

# Position sensitivity in the visual word form area

Andreas M. Rauschecker<sup>a,b,1</sup>, Reno F. Bowen<sup>b</sup>, Josef Parvizi<sup>c</sup>, and Brian A. Wandell<sup>a,b</sup>

<sup>a</sup>Neurosciences Program and Medical Scientist Training Program, and <sup>c</sup>Laboratory of Behavioral and Cognitive Neurology, Department of Neurology and Neurological Sciences, Stanford School of Medicine, Stanford, CA 94305; and <sup>b</sup>Department of Psychology, Stanford University, Stanford, CA 94305

Edited by Michael Posner, University of Oregon, Eugene, OR, and approved April 11, 2012 (received for review December 28, 2011)

Seeing words involves the activity of neural circuitry within a small region in human ventral temporal cortex known as the visual word form area (VWFA). It is widely asserted that VWFA responses, which are essential for skilled reading, do not depend on the visual field position of the writing (position invariant). Such position invariance supports the hypothesis that the VWFA analyzes word forms at an abstract level, far removed from specific stimulus features. Using functional MRI pattern-classification techniques, we show that position information is encoded in the spatial pattern of VWFA responses. A right-hemisphere homolog (rVWFA) shows similarly position-sensitive responses. Furthermore, electrophysiological recordings in the human brain show position-sensitive VWFA response latencies. These findings show that position-sensitive information is present in the neural circuitry that conveys visual word form information to language areas. The presence of position sensitivity in the VWFA has implications for how word forms might be learned and stored within the reading circuitry.

tolerance | retinotopy | perception | vision | functional MRI

In humans, ventral occipitotemporal (VOT) cortex, specifically the visual word form area (VWFA) (1), is an essential part of the neural circuitry for seeing words (2, 3). The VWFA is located just posterior to the fusiform face area and lateral to the ventral occipital (VO) visual field maps, VO-1 and VO-2 (4). The VWFA responds powerfully during reading tasks; these responses develop in synchrony with learning to read (5, 6) and are relatively weak in poor readers (2, 7). Lesions of the VWFA can produce reading deficits (8).

Early experiments suggested that VWFA responses are invariant to the visual field position of the stimulus (position invariant) and left-hemisphere lateralized (1, 9). This view has been accepted ubiquitously (10–30). Position invariance is used to support the hypothesis that the VWFA is dedicated to analyzing word forms while being invariant to stimulus-dependent features such as visual field position, fonts, letter case, and size (13). The principle that the neural responses to word forms are position invariant parallels the hypothesis that object representations in human and nonhuman primate cortex are invariant and learned (31). For reading, too, invariance to certain visual features may be beneficial (32). Hence, the first descriptions of the VWFA supported the idea of a general visual neural circuitry that recognizes objects across the visual field, abstracted from position (33).

The issue of position sensitivity can be divided into two types: the relative location of letters within a word (or words within a sentence) and the absolute location of letters and words within the visual field. Relative letter and word position is essential for reading. If no absolute visual field position information is present, the neural circuitry could encode relative position implicitly by representing ordered letter strings (bigrams, trigrams, and so forth) (34, 35). On the other hand, if the cortical circuitry includes visual field position information, other cortical regions can decode the relative position. It is also possible that both visual field position information and ordered letter strings are encoded, but at different spatial scales.

Visual field (retinotopic) organization in the VWFA has not been reported using standard methods, leaving open the possibility that absolute position information is indeed discarded. The absence of such position information in the VWFA contrasts,

however, with recent measurements of retinotopic representations in VOT cortex surrounding the VWFA (figure 1 in ref. 2). The hemifield maps surrounding the VWFA include VO-1/2 (medial to VWFA) (36), parahippocampal cortex (PHC-1/2, anterior to VWFA) (37), lateral occipital cortex (LO-1/2, lateral and posterior to VWFA) (38), and temporo-occipital cortex (TO-1/2, lateral to VWFA) (39). Many of these retinotopically organized areas possess neural specializations. For example, the LO maps are known to be sensitive to the structure of complex visual objects (40, 41), and the TO-1/TO-2 maps overlap directly with the human MT complex known to be sensitive to visual motion (39).

Given that the VWFA is surrounded by retinotopically organized visual field maps (Fig. 1), we decided to reconsider the question of VWFA retinotopic organization using more sensitive multivoxel pattern classification techniques and electrophysiological recordings in the human brain. These measurements offer a more thorough assessment of position sensitivity in the VWFA.

## Results

**Location of VWFA Relative to Retinotopic Maps.** We identified ventral occipital visual field maps (V1, V2, V3, hV4, VO-1, and VO-2) and the VWFA in individual subjects using functional magnetic resonance imaging (fMRI) retinotopic mapping and a VWFA localizer scan, respectively (*Methods*). To remove any position-sensitivity effect of voxels within known retinotopic areas in our analyses, the VWFA region of interest (ROI) was restricted to voxels outside of all known retinotopic areas and anterior to hV4 (Fig. 1). In all subjects, the vast majority of voxels significantly responsive to the VWFA localizer contrast (words > phase-scrambled words) were anterior or lateral to the visual field maps VO-1 and VO-2, typically in the left lateral occipitotemporal sulcus [individual Montreal Neurological Institute (MNI) coordinates in Table S1; mean MNI coordinates: (−42.9, −56.9, −23.0)].

In six of seven subjects, the contrast (words > phase-scrambled words) also produced a collection of significant voxels at the corresponding location in the right hemisphere [mean MNI coordinates: (41.3, −60.3, −20.3)]. The same selection criteria with respect to visual field maps were used to define the right-hemisphere homolog of the VWFA, which we refer to as the rVWFA.

**Spatial Pattern of VWFA Response Amplitudes Varies with Stimulus Position.** We measured fMRI responses in VOT cortex to words at multiple visual field positions. A support vector machine (SVM) based upon a linear pattern classifier (42) was trained on a portion of the data in each subject to distinguish between words in the left vs. right visual fields (VFs) or the upper vs. lower VFs. Decoding performance was evaluated on the remaining data. This procedure was repeated for data from regions in early visual

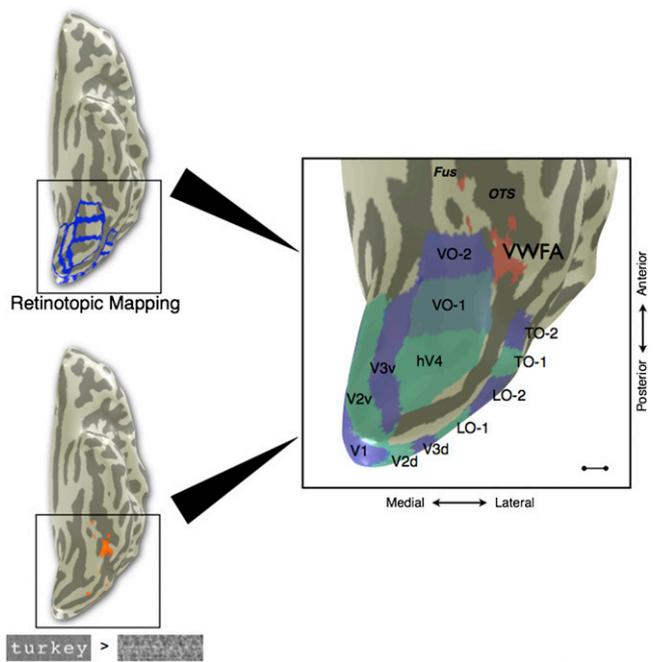
Author contributions: A.M.R. and B.A.W. designed research; A.M.R., R.F.B., and J.P. performed research; A.M.R., R.F.B., and J.P. contributed new reagents/analytic tools; A.M.R., R.F.B., and B.A.W. analyzed data; and A.M.R. and B.A.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>1</sup>To whom correspondence should be addressed. E-mail: andreas@stanford.edu.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1121304109/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1121304109/-DCSupplemental).



