M} \sum_{i=1}^N \sum_{j=1}^M f_j^i$$

$i = 1 \dots N$  (condition),  $j = 1 \dots M$  (subjects)

and then calculated the within-subject SEM using the normalized data  $f'$ .

**ROI selection and identification.** To determine the specificity of the learning effect, we identified the following ROIs in the ventral visual stream: the VWFA in the right hemisphere (rVWFA), left and right fusiform face area (FFA), and left and right lateral occipital complex (LOC). The rVWFA was localized using the same methods as above for the left hemisphere, with the resulting ROI sizes falling in the same range as in the left. In addition, the location of the rVWFA was selected by choosing a peak coordinate from this contrast that was in a homologous location in the right hemisphere. Because using the same threshold criteria yielded only a few subjects with an rVWFA ROI, we lowered the threshold criteria and selected ROI that were significant at the voxel level ( $p < 0.05$ , FDR). Using these criteria, we were able to identify an rVWFA in six subjects. The FFA was localized using the contrast of faces > objects ( $p < 0.0001$ , uncorrected,  $n = 12$ ). The FFA in the left hemisphere was selected by choosing a peak coordinate from the same contrast that was within the homologous location as the individual subject's FFA in the right hemisphere ( $n = 9$ ). The LOC was identified in the left and right hemisphere using the contrast of objects > scrambled objects ( $p < 0.0001$ , uncorrected,  $n = 12$  and 11, respectively). ROI size for FFA and LOC was between 10 and 100 voxels.

**Whole-brain group analysis.** We conducted two whole-brain analyses to examine areas of the brain that were sensitive to training effects of orthographic stimuli. In the first analysis, the DIFF > 1L smoothed contrast images (6 mm) for each subject were entered into a second-level

paired  $t$  test for TPW-pre and TPW-post, after masking by words > fixation ( $D > 1L$ ,  $p < 0.005$ , uncorrected and cluster extent of 20 voxels, words > fixation,  $p < 0.0001$ , cluster extent 20 voxels). In the second analysis, the smoothed contrast images from TPW-post DIFF > 1L and 1L > SAME ( $p < 0.005$ , uncorrected, cluster extent 20 voxels) for each subject were entered into a second-level paired  $t$  test after masking by words > fixation ( $p < 0.0001$ , cluster extent 20 voxels). We then used the MarsBar Toolbox (Brett et al., 2002) to transform the ROIs obtained for both of these contrasts, identified all voxels the two ROIs had in common, and then extracted percentage signal change in this shared ROI in the pretraining and posttraining scans.

**Hierarchical analysis.** To examine activation along the word-selective hierarchy (Vinckier et al., 2007), we identified four word-selective ROIs along the ventral visual pathway using the same contrast we used for selecting the VWFA ( $p < 0.05$ , corrected at the cluster level). We were able to identify all four ROIs in 10/12 subjects. These ROI were grouped based on the location of the  $x$ ,  $y$ , and  $z$  coordinates of the peak voxel. Because the posterior ROIs were difficult to obtain as a cluster with only 1 focus, after identifying the peak coordinate for each ROI, a sphere of a fixed size was built (radius = 5 mm) centered at the peak of each of the aforementioned ROI for each subject (van der Mark et al., 2009). For each subject, we then analyzed the activity in their four individually defined ROIs during the separate event-related scans that used the rapid adaptation paradigm.

To examine the degree of sublexical and lexical processing along the hierarchy, we computed an index to measure degree of lexicality, the “lexicality index”  $L$  (Glezer, 2009). This metric takes into consideration the statistical significance between the comparisons of interest: SAME versus 1L, SAME versus DIFF, and 1L versus DIFF as follows:

$$L = \frac{\log p_{\text{SAME-1L}} + \log p_{\text{SAME-DIFF}}}{\log p_{\text{1L-DIFF}}}$$

The index gives a measure of the statistical difference between the comparison of SAME and 1L, as well as SAME and DIFF response amplitudes (which in the case of lexical processing will be high because SAME and DIFF should be significantly different) relative to the statistical difference between 1L and DIFF response amplitudes (which in the case of lexical processing will be low because 1L and DIFF should not differ statistically for a lexical representation). The formula contains the log of the  $p$ -values instead of the raw  $p$ -values to capture the logarithmic nature of the  $p$ -values as indicators of significance [e.g., differences in  $p$ -values of 0.90000001 versus 0.91 (log values of  $-0.045757486$  vs  $-0.045757491$ ) do not matter much, whereas the same numerical differences between  $p$ -values of 0.00000001 and 0.01 do (log values of  $-8$  vs  $-2$ )]. This produces an index that is high for word-like representations and low for unselective representations. In addition, the index's grounding in the significance of signal differences between pairs of conditions rather than absolute signal levels makes it more robust against confounds by overall amplitude differences unrelated to word selectivity, for example, due to stimulus novelty.

Therefore, the  $L$  index is high in an area with a lexical response profile and low in an area exhibiting a sublexical response profile. We separately computed  $L$  for RW and PW stimuli for all four ROIs along the ventral stream in the left hemisphere. To determine whether differences in the  $L$  index between stimulus groups were significant, we conducted bootstrap analyses. Data for the different stimulus groups (RW and TPW and UTPW and TPW, respectively) were randomly shuffled to produce 10,000 pairs of  $L$  index values each. We then compared the actual difference of  $L$  index values to the shuffled distribution and significance was determined by the percentage of shuffled  $L$  values that fell above the actual difference.

