

Visual Feature-Tolerance in the Reading Network

Andreas M. Rauschecker,^{1,2,*} Reno F. Bowen,² Lee M. Perry,² Alison M. Kevan,² Robert F. Dougherty,² and Brian A. Wandell²

¹Neurosciences Program and Medical Scientist Training Program, Stanford School of Medicine

²Psychology Department

Stanford University, Stanford, CA 94305, USA

*Correspondence: andreasr@stanford.edu

DOI 10.1016/j.neuron.2011.06.036

SUMMARY

A century of neurology and neuroscience shows that seeing words depends on ventral occipital-temporal (VOT) circuitry. Typically, reading is learned using high-contrast line-contour words. We explored whether a specific VOT region, the visual word form area (VWFA), learns to see only these words or recognizes words independent of the specific shape-defining visual features. Word forms were created using atypical features (motion-dots, luminance-dots) whose statistical properties control word-visibility. We measured fMRI responses as word form visibility varied, and we used TMS to interfere with neural processing in specific cortical circuits, while subjects performed a lexical decision task. For all features, VWFA responses increased with word-visibility and correlated with performance. TMS applied to motion-specialized area hMT+ disrupted reading performance for motion-dots, but not line-contours or luminance-dots. A quantitative model describes feature-convergence in the VWFA and relates VWFA responses to behavioral performance. These findings suggest how visual feature-tolerance in the reading network arises through signal convergence from feature-specialized cortical areas.

INTRODUCTION

During successful reading, the visual system efficiently transforms a complex input of contrast-defined strokes of ink into phonological and semantic word representations. After entering primary visual cortex (V1), visual information about words undergoes several transformations in extrastriate cortex, including regions localized to ventral occipitotemporal (VOT) cortex (Dehaene et al., 2005; DiCarlo and Cox, 2007). This process includes a neural representation of letter strings in a specific region of VOT cortex that has been labeled the visual word form area (VWFA; Cohen et al., 2000). The VWFA is the primary candidate neural site for the long-hypothesized visual word lexicon (Dejerine, 1892; Warrington and Shallice, 1980; Wernicke, 1874), although debates about its specific role continue (Dehaene and Cohen, 2011; Price and Devlin, 2011; Wandell et al., 2010).

Ultimately, the VWFA is thought to communicate directly with language-related regions (Devlin et al., 2006). These language cortices presumably require a common input format that is insensitive to particular visual features. The VWFA may act as an essential link between visual and language cortices by providing such a common input format (Jobard et al., 2003). Alternatively, the collection of visual areas may have separate access to the same network with the potential to bypass the VWFA (Price and Devlin, 2011; Richardson et al., 2011).

We took a fresh look at this question by measuring responses to word stimuli intended to target different feature-specialized visual cortical regions (Figure 1). Specifically, we designed word stimuli whose shape is defined using atypical features: dots rather than line contours. The dots carried word information by spatially varying dot luminance, dot motion direction, or both. Current hypotheses suggest that the VWFA, through reading experience, becomes specialized for detecting particular line contour configurations (Dehaene and Cohen, 2011; Szwed et al., 2009; Szwed et al., 2011). Thus, the VWFA may not be expected to respond to dot-defined word stimuli that contain no line contours. Motion-defined words, for example, are expected to be processed by a motion-specialized cortical region (hMT+) located in the canonical dorsal visual pathway (Ungerleider and Mishkin, 1982) and may not depend on the VWFA in the ventral visual pathway.

Previous literature suggests an important role for the human motion complex (hMT+) in reading. Following the description of behavioral and anatomical motion processing deficits in dyslexia (Galaburda and Livingstone, 1993; Livingstone et al., 1991; Martin and Lovegrove, 1987), hMT+ was found to be underactivated in dyslexics in response to motion stimuli when measured using functional magnetic resonance imaging (fMRI) (Eden et al., 1996). Further studies revealed that the extent of hMT+ response to visual motion correlates with reading ability more generally (Ben-Shachar et al., 2007a; Demb et al., 1997, 1998). Based on these results, one might speculate that hMT+ serves a crucial role in reading. However, the nature of that role and its relationship to the VWFA have not been elucidated. By measuring (using fMRI) and disrupting (using transcranial magnetic stimulation, TMS) neural activity in hMT+, we tested its causal role in seeing words.