**Fig. 1.** Location of VVFA in relation to the visual field maps in a typical subject. (Upper Left) fMRI retinotopic mapping was performed to define visual field map boundaries (blue) for V1, V2, V3, hV4, VO-1, VO-2, LO-1, LO-2, TO-1, and TO-2. (Lower Left) Separate runs of the VVFA localizer were performed, and voxels were chosen on the basis of the contrast words > phase-scrambled words at a threshold of  $P < 0.001$  (uncorrected, orange overlay). VVFA voxels were restricted to be outside of known retinotopic areas and anterior to hV4 in ventral occipito-temporal cortex. (Right) The cortical surface of the posterior left hemisphere is rendered as a mesh and computationally inflated to visualize the sulci (dark gray). The majority of VVFA voxels (red) are located lateral or anterior to VO-1/VO-2 in all subjects. The TO-1/TO-2 maps overlap with the human motion complex (hMT+). Fus, fusiform gyrus; OTS, lateral occipitotemporal sulcus. (Scale bar: 10 mm.) Subject is S1.

cortex (V1/V2), the VVFA, the rVVFA, and Broca's area. Data from V1/V2 served as a positive control, because early visual areas are known to be retinotopically organized and therefore provide a good approximation of ceiling performance for the classifier.

As expected, the classifier accurately decodes VF position information from V1/V2 data (Fig. 2). Mean decoding accuracy is 94% for left–right VF classification and 92% for upper–lower VF classification, well above the chance level of 50%. Classification accuracy in every subject exceeds the significance threshold of 62.5% (Methods).

Classifier performance based on VVFA responses provides strong evidence that visual field position is also encoded in the VVFA. Mean decoding accuracy is 76% for left–right and 74% for upper–lower VF classification, and again VF classification accuracy is significantly above chance in every subject (Fig. 2). Decoding performance based on rVVFA responses is significantly above chance as well (mean classification for left–right VF, 79%; mean classification for upper–lower VF, 68%). Between-subject variability in classification performance in rVVFA is higher than for the VVFA.

Combining VVFA and rVVFA data increases classification accuracy to a level higher than either ROI alone (Fig. 2). There is an especially large increase in accuracy for the left vs. right VF classification (77% for VVFA, 79% for rVVFA, and 92% for VVFA and rVVFA voxels combined), suggesting that these ROIs contain complementary visual position information and together may form a complete retinotopic map. The more modest increase in decoding accuracy for the upper–lower VF classifica-

tion (75% for VVFA, 68% for rVVFA, and 78% for combined ROI) is likely explained by having more relevant features (voxels) available to the classifier, rather than a complementary mapping of VF position.

Classification of VF position using data from Broca's area is at chance (50% for left–right classification and 49% for upper–lower classification), suggesting that VF position information is lost by the time it reaches this level of processing.

Searchlight analyses (43) in individual subjects confirm these findings. We conducted the searchlight procedure using small disk-shaped (5 mm radius) ROIs defined on the cortical gray matter surface to avoid pooling data that cross sulcal boundaries or extend into white matter (44). Even with such small ROIs, high decoding accuracy for VF position extends into VOT cortex and the VVFA (Fig. S1).

#### Position Information Is Retinotopically Organized Within the VVFA.

To identify the spatial pattern of voxel responses that the classifier uses to differentiate stimuli in the upper and lower VF, we mapped the classifier weights onto the cortical surface. These classifier maps indicate the influence of each voxel on the classifier's VF prediction. In V1/V2, positive and negative weights (corresponding to predictions of upper and lower VF, respectively) are well segregated on the cortical surface (Fig. 3). Voxels with positive weights (red) are located along the V1/V2v boundary, corresponding to the vertical meridian representation for the upper VF. Similarly, voxels with negative weights (blue) are located along the V1/V2d boundary, corresponding to the vertical meridian representation for the lower VF. The position of the largest weights shifts anteriorly along the V1/V2 boundary when classifying words presented at different visual field eccentricities (1°, 3°, or 6° away from fixation). This pattern is expected on the basis of the well-established retinotopic maps in V1 and V2 (2, 45).

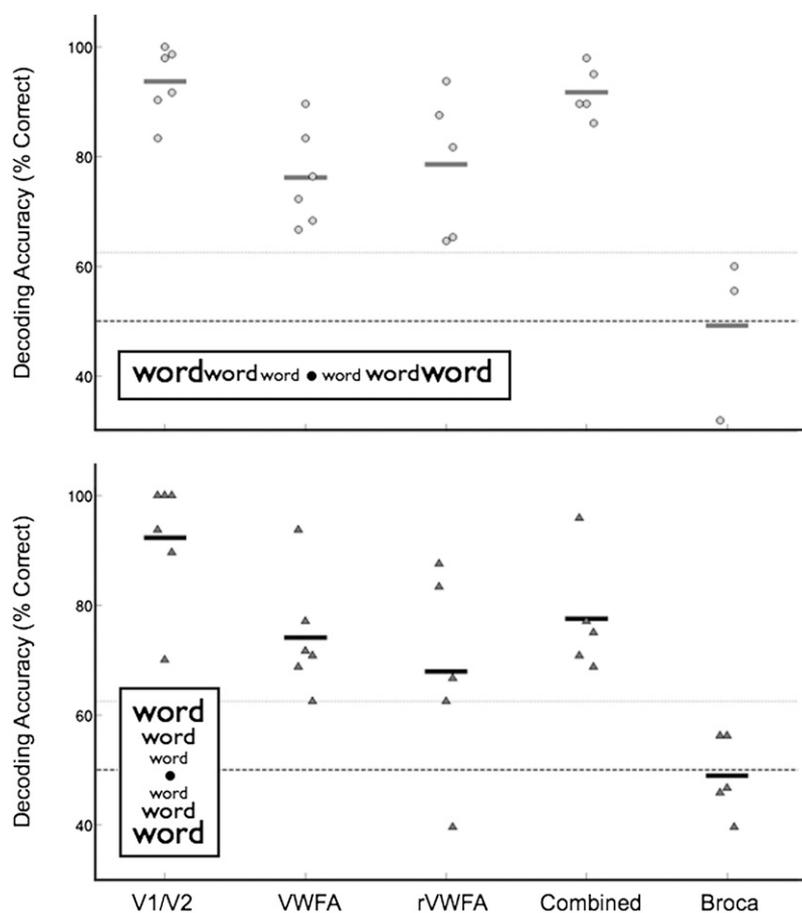
The VVFA upper–lower classifier map in some subjects is also spatially organized. Classifying between all upper and all lower VF positions, the classifier map has positive and negative weights clustered in anterior and posterior VVFAs, respectively. When comparing upper vs. lower classifier maps from the smaller set of stimuli at individual eccentricities (1°, 3°, or 6°), the same trend can be detected.

We measured the upper–lower VVFA classifier maps in each subject (Fig. S2). The VVFA classifier maps in some subjects contain one cluster of positively and one cluster of negatively weighted voxels, as the subject in Fig. 3. Other subjects' VVFAs contain several clusters of each sign. With current signal-to-noise and spatial resolution, VVFA classifier maps are not as reliable as classifier maps derived from early visual cortex data.

Further evidence for retinotopic organization in the VVFA is present in the pattern of misclassifications when classifying between all six possible spatial positions. If the VVFA visual field representation is retinotopic, then classification errors should follow a predictable pattern: Stimuli neighboring in visual space should be misclassified more often than stimuli that are far apart in visual space. The confusion matrix of classifications across all subjects shows exactly this pattern (Fig. 4). For V1/V2 there are very few errors at all. For the VVFA, there are more errors but these rarely confuse stimuli that are far apart in the visual field. For example, stimuli presented 6° in the lower VF (Fig. 4, Right, bottom row) are sometimes labeled as being 3° in the lower VF, but they are infrequently classified as falling in the upper VF.

The map seems to be sharpest near the fovea, with the best classification occurring for stimuli 1° in the upper or lower VF. The decreased blurring nearest the fovea may be a reflection of cortical magnification (46).