**Pattern analysis.** To assess training-induced changes in the VWFA activation patterns corresponding to the TPW relative to the UTPW set, we first obtained the  $\beta$  values of each voxel for each condition in the individually defined VWFA from each run of Scan 2 (pretraining) and Scan 3 (posttraining). We then averaged  $\beta$  values across conditions and runs for each stimulus group, yielding four average activation patterns (sets of  $\beta$  values) for each subject  $\beta_i^{\text{TPW, pre}}$ ,  $\beta_i^{\text{UTPW, pre}}$ ,  $\beta_i^{\text{TPW, post}}$ , and

$\beta_i^{UTPW, post}$  We then calculated separately for Scan 2 and Scan 3 the Euclidean distance of the activation patterns to get a measure of the dissimilarity of the activations patterns of the TPW and UTPW sets pretraining versus post-training. For example:

$$\text{Distance}^{pre} =$$

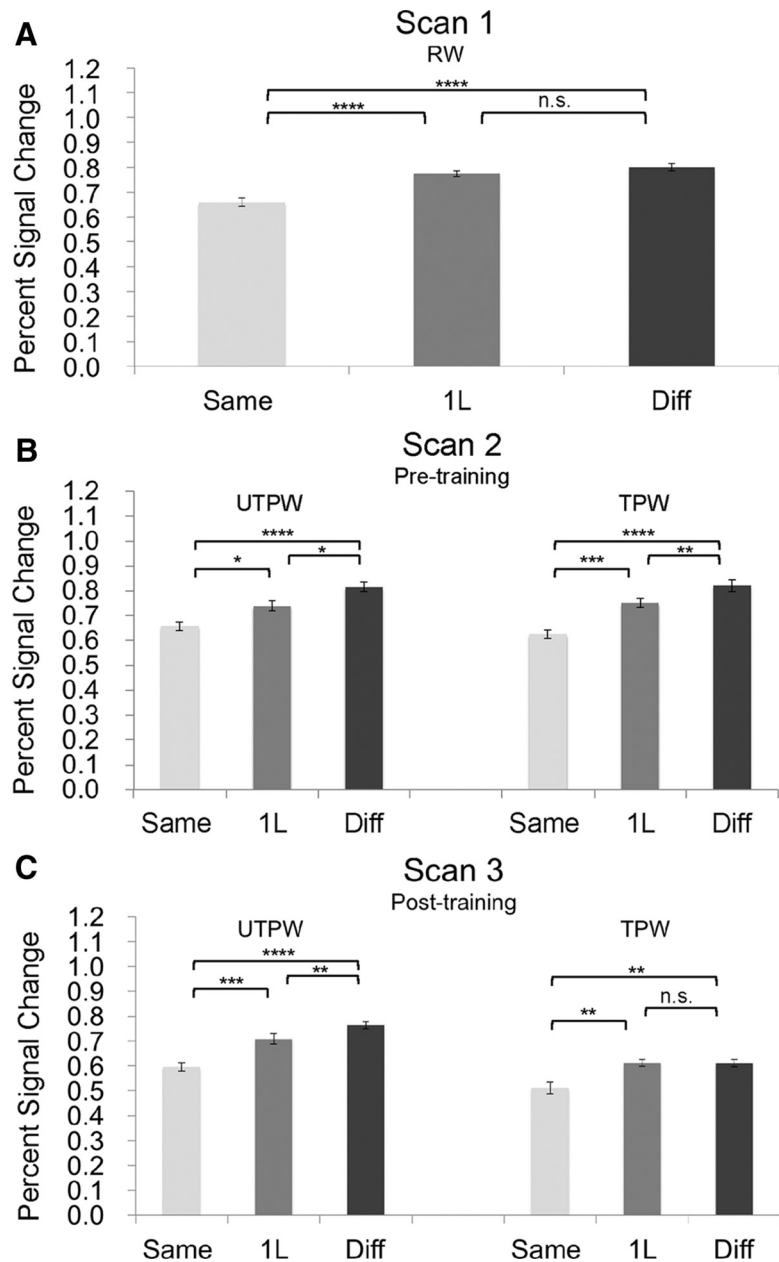
$$\sqrt{\frac{\sum_{i=1}^N (\beta_i^{TPW, pre} - \beta_i^{UTPW, pre})^2}{N}}$$

Where  $N$  is the number of voxels in the VWFA.

## Results

The study involved three fMRI-RA scans performed on three different days (two pretraining and one posttraining) and behavioral testing and training, which were administered over an average of 2–3 weeks. Scan 1 included RW and served as a baseline measure to probe neural selectivity for high-frequency words, as in our previous experiments (Glezer et al., 2009). Scan 2 included two sets of novel PW. For each subject, one set of PWs was pseudo-randomly assigned for training. This set was termed the TPW set and the other set was termed the UTPW set. A localizer scan was run in Scan 2 immediately after the fMRI-RA runs to identify the VWFA using our previously published methods (Glezer et al., 2009; Glezer and Riesenhuber, 2013) and the other ROI (see Materials and Methods). After the pretraining scans, subjects participated in behavioral testing to determine baseline recognition of TPWs and RWs before training. After this, subjects were trained to recognize TPWs and tested on their recognition ability after each training session. Once subjects reached at least 80% accuracy over two sessions, they participated in the posttraining scan in which the TPWs and a novel, previously unseen set of UTPWs were presented while subjects again performed an oddball detection task. After the postscan, subjects were assessed again for overall learning.