The results suggest that the VWFA serves as an essential link between vision and language by representing visual letter strings in a common format, independent of the particular defining visual features. To achieve this feature-tolerant shape representation, the VWFA has flexible input connectivity from feature-specialized visual areas, including hMT+.

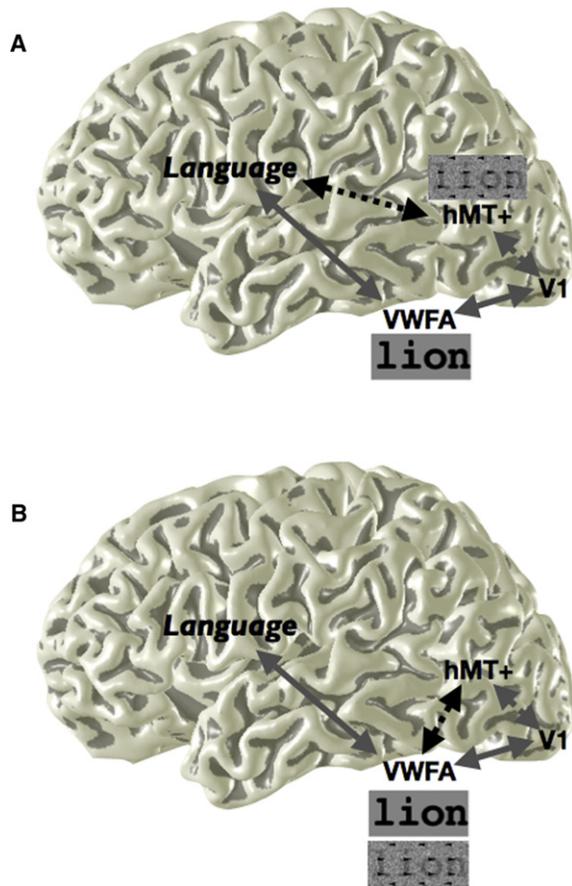


Figure 1. Alternative Hypotheses of How Information Is Communicated from V1 to Language Circuits

Different visual features are processed by functionally specialized regions in visual cortex. For example, words defined purely by motion cues may be processed by area hMT+. In hypothesis A, different cortical areas have separate access to the language network. In hypothesis B, all word stimuli, regardless of feature type, are converted to a common representation en route to the VWFA in VOT, which has unique access to the language network. Dotted connections represent communication between regions specifically for motion-defined stimuli, and solid connections represent communication for words defined by line contours. The response to different stimulus types in VWFA and hMT+, based on the difference in the black dotted connection, differentiates the two hypotheses. Schematic line contour and motion-dot stimuli are shown.

RESULTS

The VWFA Responds to Words Defined by Different Features

In an event-related fMRI design, we measured VWFA blood oxygen-level-dependent (BOLD) responses to increasing levels of word visibility while subjects were engaged in a lexical decision task. The visibility of words defined by line contours (i.e., standard words) was controlled by phase-scrambling (see [Experimental Procedures](#)). These event-related measures confirm that the VWFA response increases with word visibility (“word visibility response function”; [Figure 2](#)). Similar response functions have been observed in block-design fMRI during an

incidental reading task ([Ben-Shachar et al., 2007b](#)), and also using magnetoencephalography ([Tarkiainen et al., 1999](#)).

Word stimuli created by replacing the line-contour features with dots of spatially varying luminance or motion-direction (“luminance-dot” and “motion-dot” stimuli; see [Experimental Procedures](#) for details) produce similar word visibility response functions in the VWFA. In all three cases the peak response modulation is quite high—reaching about 1% for the highest visibilities ([Figure 2](#)). Thus, the VWFA is responsive to word visibility even when words are defined by unconventional and unpracticed stimulus features. The onset and time to peak of the BOLD signal response time courses are similar for the different stimulus features ([Figure 2](#), right column).

