

Differential Sensitivity to Words and Shapes in Ventral Occipito-Temporal Cortex

Michal Ben-Shachar^{1,2}, Robert F. Dougherty^{1,2}, Gayle K. Deutsch¹ and Brian A. Wandell^{1,2}

¹Stanford Institute for Reading and Learning and ²Psychology Department, Stanford University, Stanford, CA 94305, USA

Efficient extraction of shape information is essential for proficient reading but the role of cortical mechanisms of shape analysis in word reading is not well understood. We studied cortical responses to written words while parametrically varying the amount of visual noise applied to the word stimuli. In only a few regions along the ventral surface, cortical responses increased with word visibility. We found consistently increasing responses in bilateral posterior occipito-temporal sulcus (pOTS), at an anatomical location that closely matches the “visual word form area”. In other cortical regions, such as V1, responses remained constant regardless of the noise level. We performed 3 additional tests to assess the functional specialization of pOTS responses for written word processing. We asked whether pOTS responses are 1) left lateralized, 2) more sensitive to words than to line drawings or false fonts, and 3) invariant for visual hemifield of words but not other stimuli. We found that left and right pOTS response functions both had highest sensitivity for words, intermediate for line drawings, and lowest for false fonts. Visual hemifield invariance was similar for words and line drawings. These results suggest that left and right pOTS are both involved in shape processing, with enhanced efficiency for processing visual word forms.

Keywords: false fonts, line drawings, parametric fMRI, phase scrambling, word reading

Introduction

The human visual system efficiently extracts and identifies shapes and forms, including letters, digits, signs, and line drawings. In an important application of this visual skill, children learning to read become proficient at identifying and discriminating letters and words. A variety of neuropsychological and neuroimaging studies demonstrate that the ventral occipito-temporal cortex is essential for many aspects of visual form processing, including face, object, and word recognition (Arguin and others 1996; Puce and others 1996; Kanwisher and others 1997; Grill-Spector and others 1998; De Renzi 2000; Hasson and others 2002; James and others 2003; Farah 2004; Hanson and others 2004; Salvan and others 2004). Some authors have argued for a modular organization of ventral occipito-temporal cortex in which different parts are involved in the recognition of specific object categories (Tranel and others 1997; Spiridon and Kanwisher 2002) but the extent to which these functions are modularly organized is debated (Chao and others 1999; Haxby and others 2001).

The principle of a cortical module for visual word forms was introduced by Dejerine (1892) and later reformulated in the neuropsychological literature (e.g., Warrington and Shallice 1980). Recent neuroimaging measurements have been interpreted as supporting a visual word form module in left fusiform gyrus/

occipito-temporal sulcus (OTS) (Cohen and others 2000, 2002, 2004; Dehaene and others 2005), though the location of the proposed module differs from the left inferior parietal (angular gyrus) region described by Dejerine and others. The existence of a region specialized for processing visual word forms has become the focus of a debate spanning both sensory modality and stimulus specificity (Price and Devlin 2003, 2004; Vigneau and others 2005).

In the current study, we investigated the issue of cortical modularity by parametrically controlling the amount of visual noise added to shapes and measuring blood oxygen level-dependent (BOLD) responses to several stimulus types. Specifically, we varied the amount of phase scrambling applied to words, line drawings, and false-font strings and compared the resulting response function for each stimulus type (see Avidan and others 2002; Horowitz and others 2004, for parametric manipulations of contrast and noise in face and object recognition). This design maps responses across a range of noise levels and separately estimates the sensitivity of a cortical region to each stimulus type. Using a parametric design limits the number of stimulus categories which can be compared (as each stimulus category is repeated in all noise levels). However, a parametric design allows us to compare the full response function for different stimulus types, and thus could reveal differences that are invisible to the direct comparison of stimulus categories at zero noise.

Our initial experiments established that parametric manipulation of noise level added to words gives rise to consistent modulation of the responses in left posterior OTS (pOTS), matching the anatomical location of the “visual word form area” (VWFA) (Cohen and others 2000, p. 302; Dehaene and others 2005, p. 335). We then performed additional experiments and analyses to assess whether the responses in left pOTS are word specific. First, we asked whether the word-response functions in pOTS are left lateralized, and thus tightly coupled to language processing. Second, we asked whether left pOTS response functions are specific to words or whether they generalize to other contour defined shapes (nameable—line drawings, and non-nameable—false-font strings). Third, we asked whether the known invariance of left pOTS responses to visual hemifield presentation is unique to word stimuli, or whether it extends to line drawings as well. We conclude that both left and right pOTS play a general role in form extraction with enhanced sensitivity to word stimuli.

Materials and Methods

Subjects

Eleven healthy, native English speakers, right-handed subjects participated in the experiments (experiment 1: S1-9; experiment 2: S1-7;

experiment 3: S1-4, S6-7; experiment 4: S1-2, S5-7, S10-11). Subjects (4 females, 6 males, 23-53 years old) were graduate students or lab members, and had corrected to normal vision with no reported cognitive deficits. All subjects gave informed consent to participate in the study; 2 were monetarily compensated for their participation. The experimental protocol was approved by the institutional review board at Stanford University.

Stimuli

Experiment 1

Word stimuli ($N = 144$) were selected from a pool of 192 English 4-letter nouns (mean frequency: 65) (Kucera and Francis 1967), such as: roof, town, cake, bone, duck, etc. Words were rendered in white font "MS Sans Serif" (18 pt), within a gray rectangular frame. The frame and the letters were centered on the fixation mark, such that there were always 2 letters on each side of the cross, with minimal overlap between the letters and the cross (see Fig. 1A). The rectangle width was 7° visual angle; the words spanned $2.2-4.42^\circ$ (mean 3.5°) around the fixation cross.

A randomized-phase stimulus was created for each word by computing the 2-dimensional Fourier transform of the word image, randomizing the phase, and then applying the inverse Fourier transform to the image. The new phase, ϕ_N , was computed from the original phase, ϕ_O , by adding a scaled random variable, \tilde{U} , uniformly distributed between $[-\pi, \pi]$.

$$\phi_N = \phi_O + s\tilde{U}$$

The amount of noise was controlled by the scalar, s , which was set to [1, 0.82, 0.72 or 0.3]. The s -levels were chosen during preliminary behavioral testing to create images ranging from full noise (1) to plainly visible shape (0.3).