Before training, subjects should show high accuracy in identifying RWs as familiar and both TPWs and UTPWs as unfamiliar. Subjects accurately identified RWs as a familiar letter string with 99% accuracy. The to-be-trained PWs (TPWs) were identified as familiar 32% of the time, suggesting that subjects recalled some of the words presented during the pretraining scan [indeed, a portion of the subjects ( $n = 7$ ) also tested on never-seen-before PWs rated these as familiar in 21% of cases on average, significantly lower than TPW familiarity ratings postscan 1,  $p = 0.004$ , paired  $t$  test]. On average, subjects participated in 7.5 training sessions (range 3–15). After training and before the postscan, subjects were on average 91% accurate (range 84%–99%, SD 4.6) at discriminating the TPWs from a 1-letter different foil. In testing after the posttraining scan, subjects were highly accurate at identifying the TPWs as a familiar letter string (94%) comparable to familiarity ratings for

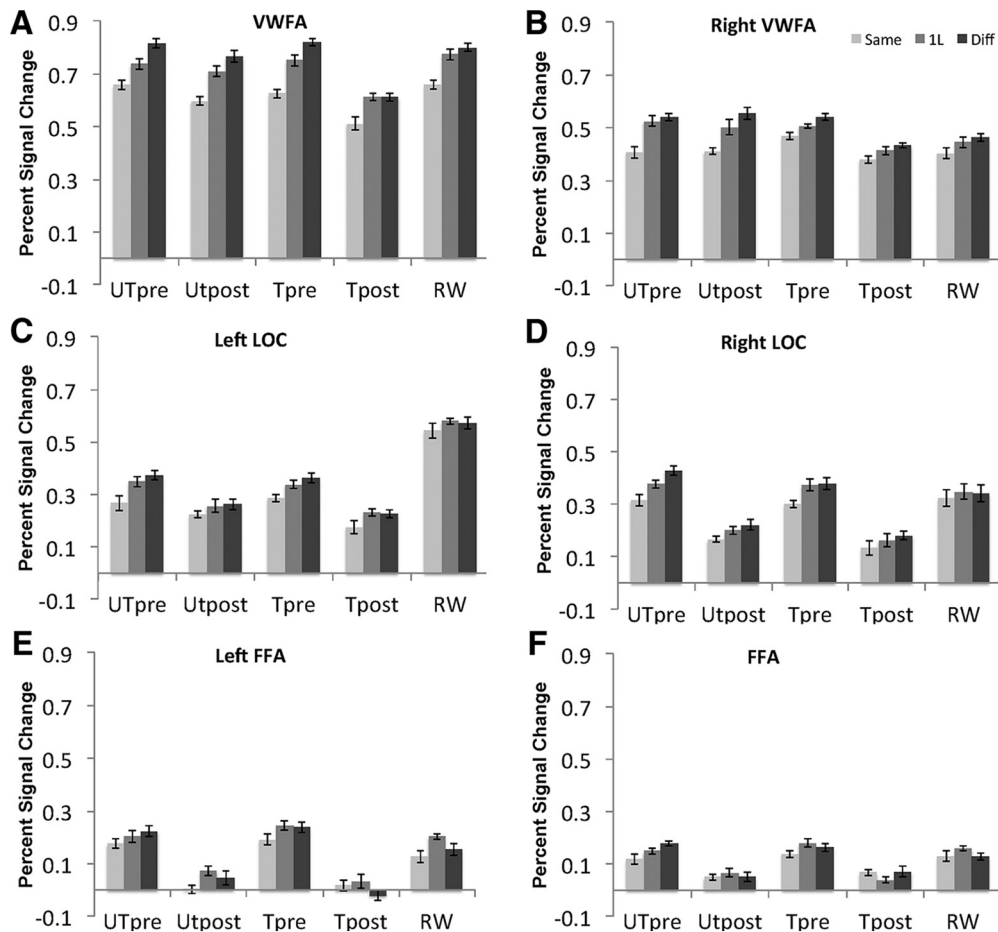


**Figure 1.** Plots of mean percentage signal change in relation to orthographic similarity in the fMRI-RA scans for Scans 1, 2 and 3. Shown are results from RW pairs (**A**), novel UTPWs and novel to-be-trained TPWs before training (**B**), and novel UTPWs and TPWs after training (**C**). Error bars represent within-subject SEM. \*\*\*\* $p < 0.0001$ ; \*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$ ; n.s., not significant.

RWs (97% average); in contrast, only 27% of UTPWs were judged to be familiar.

The VWFA regions were identified for each individual subject independently through localizer scans as described previously (Glezer et al., 2009 and see Materials and Methods). The average location of the thus-defined VWFA ROI was Talairach coordinates ( $-46 \pm 7 -55 \pm 7 -18 \pm 6$ ). For each subject, we then analyzed the activity in their individually defined ROIs during the separate fMRI-RA event-related scans.