We used a mixed effects linear model, with subject considered as a random effect, to statistically compare the motion-dot stimulus responses to the other stimulus types (line contour and luminance-dot). Contrasts were defined to compare the motion-dot stimulus responses to the other group. There is a significant linear effect ($t = 7.67$, $p < 0.001$) across all stimuli such that BOLD response increases with visibility. There is also a significant overall quadratic effect ($t = 3.12$, $p < 0.001$), indicating that the BOLD response is increasing at a decreasing rate. A significant main effect of feature type ($t = 4.8$, $p < 0.001$) indicates that the line contour and luminance-dot stimuli had a higher average response across visibilities than the motion-dot stimuli. There are no significant linear or quadratic interactions, indicating that the effects do not differ between the motion-dot stimuli and the other feature-type stimuli.

The VWFA’s tolerance to basic stimulus features does not imply that it responds exclusively to words ([Ben-Shachar et al., 2007b](#); [Brem et al., 2006](#)). For example, the fully phase-scrambled line contour stimuli (lowest visibility) are not recognizable as word forms and yet the VWFA BOLD response is more than 0.5%. To see and control for any effects of motion coherence in producing responses, we separately measured VWFA responses in four subjects to moving dots in the shape of a rectangle, as a coherent field of moving dots, and as an incoherent field of moving dots. There is a trend toward a significant VWFA response modulation to rectangles defined by coherent motion (0.33% BOLD modulation, $t[3] = 2.88$, $p = 0.06$), as well as a significant response to a field of incoherently moving dots (0.36% BOLD modulation, $t[3] = 3.18$, $p = 0.05$), compared to fixation. The mean VWFA response (0.19%) to a field of coherently moving dots was non-significant ($t[3] = 1.73$, $p = 0.18$). All of these responses are much smaller than the response to words defined by motion-dots (0.98% BOLD modulation, $t[3] = 6.59$, $p < 0.01$; [Figure S1A](#), available online). In sum, the VWFA response is larger to words than other stimuli ([Ben-Shachar et al., 2007b](#)). A novel finding in this study is that this word response advantage is present for words defined by atypical and unpracticed stimulus features.

In the VWFA, BOLD response modulation is positively correlated with subjects’ lexical decision performance on all stimulus feature types ([Figure 3A](#)). When subjects achieve a high performance level ($> 75\%$ correct), normalized VWFA modulation is high (median normalized BOLD signal 0.82; range 0.42 – 1.0). VWFA modulation for low performance ($\leq 60\%$ correct) is lower on average and highly variable (median normalized