Experiment 2

Stimuli for experiment 2 included 144 line drawings, selected from 168 black on white line drawings downloaded with permission from the International Picture Naming Project database (Szekely and others 2004). This experiment further included the word stimuli as in experiment 1. Line drawings are similar to words in that both are shapes defined by contours, and both are nameable. These categories are different from each other in exposure level (words > line drawings) and in the fact that words consist of spatially distinguished letters. Line-drawing stimuli depicted objects (e.g., phone, scissors) and animals (e.g., pelican, bat, alligator). Line drawings were rotated and scaled to fill the same bounding box covered by the word stimuli. We then reversed the

stimulus contrast (white line drawings on gray background) and adjusted it so that the overall contrast and mean were similar to the word stimuli. We applied the same phase-scrambling procedure as in experiment 1 to create stimuli with a range of noise levels (see Fig. 1B). This resulted in 4 experimental conditions for each stimulus category (words and line drawings) distinguished by the amount of phase noise.

Experiment 3

Stimuli for this experiment were strings of 4 false fonts processed and presented as in experiment 1. We used false fonts to minimize the naming response, which may be evoked by words and line drawings. In contrast with words and line drawings, false-font strings do not map to a lexical item. Therefore, increased activation for false fonts as a function of stimulus visibility is hard to explain as a result of naming responses. A set of 26 false letters were created as follows: First, 26 real letters were produced using a simple set of b-spline control points for each letter. The control points were then randomly jittered to create the false fonts. Several sets of false fonts were created, and any that happened to form a recognizable letter or shape was eliminated by visual inspection. With this method, the false fonts have a similar complexity to the real letters (e.g., the same number of strokes and inflection points). The b-splines were rendered with a stroke thickness similar to typical fonts. Once rendered, we applied the same phase-scrambling procedure as in experiment 1 to create 4 experimental conditions with increasing noise levels (see Fig. 1C).

Experiment 4

We used a subset of the word and line-drawing stimuli from experiment 2 (72 words and 72 line drawings), without phase scrambling. We also presented 72 gray/white checkerboards, matched to the words and line drawings in frame size.

Stimulus Presentation and Task

Stimulus presentation and response collection were controlled by E-Prime (Psychology Software Tools, Pittsburgh, PA). Stimuli were presented through a backlit projection screen, visible to the subject by a mirror mounted on the top of the head coil. Responses were collected using a magnetic resonance imaging (MRI) compatible response box.

In all 4 experiments, stimuli were presented in a block design. Experimental blocks (12 s long) were interleaved with fixation blocks of the same length. A fixation mark (+ sign, Arial font size 16, bold), either dark blue or black, was present in all trials (including fixation blocks). Each experimental run began with an experimental block (excluded from analysis, see below) and ended with a fixation block. Runs included 7 experimental and 7 fixation blocks.

In experiments 1-3, stimuli (with fixation mark) were presented centrally for 1700 ms, followed by a blank screen (300 ms), resulting in an inter stimulus interval (ISI) of 2000 ms (6 stimuli per block). Subjects viewed 5 runs for each stimulus category (words, line drawings). Subjects judged the color of the fixation mark in each trial and responded with the button press assigned to each color. Thus, reading was implicit. Implicit reading tasks yield robust functional MRI signals (e.g., Moore and Price 1999; Turkeltaub and others 2003).

Stimuli in experiment 4 were presented for 200 ms with an ISI of 1000 ms (12 stimuli per block), a presentation rate similar to the original experiment by Cohen and others (2000). Stimuli (words, line drawings and checkerboards) were presented to the left or right of fixation, centered on a visual angle of 4° , covering $2-6^\circ$ off fixation. The fixation judgment task was modified to fit the faster presentation rate. Subjects detected brief color changes by a button press (2 per block at random times). Subjects viewed 6-10 runs.

MRI Data Acquisition

Functional MRI (fMRI) measurements were performed on a 3-T General Electric scanner (GE Medical Systems, Milwaukee, WI) with a custom built volume head coil. Head movements were minimized by padding and tape. fMRI data were acquired with a spiral pulse sequence (Glover 1999). Twenty-six oblique slices were prescribed approximately along the anterior commissure - posterior commissure (ac-pc) plane, to cover the whole occipital and temporal lobes. The sampling rate of the BOLD

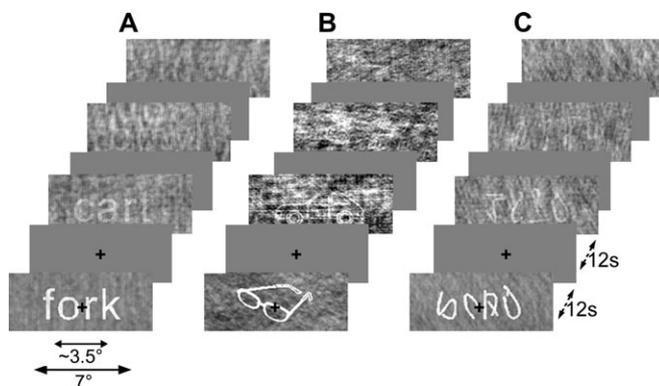


Figure 1. Experimental paradigm. Sample stimuli for each noise level in the (A) words, (B) line drawings, and (C) false-fonts experiments. Shape visibility was controlled parametrically (4 noise levels). Stimuli were presented in a block design. Within each block the noise level was constant. Experimental blocks (12 s long, 6 stimuli) were interleaved with fixation blocks (12 s, uniform gray rectangles). A black or dark blue fixation cross was refreshed every 2 s, throughout all experimental and fixation blocks (shown only in the first 2 blocks). Subjects indicated the color of the fixation cross (see Methods).

signal was 3 s (experiments 1–2) or 2 s (experiments 3–4). The voxel size was $2.5 \times 2.5 \times 3$ mm.

A set of in-plane anatomical images were acquired before the functional scans using a fast spoiled gradient recalled (SPGR) sequence. These T_1 -weighted slices were taken at the same location of the functional slices, and were used to align the functional data with high-resolution anatomical data acquired in a separate session.

We further collected high-resolution whole-brain anatomical images for each subject. These images were acquired on a GE 1.5-T Signa LX scanner using a 3-D SPGR pulse sequence. We acquired and averaged several T_1 -weighted anatomical data sets in the axial and sagittal planes for each subject.