The difference in classification accuracy between V1/V2 and the VVFA appears to arise both because the VVFA covers a smaller cortical surface area and because it has noisier responses. We can test the hypothesis that VVFA classification errors are



**Fig. 2.** Classification of visual field position using VWFA responses. The classifier was trained using data from several regions of interest (ROI): V1/V2, VWFA, rVWFA, and Broca's area. The "combined" region of interest is the union of voxels from VWFA and rVWFA. Classification accuracy was measured in individual subjects. When tested on data from untrained trials, the classifier performance is above chance for left vs. right (*Upper*) and upper vs. lower (*Lower*) visual field (VF) classification in all regions except Broca's area. The short horizontal lines indicate the mean classifier performance within an ROI and across subjects. The dashed line indicates chance performance (50%), and the line with light shading indicates the significance level ( $P < 0.05$ ) derived by a bootstrap procedure. In some sessions, because of the MR slice prescription, a Broca's area ROI was not measured.

explained entirely by noise through simulation. We added Gaussian noise at a level that equates (same fraction of correct classifications) the V1/V2 performance with the VWFA performance in each subject (Fig. S3). In each subject, we calculate the confusion matrix and average these matrices across subjects. Adding noise produces errors that are distributed across space and do not reflect retinotopic organization. The proportion of classifications assigned to visual field locations adjacent to the correct location is  $0.25 \pm 0.088$  (99% bootstrap confidence interval) in the noisy V1/V2 simulation, which is smaller than the fraction assigned to adjacent visual field locations (0.34) from VWFA data. Thus, at least some fraction of VWFA classification errors can be explained by the compression of the retinotopic map into a smaller cortical space, rather than noisy signals.

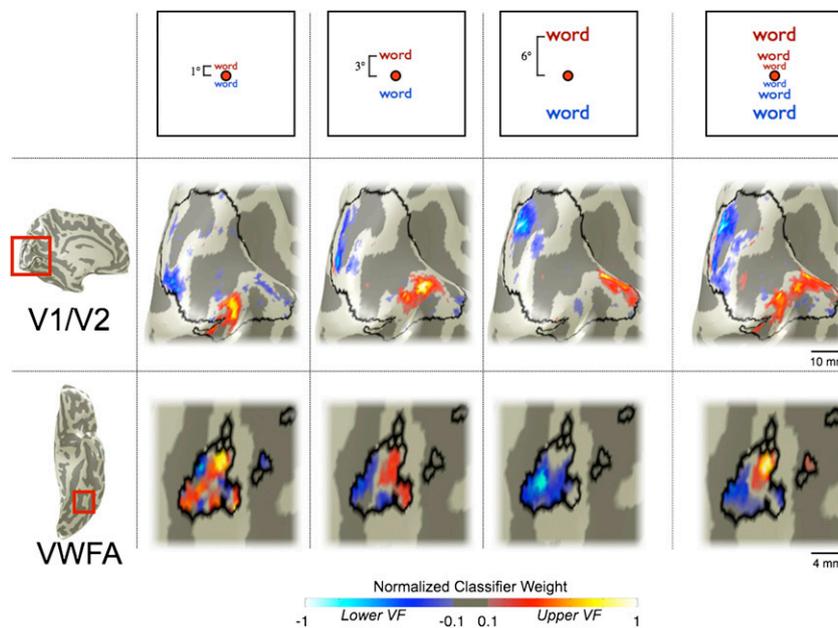
**VWFA and rVWFA Are More Responsive to Contralateral than Ipsilateral Stimuli.** The high classification accuracy for left–right VF is explained by the difference in the mean response amplitudes in the VWFA and its right-hemisphere homolog, the rVWFA. The responses in these two regions are of equal magnitude but with opposing visual hemifield preferences. The response modulations in both regions are approximately twice as large for contralateral (1.2%) compared with ipsilateral stimuli (0.6%) (Fig. 5). Note that with a conventional region-of-interest analysis, as in Fig. 5, these differences in response amplitude to words in the left and right visual fields can be observed.

We defined the lateralization index (LI) for each area, within each subject, as

$$LI = 1 - (R_I/R_C),$$

where  $R_I$  is the response [mean general linear model (GLM)  $\beta$ -weight] to ipsilateral stimuli and  $R_C$  is the response to contralateral stimuli. This index is a number between 0 and 1, where 0 indicates no difference in response to contralateral vs. ipsilateral stimuli, and 1 indicates no ipsilateral (or complete contralateral) responsiveness. The mean LI across subjects for the VWFA is 0.48 (SEM = 0.08), indicating that the response to contralateral stimuli is approximately twice as strong as to ipsilateral stimuli. The mean LI for the rVWFA is 0.51 (SEM = 0.11), and there is no significant difference between the LIs for the VWFA and rVWFA across subjects ( $t$  test,  $t = 0.25$ ,  $P = 0.81$ ).

The VWFA and the rVWFA both respond to words in the upper and lower VF. However, the response amplitudes to these stimuli do not differ as they do for the left–right VF difference (Fig. 5). Response amplitudes to upper–lower VF stimuli are similar in the two ROIs, reaching a maximum amplitude of  $\sim 1\%$  for words in either the upper or the lower VF. Thus, for successful classification of upper–lower VF position, the pattern classifier must rely on the higher-resolution spatial pattern of voxel responses within each ROI, rather than the mean amplitude of responses across voxels across the ROI.



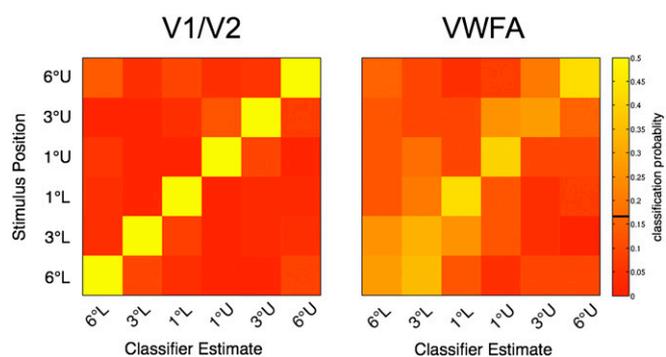
**Fig. 3.** Retinotopic organization of classifier weight maps in single subjects. Voxels are colored according to the weight assigned by the classifier. Positive weights (corresponding to voxels that identify upper VF stimuli) are red, and negative weights (corresponding to voxels that identify lower VF stimuli) are blue. Brightness represents the weight (see color bar). Values are normalized within each region of interest and thresholded at 10% of the maximum value. The columns show the classifier weight maps for words at 1°, 3°, and 6° above or below fixation separately. (Top) Stimulus eccentricity. (Middle) Maps from V1/V2. (Scale bar, 10 mm.) (Bottom) Maps from VWFA. (Scale bar, 4 mm.) The classifier weight map for differentiating all upper from all lower VF stimuli is shown in the rightmost column. (Middle row, subject S2; Bottom row, subject S3)

#### VWFA Response Latencies [electrocorticographic (ECoG) recordings] Are Delayed to Ipsilateral Compared with Contralateral Stimuli.

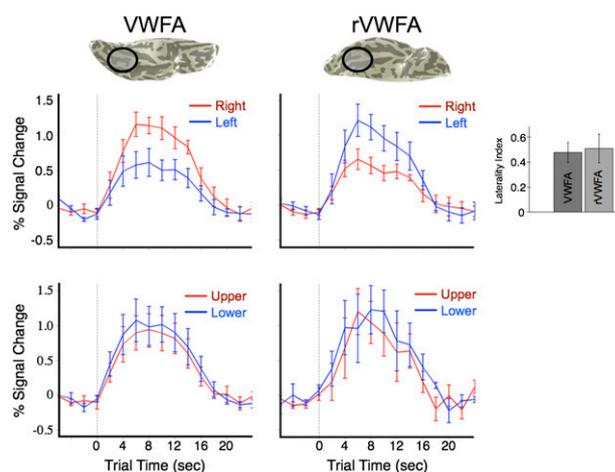
Precise response timing information can add significantly to our understanding of the reading circuitry. For example, if a cortical region shows substantially different response latencies to different stimuli, it can be inferred that these stimuli are being routed through different intermediate circuitry. Hence, we made a set of ECoG recordings to measure response timing to letter strings presented to several locations to the left and right of fixation.

Subdural electrodes were implanted in a patient as part of routine neurosurgical evaluation for epilepsy. Before the implant, the patient's VWFA was determined by a functional MRI localizer (Fig. 6A). In an electrode overlying the VWFA, letter strings produced a very high amplitude response in the high-gamma band (80–150 Hz; Fig. S4A). Neighboring electrode responses to the same stimuli are much weaker.