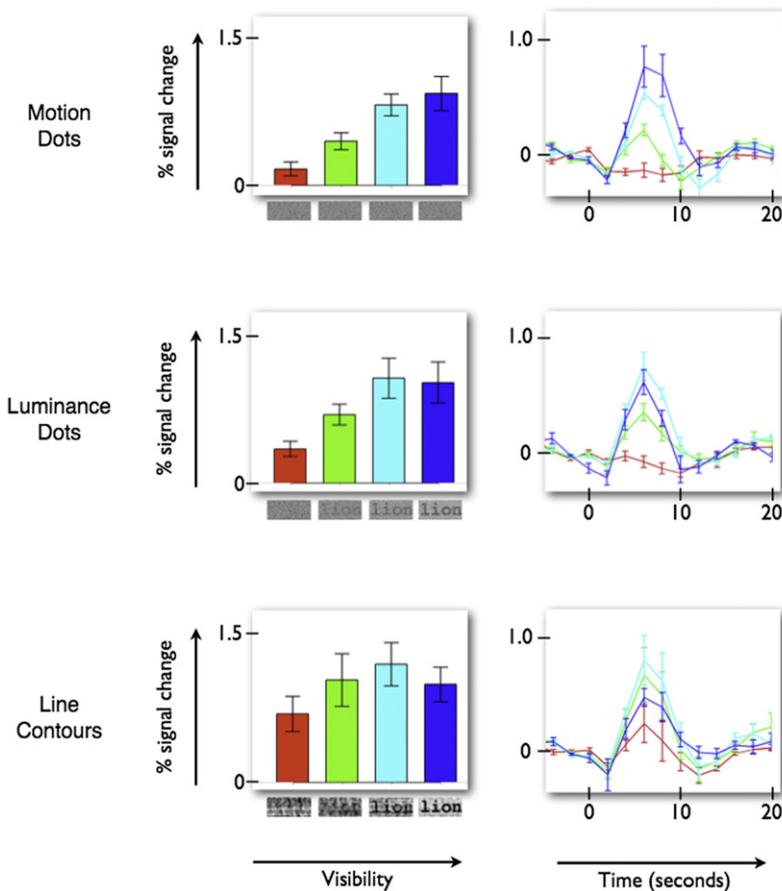


Figure 2. VWFA BOLD Amplitude Increases with Visibility of Words Defined by Different Visual Features

(Left column) Percent signal change for the stimulus events, as measured by the weight of the linear regressor (beta-weight), increases with word visibility. The three panels on the left show the VWFA response increase for words defined by motion-dots, luminance-dots, and line contours.

(Right column) The response time course, averaged across all subjects, peaks at the same time and reaches a similar peak level for the three feature types. The colors of the bars and time course lines indicate corresponding conditions. The baseline (0% level) is defined by the average of the three values prior to stimulus onset.

Error bars are ± 1 SEM across subjects. See also Figure S1A for a related experiment and Figure S2 and Movies S1 and S2 for example stimuli.

response statistical analysis, we used a mixed effects linear model, with subject as a random effect, to compare the response of motion-dot words to the other stimuli. In hMT+, there is an overall significant linear effect ($t = 5.68$, $p < 0.001$), but there is no significant quadratic effect. There is also a significant effect of feature type ($t = 2.74$, $p < 0.01$), indicating that the mean response to motion-dot words is higher than to the other stimuli. Most importantly, in contrast to the findings in the VWFA, there is a significant linear interaction between the response to motion-dot words and the other stimulus types ($t = 3.08$, $p < 0.001$), indicative that the increasing response to higher visibility is only present for the motion-dot words.

Visibility of the motion-dot stimuli is tied to motion coherence. Depending on stimulus parameters, hMT+ responses may increase simply due to motion coherence (Braddick et al., 2001). To test whether the increase in hMT+ responses with word visibility is caused by the increase in coherence alone, we separately measured responses to coherent and incoherent moving dots that did not define a word shape. The dots' motion direction was coherent or incoherent, and their other motion parameters were matched to those of the motion-dot words. There is no significant hMT+ response difference between responses to coherent and incoherent motion-dots (Figure S1B; paired t test, $t[3] = 0.59$, $p = 0.60$). However, as in the event-related paradigm in the main experiment, there is a significant hMT+ response difference between motion-dot words and incoherent motion (paired t test, $t[3] = 5.47$, $p < 0.05$).

Performance on the lexical decision task is strongly correlated (Pearson $r = 0.77$, $p < 10^{-4}$) with hMT+ BOLD response modulation for motion-dot stimuli (Figure 4B), again suggesting the importance of hMT+ activity in correctly parsing feature patterns when stimuli are defined by motion. There is also a correlation ($r = 0.54$, $p < 0.01$), although weaker, between lexical decision performance and hMT+ BOLD responses to luminance-dot stimuli. There is no significant correlation ($p = 0.35$) between hMT+ responses and performance on words defined by line contours.

BOLD signal 0.43; range -0.13 to 0.97). Hence, a high VWFA response does not guarantee good performance, perhaps because processing errors can occur anywhere along the pathway from early visual cortex to downstream language areas. A low VWFA response, meanwhile, is predictive of poor performance, presumably because low activation implies that the VWFA response is failing. Thus, VWFA response is necessary but not sufficient for high reading performance of words composed of any feature type.