Data Analysis

Preprocessing

fMRI data were analyzed using the mrVista tools (<http://white.stanford.edu/software/>). Data were analyzed voxel by voxel in individual subjects with no spatial smoothing. Baseline drifts were removed from the time series by high pass temporal filtering. Between scan motion was detected in 2/18 sessions. In those cases, motion compensation was implemented by applying rigid and nonlinear transformations that minimized the mutual information between the mean maps of the scans.

High-resolution whole-brain T_1 anatomies were aligned, averaged, and resampled into a $1 \times 1 \times 1$ mm resolution 3-dimensional anatomical volume (Dougherty and others 2005). We applied inhomogeneity correction and rotation to the ac-pc plane using tools developed at the Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB, University of Oxford, <http://www.fmrrib.ox.ac.uk/fsl/>). Functional data were registered to the high-resolution anatomical volume for region of interest (ROI) definition, Montreal Neurological Institute (MNI) space alignment and visualization. Alignment was computed between the in-plane anatomicals and the high-resolution volume anatomy using a mutual-information coregistration algorithm (from SPM2). This transformation was then applied to the functional data. For the purpose of MNI coordinate assignment, a nonlinear transformation was computed between the volume anatomy and the MNI template (ICBM152) using the spatial normalization tools from SPM2 (<http://www.fil.ion.ucl.ac.uk/spm/>).

The high-resolution volume anatomy of 8 subjects was segmented into gray and white matter using custom software and then hand edited to minimize segmentation errors (Teo and others 1997). The surface at the white/gray border was smoothed and rendered as a 3-dimensional surface using the visualization toolkit (VTK) software (<http://www.vtk.org/>). Data from all gray layers were mapped to this surface, and the maximum value of those was assigned to each triangle on the surface. For visualization purposes only, spatial smoothing along the cortical surface was applied using an iterative neighborhood average roughly equivalent to a 4 mm full width at half maximum Gaussian.

ROI Analysis

ROIs were defined in individual subject brains according to the following anatomical guidelines: posterior OTS, in the depth of the OTS, at its posterior part where OTS meets the inferior temporal sulcus; V1 (foveal), at the posterior third of the calcarine sulcus. Within these anatomical borders, we collected time course data from voxels that passed the threshold ($p < 10^{-3}$, uncorrected) for a functional localizer contrasting all word + noise conditions (excluding full noise) versus fixation. Statistical parametric maps were computed by fitting a general linear model (GLM) to each voxel's time course, and estimating the relative contribution of each experimental condition to the time course. GLM predictors were constructed as a boxcar for each condition convolved with a standard hemodynamic response model (Boynton and others 1996). Additional predictors were added for each run, to model between run variations. Contrast maps were computed as voxel-wise t -tests between the weights of the predictors of the relevant experimental conditions. Definition of ROIs was checked in the volume to corroborate the MNI location and to make sure the voxels constitute one coherent cluster. Time course data from the clusters of functional activation were always collected in the original in-plane, noninterpolated data.

The BOLD contrast measure used to plot response functions was computed for each subject and each ROI separately, as follows. First,

obtain the responses to all experimental blocks, B_i . A temporal block is defined as including all of the BOLD measurements from 6 s prior to the first stimulus and extending to 11 s past the final stimulus presentation. Find the average percent modulation of the BOLD time series within each block across all blocks:

$$u(t) = \frac{1}{N} \sum_{i=1}^N (B_i(t) - \bar{B}_i)$$

Then transform the mean percent modulation into a unit length vector $u/|u|$. This quantity represents the average normalized time series across all experimental blocks. It accounts for the hemodynamic response and block duration. Then, we computed the mean time course of the responses in the j th experimental condition, u_j . We compute the scalar response contrast for this condition as

$$c_j = \left\langle u_j, \frac{u}{|u|} \right\rangle,$$

where angle brackets represent the inner product operator. The BOLD contrast for the j th condition, c_j , has units of percent modulation. We compared the shape of the response functions obtained with this BOLD contrast measure with the ones obtained with a standard percent signal change from fixation (Fig. S2).

The response functions (Figs. 3 and 4) plot the difference between the BOLD contrast in each condition and the full-noise response contrast, $c_j - c_1$ as a function of the inverse noise level (controlled by the phase-scrambling parameter s). Inverse noise level (noise level $^{-1}$) was chosen because rising response functions are easier to understand. The smooth curves (Fig. 3) are cumulative Gaussians fit to the mean data. The curves are fit using 5 parameters: to avoid overfitting, a common asymptote BOLD response and slope were found for all curves, and a separate Gaussian mean (horizontal shift) was determined for each curve. Variability of the mean (Fig. 3) was estimated by bootstrapping (Efron and Tibshirani 1993). We resampled with replacement (100 resamples) from the subject pool for each experiment. We then found the 5 parameters and computed the standard deviations of the Gaussian means.

To evaluate the homogeneity of the BOLD response within the pOTS ROI, we performed 2 further analyses. 1) A bootstrap analysis on voxel selection: subsets of N voxels (where N = original ROI size) were resampled iteratively from the ROI, and the mean BOLD contrast was computed for each noise level and each stimulus type. We computed the standard deviation of these means (i.e., the standard error). These are shown as the error bars in Figure 4. 2) For each subject who performed the 3 experiments, we split the ROI in half by the sagittal (left/right), coronal (anterior/posterior), and axial (superior/inferior) axes and repeated the analysis. We plotted and compared the response functions in these subregions to examine any systematic differences between them.

Results

Experiments 1–3

Subjects performed the fixation color-judgment task with high accuracy (higher than 90% in all conditions). Comparing BOLD signals during experimental word conditions versus fixation blocks produced a widespread and robust response along the ventral occipito-temporal cortex (Fig. 2). The BOLD response extended through the central representations of the visual field maps V1, V2 (ventral), hV4, and VO-1 (Brewer and others 2005). Consistent activations were found in the left pOTS, at an anatomical location similar to that of the VWFA reported by Cohen and others (2000, 2002, 2004, see Fig. 2). Additional activation was present also in right pOTS, bilateral intraparietal sulcus, and precentral sulcus. The spatial distribution shown in Figure 2 is typical of all of our subjects.