Results (Fig. 1) show that, before training, responses in the VWFA showed highly selective tuning for RWs and broad tuning for PWs, replicating (Glezer et al., 2009). In particular, responses to UTPWs and to-be-trained TPWs showed gradual release from adaptation (SAME < 1L < DIFF,  $p < 0.017$  or better, paired  $t$



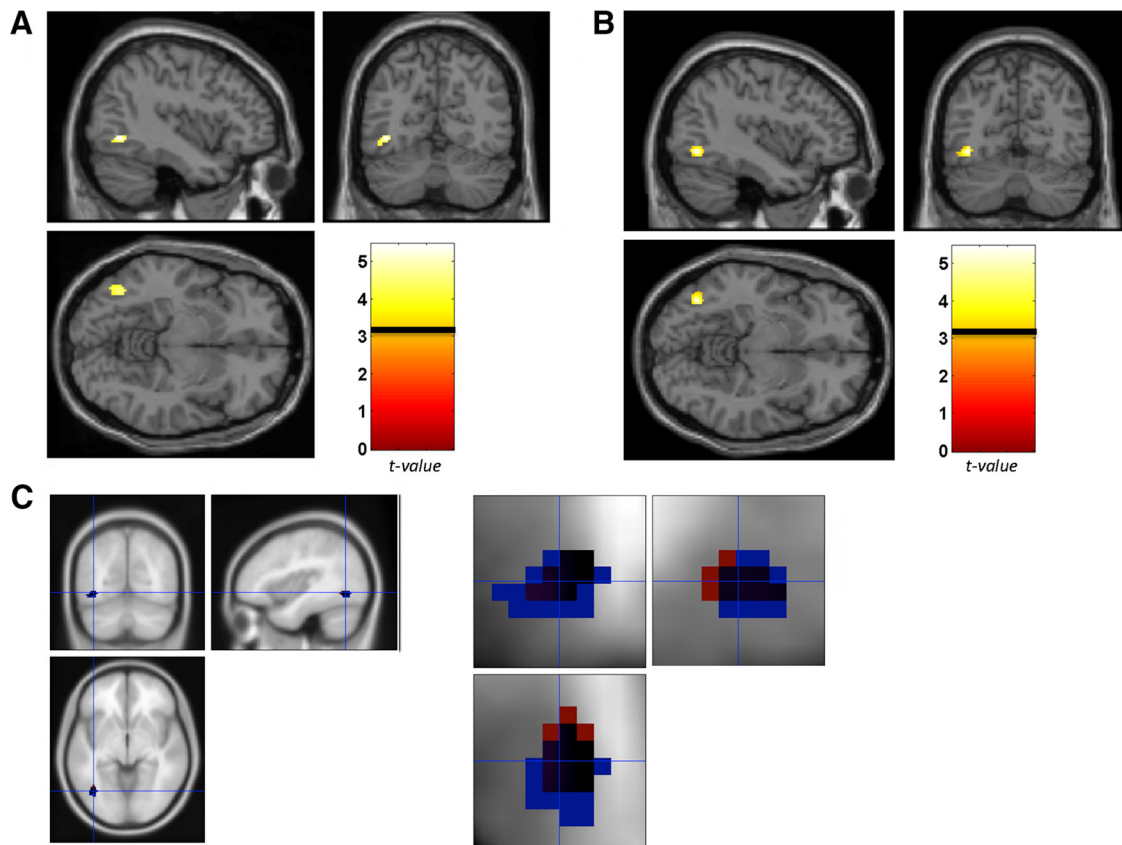
**Figure 2.** Plots of mean percentage signal change in relation to orthographic similarity in the fMRI-RA scans for Scans 1, 2, and 3 for bilateral VWFA (**A, B**), LOC (**C, D**), and FFA (**E, F**).

test), whereas responses to RWs showed full release from adaptation for word pairs that differed by only one letter (SAME < 1L,  $p < 0.0003$ ; 1L = DIFF,  $p = 0.14$ , paired  $t$  test). A  $2 \times 3$  repeated-measures ANOVA with two factors, word set (TPWs vs UTPWs) and experimental condition (SAME, 1L, and DIFF) revealed no difference before training in responses to the to-be trained TPWs and UTPWs ( $F_{(1,11)} = 0.143$ ,  $p = 0.726$ ). In addition, consistent with our previous finding (Glezer et al., 2009), a repeated-measures ANOVA with two factors, stimuli (RWs vs PWs with TPWs and UTPWs collapsed together) and conditions (1L vs DIFF, as we predicted adaptation to differ between these two conditions for RW vs PW, with DIFF > 1L for PWs, but DIFF = 1L for RWs) revealed a significant interaction between stimuli and conditions ( $F_{(1,11)} = 6.246$ ,  $p = 0.030$ ), suggesting different tuning to RWs and PWs in the VWFA.

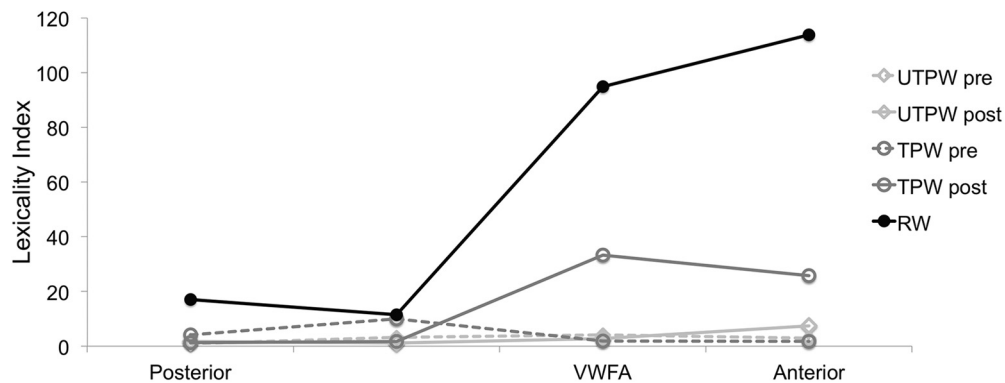
In contrast, after training, responses to TPWs showed RW-like selectivity, indicating learning-induced neural tuning to TPWs, whereas the activation to UTPWs continued to show broad tuning (UTPWs: SAME < 1L < DIFF, at least  $p = 0.006$ ; TPWs: SAME < 1L,  $p = 0.002$  and 1L = DIFF  $p = 0.95$ , paired  $t$  test). A  $2 \times 3 \times 2$  repeated-measures ANOVA with within-subjects factor word set (TPWs vs UTPWs), experimental condition (SAME, 1L, and DIFF), and training (pretraining vs posttraining) revealed significant main effects for word set ( $F_{(1,11)} = 23.874$ ,  $p < 0.001$ ) and experimental condition ( $F_{(2,22)} = 36.683$ ,  $p < 0.001$ ), a significant 3-way interaction ( $F_{(2,22)} = 4.352$ ,  $p = 0.036$ ), and significant 2-way interactions training  $\times$  word set ( $F_{(1,11)} = 25.530$ ,  $p < 0.001$ ) and