This same argument might be applied to responses in primary visual cortex (V1); yet, we found no significant correlations between the overall BOLD signal in V1 and subject performance on the lexical decision task for any stimulus types (Figure 3B). The reason for this appears to be that there is little variation in the V1 response. We presume that if the V1 response failed, subjects would fail to see the words.

hMT+ Responses Increase to Word Form Visibility with Motion-Dots Only

In hMT+, words defined by motion-dot features are the only stimuli to produce responses that increase reliably with word visibility (Figure 4A; one-way ANOVAs for motion: $F[3,13] = 3.43$, $p < 0.05$; luminance $F[3,13] = 1.45$, $p = 0.26$; line contours $F[3,13] = 0.62$, $p = 0.61$). The luminance-dot and line-contour stimuli produce an hMT+ response, but the responses are relatively constant as word visibility increases. Similar to the VWFA

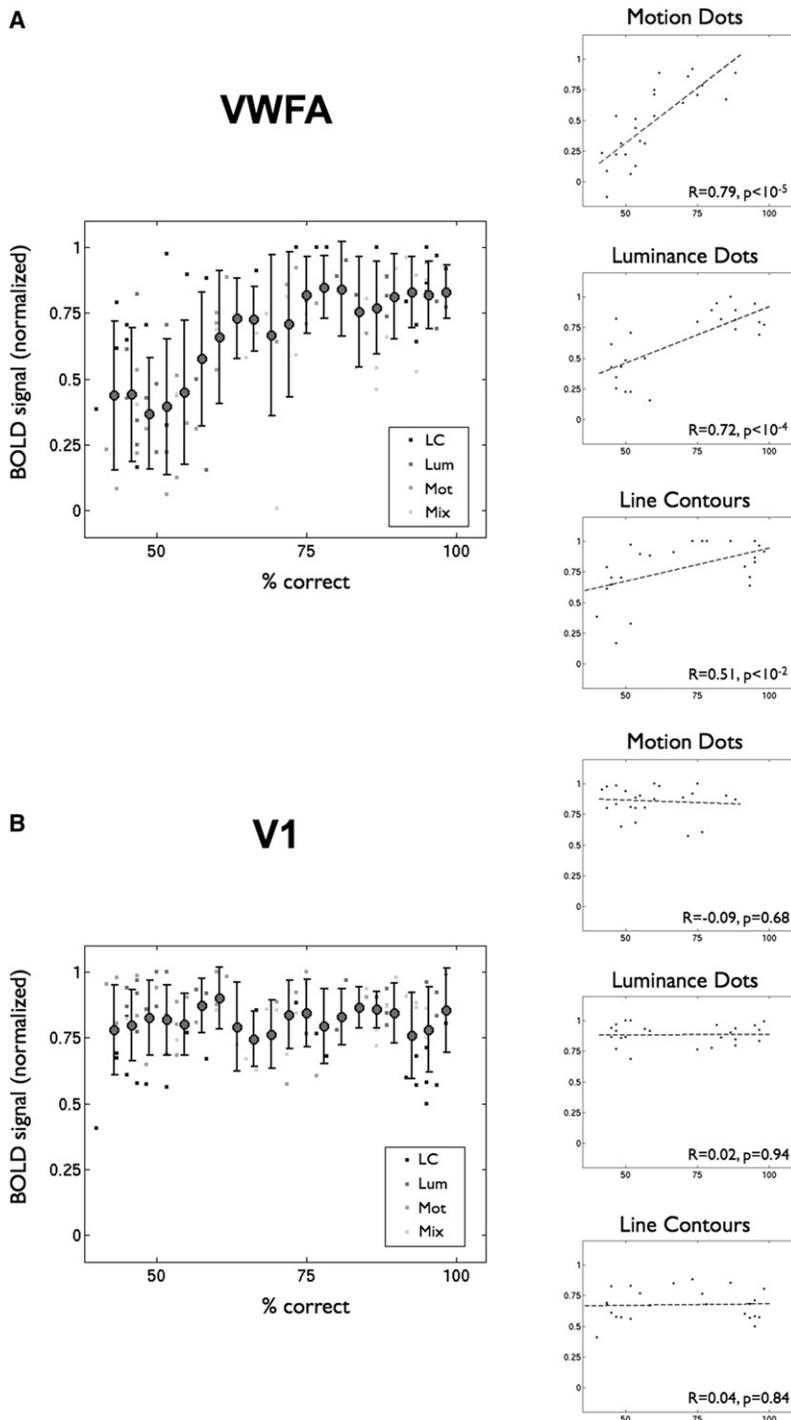


Figure 3. BOLD Response Increases with Lexical Decision Performance in VWFA but Not V1