We next compared sensitivity to words with other stimulus types. Our analysis focused on the pOTS, based on previous

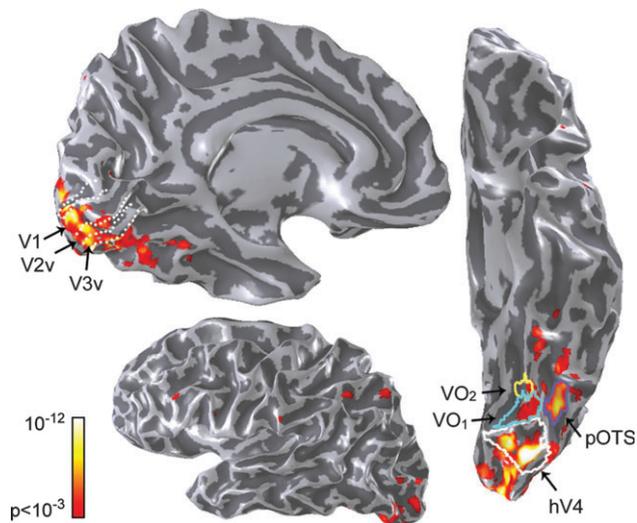


Figure 2. Localization map to define the pOTS. Activation map from a single representative subject (S6) comparing the word + noise conditions (excluding the full-noise condition) with fixation. The parametric map ($P < 0.001$, uncorrected) is overlaid on renderings of the gray-white matter surface of the left hemisphere. The surface is smoothed to expose activation within sulci; the dark and light shading indicates the position of major sulci and gyri. The borders of visual field maps were defined by standard retinotopic mapping experiments on the same subject. Visual field map borders presented: V1, V2v, V3v (dotted white), hV4 (solid white), VO-1 (cyan), VO-2 (yellow). The pOTS ROI (blue border) was defined by the activation cluster located at the intersection of the OTS with the inferior temporal sulcus.

findings for this region's specialization for visual word forms. We conducted an ROI analysis in left and right pOTS, and plotted the difference (Δ) in the mean BOLD response between each of the shape + noise conditions and the full-noise condition. ROIs were defined in individual subjects using the functional localizer (all word + noise conditions versus fixation, see Methods and Fig. 2), restricted by anatomical borders of the pOTS (see Methods). Mean MNI coordinates: left pOTS: $[-45, -65, -10]$; right pOTS: $[43, -67, -10]$; standard deviations: $[\pm 6, \pm 5, \pm 3]$; see Supplementary Figure S1 for the location of left pOTS and right pOTS ROIs in 3 individual subjects. The results of this analysis are presented in Figure 3.

Δ BOLD contrast values in both left and right pOTS varied systematically with noise level, saturating at the 2 lowest noise conditions (Fig. 3, full curves). Increasing responses were also measured for line drawings and false fonts (Fig. 3, dashed and dot-dashed curves). Highest sensitivity for words, intermediate sensitivity for line drawings, and lowest sensitivity for false fonts (variance of fitted curves was estimated using the bootstrap method; Efron and Tibshirani 1993, see Methods). Left and right pOTS showed the same relative sensitivity pattern but there was higher variability in right pOTS as shown by the larger error bars (Fig. 3, right panel).

The Δ BOLD contrast measure masks any baseline differences between subjects and between ROIs. To test for such baseline differences, we have repeated the analysis with an unscaled, standard measure of BOLD signal (% signal change from fixation). This analysis did not reveal any baseline difference between the word-response functions measured in left and right pOTS (Supplementary Fig. S2, left panel). However, there was a baseline difference in the line drawing and false-fonts curves, with higher activations in right pOTS than on the left

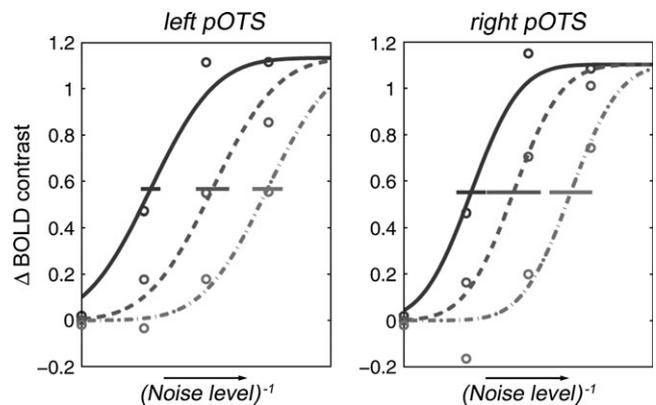


Figure 3. Response functions in pOTS. Mean Δ BOLD contrast in left pOTS and right pOTS is plotted as a function of inverse noise level. The Δ BOLD contrast is the difference between the shape + noise and noise alone. Circles represent the data, and separate curves were fitted for words (full curve, $N = 9$), line drawings (dashed curve, $N = 7$), and false-font strings (dot-dashed curve, $N = 6$). The fitted curves were constrained to have the same peak value and slope, differing only in horizontal position (sensitivity). Error bars represent the standard error of the sensitivity for each curve, computed by bootstrapping.

(Fig. S2, middle and right panels). In these cases, the shape of the response functions was similar (though shifted) in the left and right pOTS, but right pOTS baseline responses were higher to line drawings and false fonts compared with left pOTS. These curves imply that a direct contrast between words and line drawings would produce a left lateralized activation pattern in pOTS. Such a pattern would arise because left pOTS responds less to line drawings. In contrast, right pOTS responds equally high to words and line drawings.

Additional ROI analyses were carried out in left and right V1, to estimate the anatomical specificity of the pOTS response function. V1 was defined in individual subjects using the same functional localizer (as in Fig. 2), within anatomical boundaries of the posterior third of the calcarine sulcus. The activation level in left and right V1 remained high and unchanged for our amplitude-spectrum matched stimuli (see Supplementary Fig. S3). This finding demonstrates that the effects measured in pOTS are specific, and do not reflect general arousal effects present throughout visual cortex. Further whole-brain contrast maps testing the noise effect on line drawings and on false fonts provide additional support for this specificity (see Fig. S4). The constant response in V1 also suggests that for these stimuli V1 is driven mostly by the content of the amplitude spectrum, and not by phase information (see also: Oppenheim and Lim 1981; Morrone and Burr 1988).

Robust response functions were also measured in individual subjects. Figure 4 shows the response functions measured in left and right pOTS in 3 individual subjects who took part in all 3 experiments. Error bars were computed by resampling over the voxels in each ROI (see Methods). The small error bars reflect the stability of the response functions over different definitions of the ROI. Figure 4 also demonstrates higher consistency in response functions measured in left pOTS, compared with right pOTS.