training  $\times$  condition ( $F_{(2,22)} = 4.758$ ,  $p = 0.021$ ). Together, these results suggest that training selectively affected the TPWs.

To examine potential learning effects in other localizer-defined brain regions, we conducted the same  $2 \times 3 \times 2$  repeated-measures ANOVA (see above), and there were no significant 3-way interactions: right VWFA (rVWFA) ( $F_{(2,22)} = 0.837$ ,  $p = 0.463$ ), bilateral FFA (left FFA,  $F_{(2,22)} = 1.117$ ,  $p = 0.349$ , right FFA,  $F_{(2,22)} = 2.713$ ,  $p = 0.095$ ), and bilateral LOC (left LOC,  $F_{(2,22)} = 0.884$ ,  $p = 0.394$ , right LOC,  $F_{(2,22)} = 0.408$ ,  $p = 0.665$ ). In addition, a repeated-measures ANOVA on fMRI response to RW (SAME, 1L, DIFF) revealed that none of these regions showed the high degree of RW selectivity that was seen in the VWFA (Fig. 2), rVWFA ( $F_{(2,10)} = 4.163$ ,  $p = 0.058$ , left FFA ( $F_{(2,16)} = 2.636$ ,  $p = 0.129$ ), right FFA ( $F_{(2,22)} = 2.009$ ,  $p = 0.165$ ), left LOC ( $F_{(2,22)} = 0.886$ ,  $p = 0.399$ ), or right LOC ( $F_{(2,20)} = 0.224$ ,  $p = 0.715$ ). In addition, repeated ANOVAs with the following four within-subject factors, conditions (SAME vs 1L vs DIFF), ROI (VWFA vs LLOC vs FFA), stimuli (TPW vs UTPW), and training (pretraining vs posttraining) revealed a significant effect of conditions ( $F_{(2,22)} = 18.016$ ,  $p < 0.001$ ), ROIs ( $F_{(2,22)} = 13.009$ ,  $p = 0.001$ ) and stimuli ( $F_{(1,11)} = 6.768$ ,  $p = 0.025$ ), a marginal effect of training ( $F_{(1,11)} = 2.161$ ,  $p = 0.170$ ), and significant interactions between stimuli and ROI ( $F_{(2,22)} = 10.329$ ,  $p = 0.002$ ), between stimuli and training ( $F_{(2,22)} = 16.444$ ,  $p = 0.002$ ), between training and conditions ( $F_{(2,22)} = 3.718$ ,  $p = 0.047$ ), between ROIs and conditions ( $F_{(2,22)} = 6.426$ ,  $p = 0.003$ ), and a significant 4-way interactions between the 4 factors ( $F_{(4,44)}$



**Figure 3.** Results from whole-brain analyses of training effects for the TPW word set. **A, B**, DIFF-1L pretraining > DIFF-1L posttraining (**A**) and 1L-SAME > DIFF-1L posttraining (**B**). In both analyses, the VWFA region is the only consistent location that shows a change in word selectivity as a result of training. **C**, Overlap (black) of the ROI from **A** (red) and **B** (blue). Color bar indicates *t*-value; horizontal line across the color bar shows the threshold that was applied for the analysis (see Materials and Methods).



**Figure 4.** Responses along the word-processing hierarchy to RWs and UTPW/TPWs pretraining and posttraining. Shown are *L* values for RWs and novel PWs (TPW pre and UTPW pre) before training and TPWs and a new set of UTPWs after training (TPW post, UTPW post).

= 3.368, *p* = 0.034). These results suggest that the selectivity to RWs and training-induced change in selectivity to TPWs is specific to the left VWFA.

To probe for evidence of training-induced changes in selectivity across the whole brain, we also conducted a series of whole-brain group analyses to determine whether other areas of the brain showed adaptation to the experimental stimuli. Specifically, we first focused on the training-induced change in the difference between the 1L and DIFF conditions because release from adaptation in the 1L condition should increase as training increases the selectivity of the representation in the VWFA for the training words. Specifically, for the TPWs, we performed the contrast of DIFF-1L pretraining >

DIFF-1L posttraining (*p* = 0.005, uncorrected, cluster extent 20 voxels). In a second analysis, we focused on the posttraining scan and the predicted effect of training—which should not only lead to small signal differences between 1L and DIFF, but also to large signal differences between SAME and 1L (because increased selectivity post-training would be expected to produce less adaptation for the 1L condition)—performing the contrast of 1L-SAME > DIFF-1L for the TPW set (*p* = 0.005, uncorrected, cluster extent 20 voxels). Both analyses showed that the only consistent region for a training effect occurred around the VWFA region (the peak voxel that was consistent for both contrasts was almost identical:  $-40 -64 -6$  and  $-40 -66 -8$ , respectively; Fig. 3) and we did not observe consistent

effects in other brain regions. We used the MarsBar toolbox (Brett et al., 2002) to transform the ROI obtained for both of these contrasts, identified all voxels the two ROIs had in common (27 voxels), and then extracted percentage signal change in this shared ROI in the pretraining and posttraining scans. A repeated-measures ANOVA (pretraining vs posttraining  $\times$  S, 1L, DIFF) showed a main effect for condition ( $F_{(2,22)} = 10.294, p = 0.003$ ) and a significant interaction for the TPW set ( $F_{(2,22)} = 6.194, p = 0.011$ ), consistent with results from the individually defined ROI analysis.