(A) The left panel shows percent correct in the lexical decision task and normalized BOLD signal amplitude for every subject, visibility level, and feature type (LC = line contour, Lum = luminance-dot, Mot = motion-dot, Mix = motion and luminance dots combined). The filled circles are the mean (± 1 SD) averaged across lexical performance bins (width = 6%). The BOLD signal is normalized by the maximum BOLD signal within that ROI for each subject across feature types and visibility. The right panels show the same points separated by feature type. The dashed lines are linear regression fits, and the insets show the regression coefficient (R) and significance levels (p). (B) The same analysis as in panel (A), but for a region of interest in left V1.

matched across stimulus types at 82% correct performance for each feature type (top dashed line in each plot in Figure 5). Applying TMS to left hMT+ disrupts baseline performance only for stimuli defined by motion features (Figure 5), but not for stimuli defined by other visual features.

We used a linear mixed effects model, with subject intercept considered a random factor, to estimate the effect of TMS at different stimulus-pulse onset asynchronies (SOAs) on performance. A significant decrease in performance occurs only at an SOA of 87–132 ms ($t[42] = -5.14, p < 0.001$). These latency values are consistent with timing between stimulus onset and neural responses in area MT of the human (Prieto et al., 2007) and nonhuman primate (Raiguel et al., 1999). There is no significant effect at any SOA on performance with the luminance-dot stimuli or the line contour stimuli. Individual one-way ANOVAs also confirmed that there is a main effect of SOA for motion-dot stimuli ($F[5,7] = 5.19, p = 0.0009$) but not for line contour stimuli ($F[5,7] = 0.55, p = 0.735$) or luminance-dot stimuli ($F[5,7] = 1.06, p = 0.395$). Thus, hMT+ is necessary only for reading motion-dot stimuli and not all words.

Responses to Word Forms in Retinotopic Maps

To identify which visual areas are sensitive to motion-defined word forms, we measured the word visibility response function in multiple

hMT+ Is Necessary for Seeing Motion-Dot Words but Not Line-Contour Words

We used transcranial magnetic stimulation (TMS) to test the necessity of area hMT+ for processing word stimuli. Specifically, we identified the location of hMT+ in each individual and then used TMS to disrupt neural activity in that region while the subject performed the lexical decision task (see Experimental Procedures for details). Subjects' baseline performance was

left-hemisphere visual area regions of interest (Figure 6). In addition to the VWFA and left hMT+, left hV4 responses increase with word visibility (one-way ANOVA, $F[3,20] = 3.08, p = 0.05$). However, the slope of the hV4 response function is lower than the slope in the VWFA. There is no response dependence on word visibility in left V1 and V2v to motion-dot words. The V3v and VO-1 responses increase monotonically with word visibility, but these increases are not statistically significant. The right