We further examined the response functions in 6 sub-ROIs (created by splitting the left pOTS ROI along each of the 3 cardinal planes). Within subject, the curves were highly similar in all sub-ROIs. In particular, in the noisy but visible condition (third from left in Fig. 3), word responses were always higher

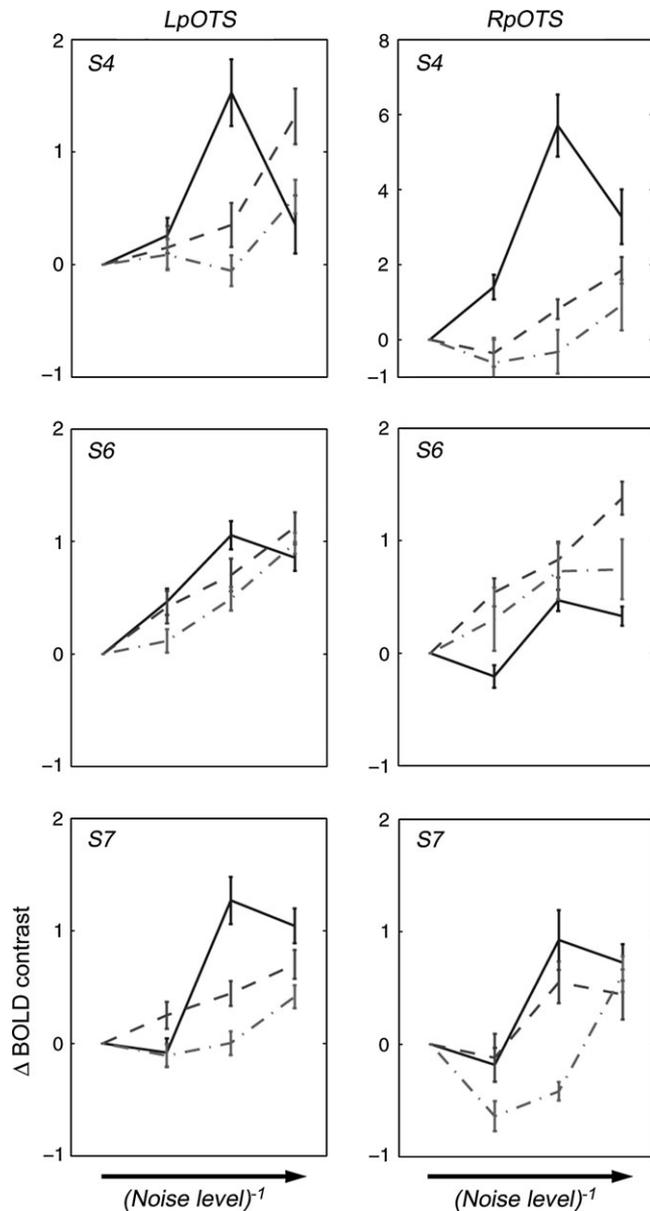


Figure 4. Response functions in pOTS for individual subjects. Mean Δ BOLD contrast is plotted for each noise level in 3 individual subjects (S4, S6, S7). Separate curves represent the responses to words (full), line drawings (dashed), and false fonts (dot dashed). Left panels: left pOTS; Right panels: right pOTS. Error bars were calculated by bootstrapping over the voxels within the ROI (see text).

than the responses to line drawings and false fonts. Furthermore, line drawings and false fonts produced rising response curves in 6/6 sub-ROIs in 2 of these subjects, 5/6 and 4/6 in the others.

We then explored the lateralization of word responses in individual subject data. In most subjects the shape of the response curves was similar in the 2 hemispheres, with a possible shift of the overall signal level favoring either the right hemisphere (S1, S4) or the left (S9). For each subject, we also computed a volume lateralization index, $LI = (L - R)/(L + R)$, comparing the left and right pOTS volumes. LI values ranged from -0.62 to 0.28 (mean LI = -0.091; mean LI for V1 is -0.0007 with a range of -0.43 to 0.38). There was no consistent difference in the volume of activated voxels in the left and

right pOTS (see Supplementary Fig. S5). Hence, the functional and anatomical properties of left pOTS and right pOTS appear similar at the spatial resolution of these measurements.

Experiment 4

The specificity of the pOTS responses to word stimuli was further examined in experiment 4. We compared pOTS responses to words, line drawings, and checkerboards presented to the right or the left of fixation (see Methods for details).

The results of experiment 4 are shown in Figure 5. We found powerful responses to parafoveal word stimuli (compared to parafoveal checkerboards) in left pOTS, in agreement with previous findings (Cohen and others 2000, 2002) (see Fig. 5A). In 5 out of 7 subjects, this comparison activated right pOTS as well. In all 7 subjects, there is a second activation anterior and medial to the pOTS, located near the fusiform gyrus. This activation is 1–3 cm anterior to the pOTS.

The mean BOLD contrast levels were similar in left and right pOTS (Fig. 5B), with the right activations being slightly higher. Activations were significantly higher for words and line drawings than for checkerboards in both hemispheres. There was no significant difference between words and line drawings. This pattern of results is robust to the pOTS selection procedure; using other localizers to select the pOTS produces the same pattern of results.

The pOTS responses were similar for left visual field (LVF) or right visual field (RVF) stimulus presentation. There was a slight trend in the right pOTS for an increased response when words and line drawings (but not checkerboards) were presented to the contralateral visual field.

Figure 5 further compares spatial activation maps in the left hemisphere for LVF and RVF word stimuli (Fig. 5C) and for LVF and RVF line drawings (Fig. 5D). The patterns of responses to words and line drawings were quite similar. Posterior ventral occipital cortex was mostly driven by contralateral words and line drawings. The pOTS responded to stimuli in both contralateral and ipsilateral visual field (Cohen and others 2000, 2002). An additional region, further anterior and medial on the fusiform gyrus, also responded consistently to stimuli in both hemifields.

The regions responding to stimuli in both hemifields (Fig. 5C,D) aligned very closely with the regions responding to the contrast between words and checkerboards (Fig. 5A). Similar activation patterns were found in the right hemisphere (see Supplementary Fig. S6), though right activations were only found in 5/7 subjects, and the size of the right hemisphere ROIs was smaller than the left in 4/5 subjects.