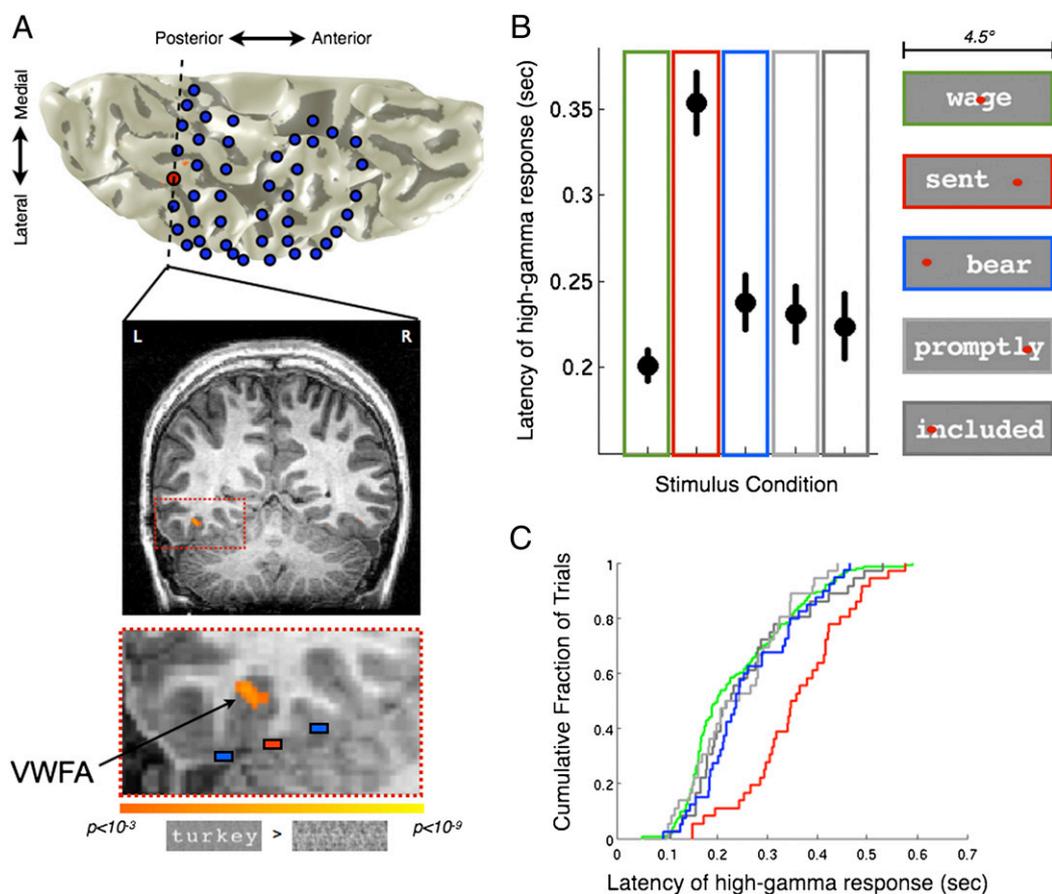
We calculated the VWFA response latency (stimulus onset to half peak power in high-gamma band) to letter string stimuli for



**Fig. 4.** Confusion matrices support a retinotopic map in VWFA. The rows represent the true stimulus position; the columns represent the classified position. The color indicates the probability of measuring a stimulus-classification pair across all trials of that stimulus position. Correct classifications are on the diagonal and misclassifications are off the diagonal. The confusion matrices shown are the average of the individual subject confusion matrices ( $n = 6$ ). (Left) V1/V2 classifier performance is very high. (Right) VWFA classification performance is lower, but misclassifications are largely confined to neighboring stimulus positions. U, upper VF; L, lower VF. One degree, 3°, and 6° refer to distances from the center of fixation. The color bar at the right represents the classification probabilities, and the black line indicates the chance level (16.7%).



**Fig. 5.** VWFA and rVWFA time course amplitudes are larger for contralateral than for ipsilateral word stimuli. (Left and Right) fMRI time courses from the VWFA and the rVWFA, respectively. (Upper) Responses to words in the left (blue) and right (red) VF. (Lower) Responses to words in the upper (red) and lower (blue) VF. Both the VWFA and the rVWFA respond preferentially to contralateral stimuli. The laterality index (main text) is shown in the *Inset* at the *Right*. There is no significant response difference between words in the upper and lower VF. Error bars are  $\pm 1$  SEM, computed across subjects.



**Fig. 6.** Electrocorticographic VWFA responses to ipsilateral stimuli are delayed compared to responses to contralateral stimuli. (A) A 3D reconstruction showing the VWFA and electrode positions. The patient's left hemisphere is shown from a ventral view. Blue disks indicate estimated electrode positions. The red disk is the electrode overlying the VWFA. The dashed line is a coronal plane, shown below. Functional MRI data contrasting words > phase-scrambled words (threshold  $P < 0.001$ , uncorrected; orange-yellow) reveal the VWFA location in this subject. The third image (Bottom) is a magnified view showing the electrode of interest (red), two neighboring electrodes (blue), and the VWFA. See also Fig. S4B. (B) Median VWFA response times to letter strings in different visual field locations. The graph shows the median response onset latency of the high-gamma signal ( $\pm 1$  SEM) across trials ( $n = 180$  for green foveal condition;  $n = 40$  for all other conditions). Colored outlines indicate the stimulus position relative to fixation, shown at the Right. Conditions included words, pseudowords, or consonant strings. (C) Cumulative distribution of single trial responses. The cumulative fraction of trials (y axis) is shown as a function of response onset latency (x axis). The curve for ipsilateral stimuli (red) is substantially to the right (delayed) compared with all other conditions. Stimulus conditions: green, foveal, four-letter, six-letter, and eight-letter; red, ipsilateral, four-letter; blue, contralateral, four-letter; light gray, ipsilateral with foveal component, eight-letter; dark gray, contralateral with foveal component, eight-letter.

every trial. For foveally presented words, the median latency is  $\sim 200$  ms (Fig. 6B), consistent with previous reports from magnetoencephalography (47) and subdural electrode recordings (48). However, there is a very large difference between the response latencies to letter strings in the ipsilateral VF and stimuli at other locations. The median ipsilateral response is more than 100 ms slower than the median contralateral response (Fig. 6B). The median contralateral response latency is also slower (20 ms) than the latency to stimuli presented in the fovea. The distribution of VWFA response latencies across trials varies from 150 to 450 ms, but the ipsilateral response latency distribution is clearly separated from the distribution derived from other types of trials (Fig. 6C).

The response amplitude is similar for ipsi- and contralateral stimuli and there is no correlation between response amplitude and response latency ( $P = 0.4$ ). The ECoG high-gamma response is simply shifted in time (Fig. S4A). Another measure of response latency, the latency to pass 10 standard deviations (SDs) of the pretrial baseline period power, also clearly separates the ipsilateral and contralateral response latencies (Fig. S4C). Although the absolute response latency depends on the

criterion measure, the distribution of ipsilateral responses is always much slower than that of contralateral responses.

Of note, the response timing effect is independent of whether the task was a lexical decision task (as shown here) or a color change fixation task. In the color change fixation task, letter string processing is incidental and unrelated to the task. The high-gamma band activity is sustained for a longer period in the lexical decision task. However, there is a large relative timing onset latency during both tasks, suggesting that the temporal delay to ipsilateral compared with contralateral word stimuli is likely due to bottom-up processing.

## Discussion

Initial measurements suggested that absolute word position information is discarded at the level of the VWFA (1, 9). Using higher-resolution methods and more sensitive pattern classification analyses, we demonstrate that the VWFA is position sensitive. The VWFA information about visual field position of words can be decoded using pattern-classification techniques in individual subjects (Fig. 2). This information is contained in the spatial pattern of blood oxygen level dependent (BOLD) amplitude responses (Fig. 3). Classification of stimulus position is possible for words in the right vs. left (contralateral vs. ipsilateral) VFs, as well as for words in

the upper vs. lower VFs (Fig. 2 and Fig. S1). Furthermore, electrophysiological responses within the VWFA to stimuli presented in the right and left visual fields have large latency differences. Hence, position sensitivity is present in the neural circuitry that conveys visual word form information to language areas. This observation has implications for how word forms might be learned and stored within the reading circuitry.

**Retinotopic Organization in the VWFA.** Our experimental design used high-resolution fMRI scans aimed at finding voxel-wise patterns within specific functional areas, such as V1/V2 or the VWFA. When the classifier weights are projected onto the cortical surface, they are organized into classifier maps that conform to the expected retinotopic organization in V1/V2. In some subjects the VWFA classifier maps are also organized retinotopically (Fig. 3), but we do not find clear retinotopic organization in the VWFA of every subject (Fig. S2).