To determine whether training also affected voxel-based activation patterns in the VWFA, we calculated the Euclidean distance of the voxelwise activation patterns to TPWs and UTPWs both pretraining and posttraining (see Materials and Methods). This analysis revealed that activation patterns for TPWs and UTPWs became more dissimilar after training ( $p < 0.0006$ , from a mean distance of 0.2,  $SD = 0.104$  to a mean distance of 0.53  $SD = 0.246$ ), suggesting that training led to more distinctive activation patterns for familiar versus novel words.

To determine the effects of novel word learning along the ventral visual stream hierarchy in the left hemisphere, we identified four ROI locations in each individual subject with the same contrast used to identify the VWFA. In the left hemisphere, this contrast typically yielded 3–4 ROIs posterior to the VWFA ( $p < 0.05$ , corrected) and 1 anterior. ROIs were selected and grouped into 4 categories (we were able to identify all 4 ROIs in 10/12 subjects): (1) area 18 (average Talairach coordinates  $-31 \pm 7 -92 \pm 5 -6 \pm 4$ ); (2) inferior occipital gyrus ( $-43 \pm 6 -79 \pm 6 -9 \pm 4$ ); (3) the VWFA ( $-48 \pm 6 -58 \pm 8 -18 \pm 6$ ); and (4) the inferior temporal gyrus ( $-43 \pm 11 -46 \pm 7 -23 \pm 4$ , i.e., anterior to the VWFA) (see Materials and Methods for details on how ROIs were derived). We then analyzed the activity in the four individually defined ROIs during the separate fMRI-RA scans. To determine the degree of sublexical and lexical processing along the hierarchy, we computed an index to measure degree of lexicality (see Materials and Methods),  $L$ . As shown in Figure 4, along the ventral stream, lexicality for RWs and TPWs posttraining showed a strong increase in the VWFA relative to posterior areas, whereas UTPW both pretraining and posttraining showed little change along the hierarchy from early visual cortex to the VWFA. Interestingly, we found an increase in  $L$  for RWs in the anterior ROI, but a decrease for TPWs, consistent with reports that regions anterior to the VWFA are involved in multimodal processing, including semantics (Moore and Price, 1999; McDermott et al., 2003; Vigneau et al., 2006; Wilson et al., 2009; Binder et al., 2009). Bootstrap testing revealed that, although lexicality indices for TPWs and UTPWs were similar in the 2 posterior areas (at least  $p > 0.15$ ), there was a significant difference in the VWFA ( $p = 0.0089$ ), with stronger lexicality for TPW than UTPW ( $L = 33.4$  vs 2.8, respectively). Together, these results provide support for theories that propose a hierarchical organization along the ventral visual processing stream, from low orthographic selectivity posterior to the VWFA to high orthographic selectivity in the VWFA (i.e., prelexical and sublexical to lexical). Our results also show directly that experience with the written word results in changes in neural tuning in the VWFA—from broad tuning to UTPW to tight tuning to TPW—and that this change in neural tuning occurs exclusively in the VWFA.

## Discussion

In sum, our study replicated findings from Glezer et al. (2009) showing tight tuning to RW and broad tuning to PW in VWFA, compatible with the theory that the VWFA contains a representation for written RWs; that is, an orthographic lexicon. Crucially, we show an experience-dependent change in selectivity in the VWFA for newly learned words after training, indicating that

the trained words were added to the orthographic lexicon. This change in tuning is specific to the VWFA. In addition, we show in the ventral visual stream a change from prelexical to lexical representations for RWs and trained words, but not for untrained words.

It has been proposed that the VWFA develops with reading acquisition as a result of the “recycling” of visual cortex, resulting in neurons dedicated to orthographic processing (Dehaene et al., 2010). Our study supports the theory (Glezer et al., 2009) that the role of the VWFA in reading is that of an orthographic lexicon in which during word learning, neurons come to be selective for the “objects” of reading, that is, whole words, enabling the rapid recognition of familiar words. These findings have interesting implications for reading remediation in individuals with phonologic processing impairments because they suggest the possibility that these individuals might benefit from visual word learning strategies to circumvent the phonologic difficulties and directly train holistic visual word representations in the VWFA. Together with other recent reports on holistic object representations in the fusiform cortex, in particular for faces in the FFA (Jiang et al., 2006), these results suggest a converging role of the fusiform cortex as an area containing high-level, whole-object visual representations that are flexibly shaped by experience, compatible with single-neuron electrophysiology studies in monkey inferotemporal cortex that have reported holistic tuning for trained stimuli (Logothetis and Pauls, 1995; Baker et al., 2002). It will be interesting in future studies to explore a number of further questions arising from this work. For example, what is the role of phonological information? Our study shows that training on visually presented words induces changes specifically in the VWFA, compatible with the VWFA's function as a visual area (Dehaene and Cohen, 2011). However, the PWs are likely also coded phonologically during training (Kronbichler et al., 2004, 2007), but it is unlikely that the findings in the VWFA reflect phonological processing. Prior studies have shown no evidence of response modulations by phonological similarity in the VWFA (Kronbichler et al., 2007; Rothlein and Rapp, 2014) and, using an fMRI-RA paradigm, we recently presented data showing that phonological similarity does not modulate responses in the VWFA (Glezer et al., 2011). It will be interesting in future studies to determine how phonological information is used in novel word learning. Another interesting question, given the “neuronal recycling” hypothesis, is whether other, non-orthographic stimuli might also lead to plasticity focused on the VWFA region as in our study or in other regions in the ventral occipitotemporal region. Remarkably, a very recent study has provided evidence for plasticity in the left fusiform cortex for a “face abet” (Moore et al., 2014), suggesting that this part of the brain as a whole might play a general role in linking visual stimuli to the language network. Determining how different types of visual object training affect selectivity in the VWFA vis-à-vis surrounding areas will help to elucidate the role of this region in reading.