Discussion

Our results suggest a new perspective on word processing in the ventral occipito-temporal cortex, both in terms of 1) the relationship between shape and word processing and 2) the lateralization of written word processing.

Shape Processing and Word Processing in pOTS

BOLD responses in left and right pOTS (but not V1) increase as a function of shape visibility. This response pattern is seen with words, line drawings, and false-font strings (experiments 1–3). The increasing response for all stimulus types suggests that pOTS extracts shape information from all these stimuli.

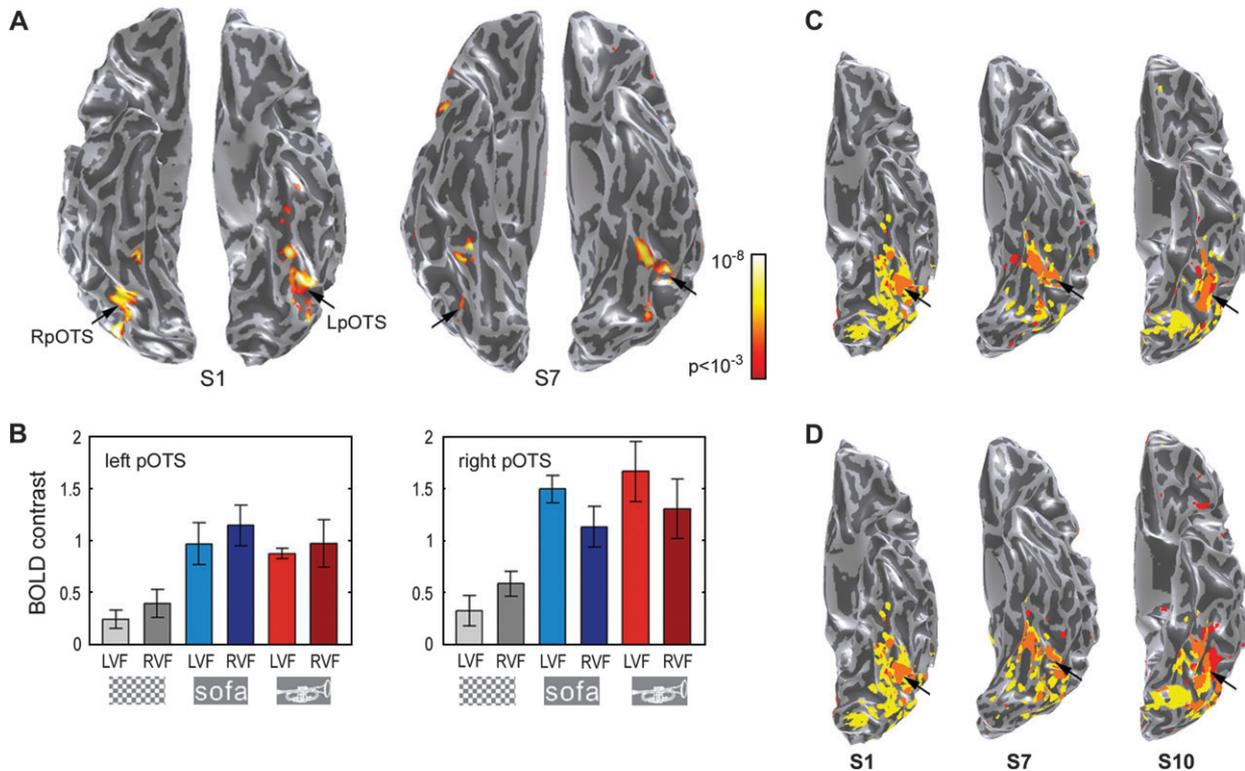


Figure 5. pOTS responses to parafoveal words and line drawings. (A) pOTS localizer. BOLD responses for words versus checkerboards (both hemifields) are shown on a ventral view of the brain for 2 typical subjects (S1, S7). The map shows locations with $P < 0.001$, uncorrected. (B) Stimulus type and hemifield effects in pOTS. Time courses from the activated voxels in left and right pOTS (for all words vs. all checkerboards, see panel A) were used to compute mean contrast values. These BOLD contrast values are presented in the bar graphs showing left ($N = 7$) and right ($N = 5$) responses. The bar graphs show mean BOLD response to checkerboards (gray), words (blue), and line drawings (red), in LVF (light shades) and in RVF (dark shades). Error bars are ± 1 standard error of the mean. For each ROI, a 2×3 within subject analysis of variance showed a significant main effect of stimulus category (left pOTS: $P < 0.0001$; right pOTS: $P < 0.005$), no main effect of hemifield and no interaction. Further planned contrasts showed significant effects in both ROIs for words > checkerboards (left/right pOTS: $P < 0.005$) and line drawings > checkerboards (left pOTS: $P < 0.0001$; right pOTS: $P < 0.02$), but not for words versus line drawings. (C) Overlapping responses to words presented in the contra- and ipsilateral visual fields. Thresholded parameter maps on the left ventral surface. The stimuli were presented in RVF or LVF; data are shown for 3 representative subjects (S1, S7, S10). Two maps were computed by contrasting responses to words in the RVF versus fixation (yellow), and words in LVF versus fixation (red). Overlap regions are shown in orange. Arrows mark the location of left pOTS (defined as in panel A). Maps were thresholded individually (S1: $P < 0.0001$; S7: $P < 0.01$; S10: $P < 0.005$, uncorrected). In each subject, the same threshold was used for all the maps. (D) Overlapping responses to Line drawings. Same conventions as in C. Overlapping responses for line drawings in left and right visual fields were found in pOTS, similar to C.

The pOTS sensitivity is significantly higher to words than to the other stimulus types. The word advantage in this area is graded (words > line drawings > false-font strings); there is no qualitative difference between the responses to words and other stimulus types, or between the responses to nameable stimuli versus false fonts. The increasing responses to false fonts in pOTS are hard to explain in terms of lexical access to a semantic or phonological representation.

We interpret the data as supporting a role of pOTS in visual shape extraction, not restricted to words. This sensitivity to shape information may be the reason why this region is recruited to support visual word recognition. The relative sensitivity to different stimulus types may be a consequence of the level of experience with each stimulus category but it may also stem from some unknown stimulus features that distinguish words from line drawings and false fonts.