The variability in VWFA classifier maps across subjects may reflect the difficulty of making complete measurements in this part of the VOT cortex (2). The VWFA is surrounded by signal dropout regions, due to the transverse dural venous sinus (49) and the auditory canals. The precise spatial relationship between these anatomical landmarks and the VWFA varies across subjects, so it is likely that different parts (and different amounts) of the VWFA are obscured in different subjects. In addition, the classifier map we estimate for each subject is simply the one with the highest decoding accuracy, with no other constraints. It is possible that there are several other solutions with a similar degree of accuracy but that are more “map-like”. In the future, analyses might be developed to allow us to find these alternative solutions without much loss of classification accuracy (50).

In addition to the VWFA classifier maps (Fig. 3), however, there is a second line of evidence that suggests that VWFA position information is in the form of a retinotopic map. The misclassifications are likely to be made into adjacent visual field positions when classifying between six unique spatial locations (Fig. 4). The pattern of misclassifications suggests that nearby visual field positions are coded in nearby regions within the VWFA.

**VWFA and rVWFA Contain Complementary Representations of Visual Space.** Previous claims of position invariance were based largely on contrasts across subjects showing a left-lateralized VWFA response to words > checkerboards presented to the left or right VFs (9). Whereas we confirm that the left VWFA does respond to ipsilateral letter strings, we also find a strong contralateral bias in the VWFA response within and between subjects (Fig. 5). In fact, the classification performance using VWFA/rVWFA data is nearly as accurate as classification using V1/V2 data (Fig. 2). These results suggest complementary coding of contralateral visual hemifields in homologous visual areas, VWFA and rVWFA.

In recording ECoG data, we identify the VWFA by means of the same standard fMRI localizer, so that we can ensure that we are recording from the same functional area. In both the fMRI (Fig. 5) and the ECoG (Fig. S4) data, the VWFA does respond to ipsilateral word stimuli. However, the ECoG response latency measurements indicate that the ipsilateral response is delayed by ~100 ms (Fig. 6B). Such latency differences are not easily measured with fMRI because of its low temporal resolution. Electrophysiological recordings in the human brain allow precise temporal measurements in response to words, and these recordings show a large hemifield-dependent effect on response latencies in the VWFA.

**Cortical Circuitry for Seeing Words.** These results constrain the possible systems-level circuitry for the early cortical stages of reading. Traditionally, both hemispheres' V1 through hV4 are thought to process word stimuli from the opposite visual hemifield (1). At that stage, signals from right hV4 are purported to

cross to the left VWFA (10, 13). This circuitry cannot account for a 100-ms delay in signal timing. To account for the delay, there must be additional processing for left VF stimuli before the information reaches the left-hemisphere VWFA.

Combining our observations from fMRI and ECoG, we build on a previous model of word processing, the local combination detector (LCD) model (11), and propose that the two hemispheres each process signals from opposite hemifields in distinct functional modules (represented by retinotopic maps) up to and including the VWFA and the rVWFA (Fig. 7). The VWFA and the rVWFA convey visual word form information to language areas. Signals may cross from the rVWFA to the VWFA via the corpus callosum, or the rVWFA may have direct connections to language areas (Fig. 7B). A tracing study in the human brain supports monosynaptic connections from right VOT cortex to Broca's and Wernicke's areas (51). Diffusion imaging may help resolve these alternative possibilities.

The source of visual field position information in the VWFA is likely shared with multiple other cortical areas. Specifically, it is likely that the position information is inherited from signals arising in nearby visual field maps, such as hV4 and VO. In addition, the information is likely communicated between other cortical regions, such as parietal cortex, on the basis of the nature of the anatomical connections of the VWFA (4).

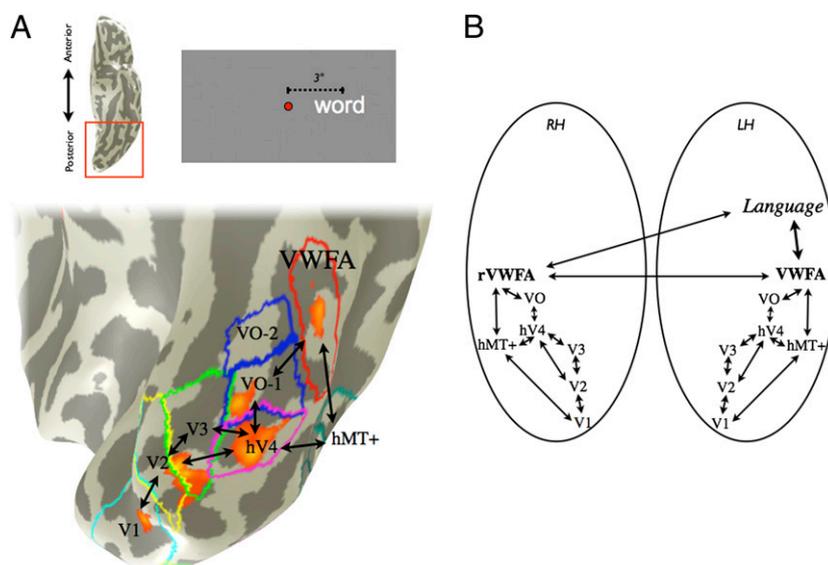
**Retinotopy and Neural Representations of Complex Visual Stimuli.** If VWFA neurons store statistical regularities about word forms, such as letter combination frequencies or ordered letter strings, the demonstration of visual field position sensitivity implies that this statistical information is encoded multiple times, once for each visual field position. That is, if VWFA neurons are trained to respond to specific types of sublexical or lexical information (24, 35), the training must be represented by multiple neural populations that are responsive to different parts of the visual field. Alternatively, the statistical regularities of the word forms may not be stored within the VWFA. Instead, the information may be provided by feedback from language-related processing modules (20) across relatively wide regions of VOT cortex (35).

A similar question of position invariance has been raised with respect to processing of other complex visual forms, such as objects and faces. In the nonhuman primate, it was observed that neuronal receptive field sizes increase along inferotemporal (IT) cortex (52), which corresponds roughly to human VOT cortex. This observation supported the idea of neural machinery that recognizes objects across the visual field, ignoring position (33).

In the macaque, the notion of visual field position-invariant representations spanning large regions of the visual field has been recently challenged (53–55). Similarly, in human the regions of cortex thought to underlie object recognition in human, such as LO, are shown to be retinotopically organized (38, 39). Such position sensitivity is further supported by psychophysical measurements showing the face aftereffect is retinotopically specific (56). Furthermore, training word or object shape perception at one retinal location does not generalize substantially to other retinal locations, as measured perceptually in the human (57, 58) and within single neurons in nonhuman primate IT cortex (59). At present, the balance of evidence suggests that visual recognition of complex stimuli, including letters and words, depends on multiple retinotopically organized neural representations.

## Conclusion

Position sensitivity is present in a key cortical circuit that conveys visual information about words to language areas. The VWFA is likely to be retinotopically organized. The discovery of retinotopic organization and the measurements of precise temporal information of VWFA signals support the hypothesis that the VWFA and the rVWFA work together as part of a unified circuit that interfaces vision and language. Finally, models of VWFA function must take into account that either statistical regularities



**Fig. 7.** A new circuit diagram for visual word form processing. (A) Parafoveal stimuli produce distinct clusters of activity in visual cortex that can be measured in individual subjects. A contrast map between words in the right and left visual field ( $3^\circ$  eccentricity, right > left) is shown on an inflated ventral posterior left hemisphere of one subject ( $P < 0.001$ , uncorrected; orange-yellow). The visual field map and VWFA positions, derived in separate experiments, are also shown. The red square on the ventral view of the left hemisphere (*Inset Upper Left*) indicates the magnified region. (B) In the new circuit diagram multiple visual field maps perform a feature-tolerant and position-sensitive transformation, which yields an abstract shape representation in the VWFA (31). The position-sensitive VWFA and the rVWFA responses may interact and communicate word form information to the language system. Black lines indicate possible processing pathways. LH, left hemisphere; RH, right hemisphere. Subject is S2.

about word forms must be stored repeatedly in distinct cortical populations or this information must be broadly distributed by feedback to these populations.