## References

- Aguirre GK, Zarahn E, D'Esposito M (1998) The inferential impact of global signal covariates in functional neuroimaging analyses. *Neuroimage* 8:302–306. [CrossRef Medline](#)
- Baayen RH, Piepenbrock R, van Rijn H (1995) The CELEX lexical database. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.
- Baker CI, Behrmann M, Olson CR (2002) Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nat Neurosci* 5:1210–1216. [CrossRef Medline](#)
- Baker CI, Liu J, Wald LL, Kwong KK, Benner T, Kanwisher N (2007) Visual word processing and experiential origins of functional selectivity in hu-

- man extrastriate cortex. *Proc Natl Acad Sci U S A* 104:9087–9092. [CrossRef Medline](#)
- Ben-Shachar M, Dougherty RF, Deutsch GK, Wandell BA (2007) Differential sensitivity to words and shapes in ventral occipito-temporal cortex. *Cereb Cortex* 17:1604–1611. [CrossRef Medline](#)
- Ben-Shachar M, Dougherty RF, Deutsch GK, Wandell BA (2011) The development of cortical sensitivity to visual word forms. *J Cogn Neurosci* 23:2387–2399. [CrossRef Medline](#)
- Binder JR, Desai RH, Graves WW, Conant LL (2009) Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex* 19:2767–2796. [CrossRef Medline](#)
- Bowers JS, Davis CJ, Hanley DA (2005) Interfering neighbours: the impact of novel word learning on the identification of visually similar words. *Cognition* 97:B45–B54. [CrossRef Medline](#)
- Brem S, Bach S, Kucian K, Guttorm TK, Martin E, Lyytinen H, Brandeis D, Richardson U (2010) Brain sensitivity to print emerges when children learn letter-speech sound correspondences. *Proc Natl Acad Sci U S A* 107:7939–7944. [CrossRef Medline](#)
- Brett M, Anton JL, Valabregue R, Poline JB (2002) Region of interest analysis using an SPM toolbox. Paper presented at the 8th International Conference on Functional Mapping of the Human Brain. June 2–6, 2002, Sendai, Japan.
- Buracas GT, Boynton GM (2002) Efficient design of event-related fMRI experiments using M-sequences. *Neuroimage* 16:801–813. [CrossRef Medline](#)
- Callan AM, Callan DE, Masaki S (2005) When meaningless symbols become letters: Neural activity change in learning new phonograms. *Neuroimage* 28:553–562. [CrossRef Medline](#)
- Chalmers KA, Burt JS (2008) Phonological and semantic information in adults' orthographic learning. *Acta Psychol* 128:162–175. [CrossRef](#)
- Cohen L, Dehaene S (2004) Specialization within the ventral stream: the case for the visual word form area. *Neuroimage* 22:466–476. [CrossRef Medline](#)
- Dehaene S, Cohen L (2011) The unique role of the visual word form area in reading. *Trends Cogn Sci* 15:254–262. [CrossRef Medline](#)
- Dehaene S, Pegado F, Braga LW, Ventura P, Nunes Filho G, Jobert A, Dehaene-Lambertz G, Kolinsky R, Morais J, Cohen L (2010) How learning to read changes the cortical networks for vision and language. *Science* 330:1359–1364. [CrossRef Medline](#)
- Fisher JE, Cortes CR, Griego JA, Tagamets MA (2012) Repetition of letter strings leads to activation of and connectivity with word-related regions. *Neuroimage* 59:2839–2849. [CrossRef Medline](#)
- Gaskell MG, Dumay N (2003) Lexical competition and the acquisition of novel words. *Cognition* 89:105–132. [CrossRef Medline](#)
- Glezer LS (2009) Investigating the neural code for single-word reading (doctoral dissertation). Washington, DC: Georgetown University.
- Glezer LS, Riesenhuber M (2013) Individual variability in location impacts orthographic selectivity in the “visual word form area.” *J Neurosci* 33:11221–11226. [CrossRef Medline](#)
- Glezer LS, Jiang X, Riesenhuber M (2009) Evidence for highly selective neuronal tuning to whole words in the “visual word form area.” *Neuron* 62:199–204. [CrossRef Medline](#)
- Glezer LS, Jiang X, Riesenhuber M (2011) Evidence that orthography, not phonology determines selectivity in the occipitotemporal cortex. Poster presented at the Society for Neuroscience Meeting, Washington, D.C.
- Grady CL, Horwitz B, Pietrini P, Mentis MJ, Ungerleider LG, Rapoport SI, Haxby JV (1996) Effects of task difficulty on cerebral blood flow during perceptual matching of faces. *Hum Brain Mapp* 4:227–239. [CrossRef Medline](#)
- Grill-Spector K, Henson R, Martin A (2006) Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci* 10:14–23. [CrossRef Medline](#)
- Henson R, Shallice T, Dolan R (2000) Neuroimaging evidence for dissociable forms of repetition priming. *Science* 287:1269–1272. [CrossRef Medline](#)