It is possible that the pOTS ROI consists of smaller patches with higher category selectivity. At the current in-plane resolution (2.5×2.5 mm) and signal-to-noise, data become very unreliable unless at least 4–8 voxels are included in a region of interest. Hence, it would be impossible to find patches smaller than 25 mm^2 of cortical surface area. The total size of the pOTS ROI is on the order of 225 mm^2 . We examined the data for such inhomogeneity in 2 ways. We used a bootstrapping

procedure to resample the voxels within the pOTS ROI randomly. Random resampling did not change the basic outcome. Second, we divided the pOTS ROI into split-halves according to anatomical position. We found no difference in the response pattern when comparing the anterior–posterior portions, medial–lateral portions, or dorsal–ventral portions. Future studies with high-resolution fMRI will be able to identify the presence of any inhomogeneities at finer scales.

On the current view, pOTS lesions should interfere with general shape processing, not just word processing. A related line of study asserts that pure alexia, a visual word recognition deficit usually ascribed to ventral occipito-temporal lesions, is frequently accompanied by a more general shape-processing deficit. For example, Behrmann and others (1998) demonstrated a deficit in line-drawing identification in pure alexics with left occipito-temporal lesions. Response times to line drawings as a function of visual complexity was slowed disproportionately compared with controls. Other perceptual effects in pure alexia include an interaction between the word length effect and visual degradation of the stimulus (Farah and Wallace 1991), high proportion of visual errors in letter identification (Hanley and Kay 1992), and poor performance on perceptual fluency tests (Farah and Wallace 1991; Sekuler and Behrmann 1996; see also Friedman and Alexander 1984 for

poor performance on tachistoscopic identification of line drawings). Taken together, the neuropsychological and functional measurements support the hypothesis that pOTS serves a role in shape processing that includes, but is not restricted to, reading.

Lateralization of pOTS Responses

Our results differ from earlier reports that describe left lateralized word responses in the pOTS. The reports of a lateralized response were based on comparisons between words versus checkerboards (Cohen and others 2002), letters versus digits (Polk and others 2002), and words versus fixation (Cohen and others 2000, Fig. 3, top). Our parametric measurements show that the left and right pOTS functional responses are similar. Both left and right response functions increase for all 3 stimulus types. The relative sensitivity to words, line drawings, and false fonts is the same in left and right pOTS. Further, the volume analysis of pOTS ROIs showed a diverse pattern of lateralization, with no consistent size difference between left and right hemispheres (see Fig. S5).

We suggest several reasons for the differing results. First, when examining the data in Figure 3 note that the response to words and line drawings in the most visible (lowest noise) condition do not differ on the right but they do differ on the left. Hence, an analysis that compares only the response to zero noise stimuli would conclude (incorrectly) that the effect is limited to the left hemisphere. Second, lateralization differences may be explained by task differences. Left lateralized feedback signals from language cortex may enhance left pOTS responses. These language-related signals may increase in passive viewing experiments (e.g., Cohen and others 2000, 2002) in which subjects are likely to engage in lexical processing in word conditions more than in baseline conditions. An overt attention-demanding task, as in the current study, decreases the responses in left lateralized language regions and therefore decreases the asymmetry between left and right pOTS responses. This is in line with the low activations we observed for our word stimuli in frontal and temporal language regions (experiment 1, see Fig. 2).

The Development of Expertise for Letter Shapes in pOTS

As children become efficient at letter shape extraction, the relative responses to words and other stimuli may change. One might imagine that the shape-processing difference between V1 and pOTS, revealed by their different response functions to shapes in noise, emerges at very young ages; increasingly fine discriminations between words, line drawings, false fonts, and other shape stimuli may develop with increasing expertise at slightly older ages (Aghababian and Nazir 2000; McCandliss and others 2003). These predictions can be tested using longitudinal measurements of pOTS word-response functions in children.

Recent behavioral studies in developmental dyslexics documented perceptual deficits in detecting patterns within visual noise (Sperling and others 2005). Further, developmental dyslexics show reduced activations in left ventral occipito-temporal cortex for both words and line drawings in a region that is very likely the left pOTS (McCrorry and others 2005). These observations suggest that pOTS functionality is impaired in children with reading deficits. Two possible explanations for this linkage are 1) lack of sufficient reading experience in poor readers (of any etiology) could lead to improper development of the

specialization in pOTS; 2) a failure of pOTS to develop properly in terms of its functional properties and its connectivity with other regions may impede reading development (Shaywitz and others 2002; McCandliss and Noble 2003). Measurements of pOTS functionality and connectivity in young children with specific reading deficits will enhance our understanding of the causal role of pOTS in developmental dyslexia.

Supplementary Material

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

Notes

Funding provided by National Institutes of Health grant EY015000 and the Schwab Foundation for Learning. We thank Alex Wade, Alyssa Brewer, Arvel Hernandez, Junjie Liu, Kalanit Grill-Spector, Mark Eckert, Polina Potanina, Rachel Kalmar, Rory Sayres, Sing-Hang Cheung, and Serge Dumoulin for their help in various stages of the study. *Conflict of Interest:* None declared.

Address correspondence to Michal Ben-Shachar, Jordan Hall, Building 420, Stanford University, Stanford, CA 94305, USA. Email: michal@white.stanford.edu.

References

- Aghababian V, Nazir TA. 2000. Developing normal reading skills: aspects of the visual processes underlying word recognition. *J Exp Child Psychol* 76:123-150.
- Arguin M, Bub D, Dudgeon G. 1996. Shape integration for visual object recognition and its implication in category-specific visual agnosia. *Vis Cogn* 3:221-275.
- Avidan G, Harel M, Hendler T, Ben-Bashat D, Zohary E, Malach R. 2002. Contrast sensitivity in human visual areas and its relationship to object recognition. *J Neurophysiol* 87:3102-3116.
- Behrmann M, Nelson J, Sekuler EB. 1998. Visual complexity in letter-by-letter reading: "pure" alexia is not pure. *Neuropsychologia* 36:1115-1132.
- Boynton GM, Engel SA, Glover GH, Heeger DJ. 1996. Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16:4207-4221.
- Brewer AA, Liu J, Wade AR, Wandell BA. 2005. Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nat Neurosci* 8:1102-1109.
- Chao LL, Haxby JV, Martin A. 1999. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat Neurosci* 2:913-919.
- Cohen L, Dehaene S, Naccache L, Lehericy S, Dehaene-Lambertz G, Henaff MA, 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* 123:291-307.
- Cohen L, Jobert A, Le Bihan D, Dehaene S. 2004. Distinct unimodal and multimodal regions for word processing in the left temporal cortex. *Neuroimage* 23:1256-1270.
- Cohen L, Lehericy S, Chochon F, Lemer C, Rivaud S, Dehaene S. 2002. Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain* 125:1054-1069.
- De Renzi E. 2000. Disorders of visual recognition. *Semin Neurol* 20:479-485.
- Dehaene S, Cohen L, Sigman M, Vinckier F. 2005. The neural code for written words: a proposal. *Trends Cogn Sci* 9:335-341.
- Dejerine J. 1892. Contribution à l'étude anatomo-pathologique et clinique des différentes variétés de cécité verbale. *Mem Soc Biol* 4:61-90.
- Dougherty RF, Ben-Shachar M, Deutsch G, Potanina P, Bammer R, Wandell B. 2005. Occipital-callosal pathways in children: validation and atlas development. *Ann N Y Acad Sci* 1064:98-112.
- Efron B, Tibshirani RJ. 1993. An introduction to the bootstrap. London, UK: Chapman & Hall.