## Methods

**Subjects (fMRI).** Seven right-handed subjects (three females; ages 23–38 y, median age 27 y) participated in the study, which was approved by the institutional review board at Stanford University. All subjects gave informed consent, were native English speakers, and had normal vision.

**Stimuli: Main fMRI Experiment.** Stimuli were projected onto a screen that the subject viewed through a mirror fixed above the head. The screen, located at the back of the magnetic bore, subtended a  $12^\circ$  radius in the vertical dimension. A custom magnetic resonance-compatible eye tracker mounted to the mirror continuously recorded (software: ViewPoint; Arrington Research) eye movements to ensure good fixation performance during scanning sessions.

All stimuli used for the main fMRI experiment were four-letter nouns with a mean frequency (per million) of 74.1 (median = 27.9, SD = 164.1) (59). The mean constrained trigram frequency (per million) of the words was 301.5 (median = 135.3, SD = 422.2). All words ( $n = 288$ , 24 words  $\times$  12 blocks) were unique within each run.

Words were presented, in the context of a block-design fMRI experiment, at  $1^\circ$ ,  $3^\circ$ , or  $6^\circ$  from a central fixation dot. Within a scanning session, words could appear either to the left or to the right of fixation, or above or below fixation, at those eccentricities. The sizes of the words at different eccentricities were chosen to achieve equal readability, on the basis of piloting psychophysical lexical decision task experiments in three subjects. Words presented at  $1^\circ$ ,  $3^\circ$ , and  $6^\circ$  had x-heights of  $0.3^\circ$ ,  $0.5^\circ$ , and  $0.8^\circ$  of visual angle, respectively. These parameters produced words that were  $1.3^\circ$ ,  $2.3^\circ$ , and  $3.9^\circ$  in width and  $0.5^\circ$ ,  $0.7^\circ$ , and  $1.2^\circ$  in height. The task was a challenging fixation task that required the subject to report, via a button press, a change in the small ( $2 \times 2$  pixels) fixation dot's color.

Words were rendered as white against a gray background in monospaced Courier font with a Weber contrast of 0.8. Stimulus presentation and response collection were performed using custom Matlab (MathWorks) scripts and controlled using the Psychtoolbox (60). Stimuli were presented for 300 ms with a 200-ms interstimulus interval, giving a total of 24 words per 12-s block. Every block was followed by a 12-s interblock fixation interval. The order of conditions was pseudorandomized across runs and across subjects.

**Scanning Parameters.** Anatomical and functional imaging data were acquired on a 3T General Electrical scanner using an eight-channel head coil. Subject head motion was minimized by placing padding around the head. Functional MR data were acquired using a spiral pulse sequence (61).

Each subject participated in three scanning sessions of  $\sim 1$  h each. The sessions were separated by days to months. In the first scanning session, the VWFA functional localizer and retinotopy scans were performed to identify regions of interest. The second and third scanning sessions consisted of the experimental runs, with stimuli either in the upper and lower VFs or in the left and right VFs. The order of the latter two scanning sessions was randomized across subjects. Two subjects participated in only one or the other session, resulting in six subjects per left–right and upper–lower VF session. These scanning sessions consisted typically of four experimental runs, although we collected six runs in two subjects for each session type. In three of the seven subjects, the retinotopy scans were collected on a separate day; the VWFA localizer was performed during the same session as one of the experimental scan sessions.

Of the 12 total experimental sessions (six subjects  $\times$  two stimulus types), data in 8 sessions were collected at  $1.7 \times 1.7 \times 2$ -mm resolution (acquisition matrix size  $128 \times 128$ ; field of view (FOV) 220 mm; 22 axial-oblique slices covering the occipital, the temporal, and part of the frontal lobe), whereas data in 4 sessions were collected at  $2.8 \times 2.8 \times 2.5$ -mm resolution (acquisition matrix size  $64 \times 64$ , FOV 180 mm, 30 coronal-oblique slices covering the occipital and part of the temporal and parietal lobes). There were no systematic differences in classifier performance that depends on this difference in resolution, although anecdotally the best-performing data came from the subject with the least head movement ( $<0.1$  voxels) and high spatial resolution. At both spatial resolutions, data were acquired with the following parameters: repetition time (TR) = 2,000 ms, echo time (TE) = 30 ms, flip angle =  $77^\circ$ .

T1-weighted anatomical images were acquired separately for each subject. For each subject, 3–5 vol (0.9 mm isotropic) were acquired and averaged together for higher gray/white matter contrast. The resulting volumes were resampled to  $1 \times 1 \times 1$ -mm voxels and aligned to anterior commissure–posterior commissure (AC-PC) space. Functional data were aligned to this volume, using a mutual information algorithm through an in-plane anatomical image acquired before each set of functional scans, or directly through the mean functional volume from the first run of each set of runs.

**General Analysis.** fMRI data were analyzed using the freely available mrVista tools (<http://white.stanford.edu/software/>; SVN revision 2649). Motion artifacts within and across runs were corrected using an affine transformation of each temporal volume in a session to the first volume of the first run. All

subjects showed <1 voxel (and generally <0.4 voxel) head movement. Baseline drifts were removed from the time series by fitting a quadratic function to each run. A GLM was fitted to each voxel's time course, estimating the relative contribution of each 12-s block within a run to the time course. Each 12-s block can be thought of as a "trial" of that condition. We included a separate regressor for each run to account for constant (DC) shifts in baseline. The resulting regressor estimates ( $\beta$ -values) for each voxel within a region of interest, for every trial, were entered into the classifier analysis.

**Classification.** Classification was performed using LibSVM (42). A leave-one-run-out cross-validation procedure was used. Specifically, a linear classifier model was built on all but one run and we determined classification accuracy on the remaining (completely independent) runs. This procedure was repeated leaving each run out once, to calculate the mean accuracy for each subject for each region of interest. In some analyses, classification accuracy was based on correctly decoding one of six spatial locations (chance =  $1/6 = 16.7\%$ ). In other analyses, classification accuracy was based on correctly decoding the left vs. right or upper vs. lower hemifield (chance =  $1/2 = 50\%$ ). For these classifications, trials at  $1^\circ$ ,  $3^\circ$ , or  $6^\circ$  eccentricity within a hemifield were treated as the same condition, because they are located within the same visual hemifield. All Matlab codes used for fMRI analyses are part of the freely available mrVista tools.

To obtain linear classifier weight maps (Fig. 3 and Fig. S2), we recorded the model weights assigned to each feature, or voxel, in the ROI for each iteration of the leave-one-run-out procedure. The mean of these weights across iterations within a voxel is projected back onto the cortex to visualize what portions of cortex were most informative for classifying between any two conditions.

A significance threshold for classification accuracy was computed by bootstrapping. Specifically, noise distributions were generated by randomly shuffling the condition labels 100 times within each subject, in each ROI, and computing the classification accuracy on the shuffled data. The 95th percentile of all classification accuracies, across all subjects and all ROIs, was chosen as the significance threshold.

Searchlight analyses (Fig. S1) were conducted in individual subjects by moving a disk-shaped ROI (5-mm radius) along the gray matter surface, similar to that in ref. 44. Each voxel where data were collected served as the center of one such ROI. Classification was performed as explained above, for each ROI, and the center voxel of the ROI was colored according to the decoding accuracy from that ROI.

To calculate the noise variance that equates the V1/V2 and VWFA classifier performance, we (a) calculated the SD of the  $\beta$ -values in the V1/V2 data and (b) multiplied a Gaussian noise matrix the same size as the V1/V2 data by multiples of this SD [multipliers were  $[0 \text{ logspace}(0, 1.5, 50)]$  executed in Matlab] for each subject. This procedure was repeated 100 times per subject with new noise matrices to estimate a mean accuracy for each noise level. We then found the amount of noise, for each subject, that was closest to the VWFA classifier performance for that subject (Fig. S3). Finally, for that particular noise level in each subject, we created a bootstrapped distribution (10,000 repeats) of the fraction of misclassifications that are neighboring visual field locations by classifying between all six upper and lower VF locations (i.e., creating 10,000 confusion matrices per subject using V1/V2 data with noise added). These distributions were averaged and compared with the mean fraction of neighbor misclassifications from the VWFA confusion matrices (one per subject).