- Jiang X, Rosen E, Zeffiro T, Vanmeter J, Blanz V, Riesenhuber M (2006) Evaluation of a shape-based model of human face discrimination using fMRI and behavioral techniques. *Neuron* 50:159–172. [CrossRef Medline](#)
- Jiang X, Bradley E, Rini RA, Zeffiro T, Vanmeter J, Riesenhuber M (2007) Categorization training results in shape- and category-selective human neural plasticity. *Neuron* 53:891–903. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Kronbichler M, Bergmann J, Hutzler F, Staffen W, Mair A, Ladurner G, Wimmer H (2007) Taxi vs taxi: on orthographic word recognition in the left ventral occipitotemporal cortex. *J Cogn Neurosci* 19:1584–1594. [CrossRef Medline](#)
- Loftus GR, Masson ME (1994) Using confidence intervals in within-subject designs. *Psychon Bull Rev* 1:476–490. [CrossRef Medline](#)
- Logothetis NK, Pauls J (1995) Psychophysical and physiological evidence for viewer-centered object representations in the primate. *Cereb Cortex* 5:270–288. [CrossRef Medline](#)
- McDermott KB, Petersen SE, Watson JM, Ojemann JG (2003) A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. *Neuropsychologia* 41:293–303. [CrossRef Medline](#)
- Medler DA, Binder JR (2005) An on-line orthographic database of the English language. Available from: <http://www.neuro.mcw.edu/mcword/>. Accessed Sept. 10, 2010.
- Moore CJ, Price CJ (1999) Three distinct ventral occipitotemporal regions for reading and object naming. *Neuroimage* 10:181–192. [CrossRef Medline](#)
- Moore MW, Durisko C, Perfetti CA, Fiez JA (2014) Learning to read an alphabet of human faces produces left-lateralized training effects in the fusiform gyrus. *J Cogn Neurosci* 26:896–913. [CrossRef Medline](#)
- Murray SO, Wojciulik E (2004) Attention increases neural selectivity in the human lateral occipital complex. *Nat Neurosci* 7:70–74. [CrossRef Medline](#)
- Price CJ, Devlin JT (2003) The myth of the visual word form area. *Neuroimage* 19:473–481. [CrossRef Medline](#)
- Rothlein D, Rapp B (2014) The similarity structure of distributed neural responses reveals the multiple representations of letters. *Neuroimage* 89:331–344. [CrossRef Medline](#)
- Salasoo A, Shiffrin RM, Feustel TC (1985) Building permanent memory codes: codification and repetition effects in word identification. *J Exp Psychol Gen* 114:50–77. [CrossRef Medline](#)
- Sandak R, Mencl WE, Frost SJ, Rueckl JG, Katz L, Moore DL, Mason SA, Fulbright RK, Constable RT, Pugh KR (2004) The neurobiology of adaptive learning in reading: a contrast of different training conditions. *Cogn Affect Behav Neurosci* 4:67–88. [CrossRef Medline](#)
- Sunaert S, Van Hecke P, Marchal G, Orban GA (2000) Attention to speed of motion, speed discrimination, and task difficulty: an fMRI study. *Neuroimage* 11:612–623. [CrossRef Medline](#)
- Szwed M, Dehaene S, Kleinschmidt A, Eger E, Valabrègue R, Amadon A, Cohen L (2011) Specialization for written words over objects in the visual cortex. *Neuroimage* 56:330–344. [CrossRef Medline](#)
- Tamminen J, Gaskell MG (2008) Newly learned spoken words show long-term lexical competition effects. *Q J Exp Psychol (Hove)* 61:361–371. [CrossRef](#)
- van der Mark S, Bucher K, Maurer U, Schulz E, Brem S, Buckelmueller J, Kronbichler M, Loenneker T, Klaver P, Martin E, et al. (2009) Children with dyslexia lack multiple specializations along the visual word-form (VWF) system. *Neuroimage* 47:1940–1949. [CrossRef Medline](#)
- Vigneau M, Beaucousin V, Hervé PY, Duffau H, Crivello F, Houdé O, Mazoyer B, Tzourio-Mazoyer N (2006) Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage* 30:1414–1432. [CrossRef Medline](#)
- Vinckier F, Dehaene S, Jobert A, Dubus JP, Sigman M, Cohen L (2007) Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron* 55:143–156. [CrossRef Medline](#)
- Wilson SM, Brambati SM, Henry RG, Handwerker DA, Agosta F, Miller BL, Wilkins DP, Ogar JM, Gorno-Tempini ML (2008) The neural basis of surface dyslexia in semantic dementia. *Brain* 132:71–86. [CrossRef Medline](#)
- Xue G, Poldrack RA (2007) The neural substrates of visual perceptual learning of words: implications for the visual word form area hypothesis. *J Cogn Neurosci* 19:1643–1655. [CrossRef Medline](#)
- Xue G, Chen C, Jin Z, Dong Q (2006) Language experience shapes fusiform activation when processing a logographic artificial language: An fMRI training study. *Neuroimage* 31:1315–1326. [CrossRef Medline](#)