- Farah M. 2004. Visual agnosia. 2nd ed. Cambridge, MA: MIT Press.
- Farah MJ, Wallace MA. 1991. Pure alexia as a visual impairment: a reconsideration. *Cogn Neuropsychol* 8:313-334.
- Friedman R, Alexander MP. 1984. Pictures, images and pure alexia: a case study. *Cogn Neuropsychol* 9:1-23.
- Glover GH. 1999. Simple analytic spiral K-space algorithm. *Magn Reson Med* 42:412-415.
- Grill-Spector K, Kushnir T, Hendler T, Edelman S, Itzhak Y, Malach R. 1998. A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Hum Brain Mapp* 6:316-328.
- Hanley JR, Kay J. 1992. Does letter-by-letter reading involve the spelling system? *Neuropsychologia* 30:237-256.
- Hanson SJ, Matsuka T, Haxby JV. 2004. Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area? *Neuroimage* 23:156-166.
- Hasson U, Levy I, Behrmann M, Hendler T, Malach R. 2002. Eccentricity bias as an organizing principle for human high-order object areas. *Neuron* 34:479-490.
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425-2430.
- Horowitz SG, Rossion B, Skudlarski P, Gore JC. 2004. Parametric design and correlational analyses help integrating fMRI and electrophysiological data during face processing. *Neuroimage* 22:1587-1595.
- James TW, Culham J, Humphrey GK, Milner AD, Goodale MA. 2003. Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study. *Brain* 126:2463-2475.
- Kanwisher N, McDermott J, Chun MM. 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17:4302-4311.
- Kucera H, Francis WM. 1967. Computational analysis of present-day American English. Providence, RI: Brown University Press.
- McCandliss BD, Cohen L, Dehaene S. 2003. The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn Sci* 7:293-299.
- McCandliss BD, Noble KG. 2003. The development of reading impairment: a cognitive neuroscience model. *Ment Retard Dev Disabil Res Rev* 9:196-204.
- McCrary EJ, Mechelli A, Frith U, Price CJ. 2005. More than words: a common neural basis for reading and naming deficits in developmental dyslexia? *Brain* 128:261-267.
- Moore CJ, Price CJ. 1999. Three distinct ventral occipitotemporal regions for reading and object naming. *Neuroimage* 10:181-192.
- Morrone MC, Burr DC. 1988. Feature detection in human vision: a phase-dependent energy model. *Proc R Soc Lond B Biol Sci* 235:221-245.
- Oppenheim AV, Lim JS. 1981. The importance of phase in signals. *Proc IEEE* 69:529-541.
- Polk TA, Stallcup M, Aguirre GK, Alsup DC, D'Esposito M, Detre JA, Farah MJ. 2002. Neural specialization for letter recognition. *J Cogn Neurosci* 14:145-159.
- Price CJ, Devlin JT. 2003. The myth of the visual word form area. *Neuroimage* 19:473-481.
- Price CJ, Devlin JT. 2004. The pro and cons of labelling a left occipitotemporal region: "the visual word form area". *Neuroimage* 22:477-479.
- Puce A, Allison T, Asgari M, Gore JC, McCarthy G. 1996. Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *J Neurosci* 16:5205-5215.
- Salvan CV, Ulmer JL, DeYoe EA, Wascher T, Mathews VP, Lewis JW, Prost RW. 2004. Visual object agnosia and pure word alexia: correlation of functional magnetic resonance imaging and lesion localization. *J Comput Assist Tomogr* 28:63-67.
- Sekuler E, Behrmann M. 1996. Perceptual cues in pure alexia. *Cogn Neuropsychol* 13:941-974.
- Shaywitz BA, Shaywitz SE, Pugh KR, Mencl WE, Fulbright RK, Skudlarski P, Constable RT, Marchione KE, Fletcher JM, Lyon GR, and others. 2002. Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biol Psychiatry* 52:101-110.
- Sperling AJ, Lu ZL, Manis FR, Seidenberg MS. 2005. Deficits in perceptual noise exclusion in developmental dyslexia. *Nat Neurosci* 8:862-863.
- Spiridon M, Kanwisher N. 2002. How distributed is visual category information in human occipito-temporal cortex? An fMRI study. *Neuron* 35:1157-1165.
- Szekely A, Jacobsen T, D'Amico S, Devescovi A, Andonova E, Herron D, Lu CC, Pechmann T, Pleh C, Wicha N, and others. 2004. A new online resource for psycholinguistic studies. *J Mem Lang* 51:247-250.
- Teo PC, Sapiro G, Wandell BA. 1997. Creating connected representations of cortical gray matter for functional MRI visualization. *IEEE Trans Med Imaging* 16:852-863.
- Tranel D, Damasio H, Damasio AR. 1997. A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia* 35:1319-1327.
- Turkeltaub PE, Gareau L, Flowers DL, Zeffiro TA, Eden GF. 2003. Development of neural mechanisms for reading. *Nat Neurosci* 6:767-773.
- Vigneau M, Jobard G, Mazoyer B, Tzourio-Mazoyer N. 2005. Word and non-word reading: What role for the Visual Word Form Area? *Neuroimage* 27:694-705.
- Warrington EK, Shallice T. 1980. Word-form dyslexia. *Brain* 103:99-112.