**Retinotopy.** We identified known retinotopic areas in each individual, using the previously described population receptive field method (62). Briefly, this method finds the best-fitting 2D Gaussian receptive field, consisting of an  $x$ ,  $y$  position, and sigma, the receptive field size, for each voxel in response to bars of flashing checkerboards sweeping across the VF. Retinotopic areas were defined by polar angle reversals (45). In one subject, wedge and ring stimuli were used instead of using the population receptive field method, because the data had been acquired for a separate experiment. We identified the following visual areas in each individual: V1, V2v/d, V3v/d, hV4, VO-1, and VO-2 (where v/d refer to ventral and dorsal subfields).

**VWFA Localizer: Stimuli.** The VWFA localizer consisted of four block-design runs of 180 s each. Twelve-second blocks of words, fully phase-scrambled words, or checkerboards alternated with 12-s blocks of fixation (gray screen with fixation dot). Stimuli during each block were shown for 400 ms, with 100-ms interstimulus intervals, giving 24 unique stimuli of one category per block. Words were six-letter nouns with a minimum word frequency of seven per million (63). The size of all stimuli was  $14.2^\circ \times 4.3^\circ$  and was presented on a partially phase-scrambled gray background. Fully phase-scrambled words

consisted of the same stimuli, except that the phase of the images was randomized. Checkerboard stimuli reversed contrast at the same rate that the stimuli changed and were the same size as other stimuli. The order of the blocks was pseudorandomized, and the order of stimuli within those blocks was newly randomized for each subject.

**VWFA, rVWFA, and Broca's Area Definition.** The VWFA and Broca's area were defined in individual subjects on the basis of functional localizers. We defined the VWFA as the activation on the ventral cortical surface from a contrast between words and phase-scrambled words ( $P < 0.001$ , uncorrected, Fig. 1). The region was restricted to responsive voxels outside retinotopic areas and anterior to hV4. The MNI coordinates of the peak voxel within the ROI were identified by finding the best-fitting transform between the individual T1-weighted anatomy and the average MNI-152 T1-weighted anatomy and then applying that transform to the center-of-mass coordinate of the VWFA. The identical process was used to define the rVWFA in each subject.

Broca's area was defined from activation in the gray matter on or near the left inferior frontal gyrus in response to the contrast of words vs. fixation ( $P < 0.01$ , uncorrected). The orientation of the slices included this region in five of the six subjects for the upper-lower classification and three of the six for the left-right classification. The typical fMRI response modulation to words in Broca's area for our implicit reading task was about 0.5%.

**ECOG Recordings: Subject.** We recorded ECoG data from a patient undergoing neurosurgical evaluation for intractable epilepsy. The epileptic focus was found in the anterior left temporal lobe. The electrodes that provided the data for the present study showed no abnormal discharges or epileptic activity. The patient was a 20-y-old female. The study did not cause additional risk to the participant, and the intracranial procedures were conducted entirely for clinical reasons to localize the source of epileptic discharges. The study was approved by the Stanford University Institutional Review Board Office for Protection of Human Research Subjects, and subjects signed informed consent for participation in our research study.

Before electrode placement, the patient underwent an fMRI localizer of the VWFA. All procedures for this localizer were the same as for the other subjects in this study, as described above.

**ECOG: Electrode Localization.** The MS08R-IP10X-000 and MS10R-IP10X-000 strips and FG16A-SP10X-000 grid made by AdTech Medical Instrument were used for clinical and research-related recording in our subject. These electrodes have the following parameters: 4-mm flat diameter contacts with 2.3-mm diameter of exposed recording area ( $4.15 \text{ mm}^2$ ) and interelectrode distance of 1 cm. Postsurgical computed tomography (CT) images indicating the location of electrodes were aligned to preoperative T1-weighted structural MRI images using a mutual-information algorithm, implemented in SPM5. The electrodes were easily identified in the CT scans (Fig. S4B) and their locations were manually marked. These images were visualized using ITKGray, a segmentation tool based on ITKSnap (64). The resulting images were manually aligned to 3D mesh renderings of the T1 anatomical images produced using mrVista, on which the fMRI activation is displayed, thereby conserving the electrode to T1 anatomical image alignment. This procedure allowed us to construct 3D visualization of electrode locations relative to the patient's cortical anatomy (Fig. 6A).

**ECOG: Task.** After implantation of the electrodes and postsurgical stabilization, two versions of the task were administered to the patient while she was in her hospital bed. Stimuli were presented on a 15-inch MacBook Pro and responses collected with a Toshiba external 10-key numerical keyboard. In the first task, the patient was asked to fixate on the central fixation dot and report (via button press) a change in the color of the fixation dot. Stimuli were presented for 200 ms, followed by a 0.9-s response time (gray mean luminance with fixation dot), followed by an intertrial interval randomly distributed between 25 and 3,000 ms, approximately following a Poisson distribution with a mean of 100 ms. In the second version of the task, which we report here, the stimuli were identical except that the response time was 2.7 s. In this task, the subject was asked to try to read the stimuli on the screen and report whether the stimulus was a real English word or not. The subject was also asked to keep fixation on all stimuli, even at the cost of incorrect responses.

**ECOG: Stimuli.** Stimuli consisted of words, pseudowords, consonant strings, and noise stimuli in an event-related design. For the purposes of this study, we analyzed words, pseudowords, and consonant strings shown at the center and to the left or right of fixation. Stimuli at the fovea were four, six, or eight letters in length, and stimuli to the left or right of fixation were four or eight

letters in length. Eight-letter stimuli centered to the left or right of fixation extended into the opposite VF by 0.5°. Four-letter stimuli centered to the left or right of fixation were fully contained in one hemifield, with the edge of the word being 1° away from the fixation dot. Each letter was ~0.5° in width.

**ECoG: Recording and Analysis.** We recorded signals at 3,051.8 Hz through a 128-channel recording system made by Tucker Davies Technologies. Offline, we applied a notch filter at 60 Hz and harmonics to remove power line noise. We removed channels with epileptic activity, as determined by the patient's neurologist. Before processing, the signal was down-sampled to 435 Hz. To visualize electrophysiological responses, we created event-related spectral perturbation (ERSP) maps on the basis of the normalized power of electrophysiological activity during each condition. A Hilbert transform was applied to each of 42 band-pass filtered time series to obtain instantaneous power. Using the Hilbert-transformed time series, time-frequency analysis was performed for event-related data. We logged the onset and duration of each trial via photodiode event markers for each experimental condition time locked with the ECoG recording. Event markers were used to precisely align the recorded signal to the stimulus presentation sequence. The ERSP was scaled by the total mean power at each frequency to compensate for

the skewed distribution of power values over frequencies and the result converted to decibel units.

To measure response timing, the instantaneous power of the signal was extracted from a frequency band comprising 80–150 Hz, which is the frequency band that showed the strongest response in the ERSP (Fig. S4A). Response timing was measured in two ways: (i) For each individual trial, we found the time point with the peak power between 30 and 600 ms after stimulus onset. The time point at which the power reached half-maximum was recorded as the response time. (ii) For each individual trial, we estimated the noise of the response by computing the SD of the signal across the 300 ms before trial onset. The time point at which the power reached 10 SDs of this baseline period for that trial was recorded as the response time.

**ACKNOWLEDGMENTS.** We thank Jonathan Winawer and Lee M. Perry for their help and thoughtful input. We also thank Mohammad Dastjerdi and Brett Foster for their help collecting ECoG data and for advice on data analysis. Jason D. Yeatman and Joyce E. Farrell provided valuable advice on the manuscript. This work was supported by the Medical Scientist Training Program (A.M.R.), the Bio-X Graduate Student Fellowship Program (A.M.R.), and National Institutes of Health Grant R01 EY015000.

- Cohen L, et al. (2000) The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* 123:291–307.
- Wandell BA, Rauschecker AM, Yeatman J (2012) Learning to see words. *Annu Rev Psychol* 63:31–53.
- Warrington EK, Shallice T (1980) Word-form dyslexia. *Brain* 103:99–112.
- Yeatman J, Rauschecker AM, Wandell B (2012) Anatomy of the visual word form area: Adjacent cortical circuits and long-range white matter connections. *Brain Lang*, 10.1016/j.bandl.2012.04.010.
- 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.
- Dehaene S, et al. (2010) How learning to read changes the cortical networks for vision and language. *Science* 330:1359–1364.
- Maisog JM, Einbinder ER, Flowers DL, Turkeltaub PE, Eden GF (2008) A meta-analysis of functional neuroimaging studies of dyslexia. *Ann N Y Acad Sci* 1145:237–259.
- Gaillard R, et al. (2006) Direct intracranial, fMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron* 50:191–204.
- Cohen L, et al. (2002) Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain* 125:1054–1069.
- Cohen L, et al. (2003) Visual word recognition in the left and right hemispheres: Anatomical and functional correlates of peripheral alexias. *Cereb Cortex* 13:1313–1333.
- Dehaene S, Cohen L, Sigman M, Vinckier F (2005) The neural code for written words: A proposal. *Trends Cogn Sci* 9:335–341.
- Grainger J, Rey A, Dufau S (2008) Letter perception: From pixels to pandemonium. *Trends Cogn Sci* 12:381–387.
- 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.
- Molko N, et al. (2002) Visualizing the neural bases of a disconnection syndrome with diffusion tensor imaging. *J Cogn Neurosci* 14:629–636.
- Pammer K, et al. (2004) Visual word recognition: The first half second. *Neuroimage* 22:1819–1825.
- Szwed M, et al. (2011) Specialization for written words over objects in the visual cortex. *Neuroimage* 56:330–344.
- Cohen L, Dehaene S, Vinckier F, Jobert A, Montavont A (2008) Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. *Neuroimage* 40:353–366.
- Kronbichler M, et al. (2004) The visual word form area and the frequency with which words are encountered: Evidence from a parametric fMRI study. *Neuroimage* 21: 946–953.
- Devlin JT, Jamison HL, Gonnerman LM, Matthews PM (2006) The role of the posterior fusiform gyrus in reading. *J Cogn Neurosci* 18:911–922.
- Price CJ, Devlin JT (2011) The interactive account of ventral occipitotemporal contributions to reading. *Trends Cogn Sci* 15:246–253.
- Dehaene S, Cohen L (2011) The unique role of the visual word form area in reading. *Trends Cogn Sci* 15:254–262.
- Reich L, Szwed M, Cohen L, Amedi A (2011) A ventral visual stream reading center independent of visual experience. *Curr Biol* 21:363–368.
- Vidyaasagar TR, Pammer K (2010) Dyslexia: A deficit in visuo-spatial attention, not in phonological processing. *Trends Cogn Sci* 14:57–63.
- 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.
- Cai Q, Lavidor M, Brysbaert M, Paulignan Y, Nazir TA (2008) Cerebral lateralization of frontal lobe language processes and lateralization of the posterior visual word processing system. *J Cogn Neurosci* 20:672–681.
- Dehaene S, et al. (2004) Letter binding and invariant recognition of masked words: Behavioral and neuroimaging evidence. *Psychol Sci* 15:307–313.
- Cohen L, Dehaene S (2004) Specialization within the ventral stream: The case for the visual word form area. *Neuroimage* 22:466–476.
- Dehaene S, Le Clec'H G, Poline JB, Le Bihan D, Cohen L (2002) The visual word form area: A prelexical representation of visual words in the fusiform gyrus. *Neuroreport* 13:321–325.
- Dehaene S, et al. (2001) Cerebral mechanisms of word masking and unconscious repetition priming. *Nat Neurosci* 4:752–758.
- Nazir TA, Ben-Boutayab N, Decoppet N, Deutsch A, Frost R (2004) Reading habits, perceptual learning, and recognition of printed words. *Brain Lang* 88:294–311.
- Cox DD, Meier P, Oertelt N, DiCarlo JJ (2005) 'Breaking' position-invariant object recognition. *Nat Neurosci* 8:1145–1147.
- Rauschecker AM, et al. (2011) Visual feature-tolerance in the reading network. *Neuron* 71:941–953.
- Riesenhuber M, Poggio T (2000) Models of object recognition. *Nat Neurosci* 3(Suppl): 1199–1204.
- Binder JR, Medler DA, Westbury CF, Liebenthal E, Buchanan L (2006) Tuning of the human left fusiform gyrus to sublexical orthographic structure. *Neuroimage* 33: 739–748.
- Vinckier F, et al. (2007) Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron* 55: 143–156.
- 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.
- Arcaro MJ, McMains SA, Singer BD, Kastner S (2009) Retinotopic organization of human ventral visual cortex. *J Neurosci* 29:10638–10652.
- Larsson J, Heeger DJ (2006) Two retinotopic visual areas in human lateral occipital cortex. *J Neurosci* 26:13128–13142.
- Amano K, Wandell BA, Dumoulin SO (2009) Visual field maps, population receptive field sizes, and visual field coverage in the human MT+ complex. *J Neurophysiol* 102: 2704–2718.
- Grill-Spector K, Kourtzi Z, Kanwisher N (2001) The lateral occipital complex and its role in object recognition. *Vision Res* 41:1409–1422.
- Malach R, et al. (1995) Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci USA* 92:8135–8139.
- Chang C, Lin C (2011) LIBSVM: A library for support vector machines. *ACM Trans Intell Syst Technol* 2(3):2721–2727.
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci USA* 103:3863–3868.
- Chen Y, et al. (2011) Cortical surface-based searchlight decoding. *Neuroimage* 56: 582–592.
- Wandell BA, Dumoulin SO, Brewer AA (2007) Visual field maps in human cortex. *Neuron* 56:366–383.
- Dougherty RF, et al. (2003) Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *J Vis* 3:586–598.
- Tarkiainen A, Helenius P, Hansen PC, Cornelissen PL, Salmelin R (1999) Dynamics of letter string perception in the human occipitotemporal cortex. *Brain* 122:2119–2132.
- Nobre AC, Allison T, McCarthy G (1994) Word recognition in the human inferior temporal lobe. *Nature* 372:260–263.
- Winawer J, Horiguchi H, Sayres RA, Amano K, Wandell BA (2010) Mapping hV4 and ventral occipital cortex: The venous eclipse. *J Vis* 10:1.
- Cuingnet R, et al. (2010) Spatially regularized SVM for the detection of brain areas associated with stroke outcome. *Medical Image Computing and Computer-Assisted Intervention: MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, eds Jiang T, et al. (Springer-Verlag, Berlin and Heidelberg), Vol 13(Pt 1), pp 316–323.
- Di Virgilio G, Clarke S (1997) Direct interhemispheric visual input to human speech areas. *Hum Brain Mapp* 5:347–354.
- Gross CG, Rocha-Miranda CE, Bender DB (1972) Visual properties of neurons in inferotemporal cortex of the Macaque. *J Neurophysiol* 35:96–111.
- DiCarlo JJ, Maunsell JH (2003) Anterior inferotemporal neurons of monkeys engaged in object recognition can be highly sensitive to object retinal position. *J Neurophysiol* 89:3264–3278.

54. Hung CP, Kreiman G, Poggio T, DiCarlo JJ (2005) Fast readout of object identity from macaque inferior temporal cortex. *Science* 310:863–866.
55. Kravitz DJ, Vinson LD, Baker CI (2008) How position dependent is visual object recognition? *Trends Cogn Sci* 12:114–122.
56. Afraz SR, Cavanagh P (2008) Retinotopy of the face aftereffect. *Vision Res* 48:42–54.
57. Yu D, Legge GE, Park H, Gage E, Chung STL (2010) Development of a training protocol to improve reading performance in peripheral vision. *Vision Res* 50:36–45.
58. Nazir TA, O'Regan JK (1990) Some results on translation invariance in the human visual system. *Spat Vis* 5:81–100.
59. Cox DD, DiCarlo JJ (2008) Does learned shape selectivity in inferior temporal cortex automatically generalize across retinal position? *J Neurosci* 28:10045–10055.
60. Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436.
61. Glover GH (1999) Simple analytic spiral K-space algorithm. *Magn Reson Med* 42:412–415.
62. Dumoulin SO, Wandell BA (2008) Population receptive field estimates in human visual cortex. *Neuroimage* 39:647–660.
63. Medler DA, Binder JR (2005) MCWord: An on-line orthographic database of the English language. Available at <http://www.neuro.mcw.edu/mcword/>. Accessed March 30, 2012.
64. Yushkevich PA, et al. (2006) User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage* 31:1116–1128.