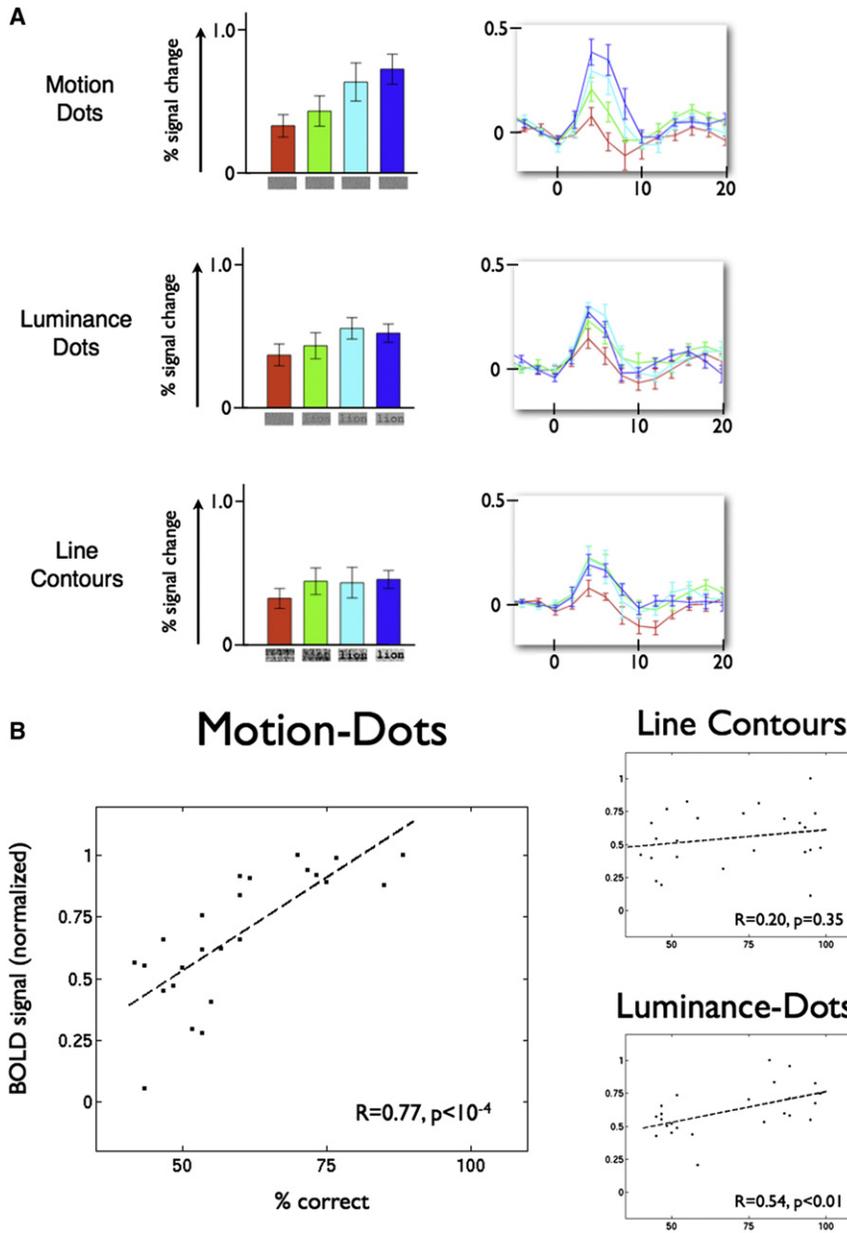


Figure 4. Human MT+ BOLD Responses Increase with Visibility and Lexical Decision Performance for Motion-Dot Words

(A) Left hMT+ BOLD responses increase with visibility for motion-dot words. The plots in this figure follow the same conventions as the VWFA analysis in Figure 2.

(B) The three plots show lexical decision performance (% correct) and normalized BOLD signal amplitude in left hMT+ separated by feature type. BOLD responses increase with lexical decision performance for motion- and luminance-dots, but not line contours. Other details as in Figure 3.

were designed to direct visual responses into distinct pathways, and both the TMS results and BOLD responses in hMT+ suggest that this manipulation succeeded. We therefore performed behavioral and functional imaging experiments to measure how these features, which diverge on a gross anatomical scale after early visual cortex, combine perceptually and in the VWFA response.

The motion and luminance coherence in our stimuli could be modulated independently, providing us with stimuli of different relative amounts of information from each feature (motion-dot and luminance-dot coherence). We measured lexical decision behavioral thresholds for words defined by these feature mixtures (Figure 7A). If motion- and luminance-dot coherence combine additively, then the coherence thresholds to the mixtures will fall on the negative diagonal dotted line. If the features provide independent information to the observer, as in a high-threshold model, thresholds will fall on the outer box. On a probability summation model with an exponent of $n = 3$ the thresholds would fall along the dashed

hemisphere homolog of the VWFA (which we name rVWFA here) was defined as a word-selective region of interest in the right hemisphere, identified by the VWFA localizer in the same manner as the VWFA (see Experimental Procedures). This rVWFA responds increasingly to word visibility ($F[3,16] = 3.67, p < 0.05$), much like the left hemisphere VWFA, apart from a larger response to the noise stimulus (lowest visibility, red bar). The results for early visual areas (V1-hV4) are unchanged when including right hemisphere homologs (not shown).

A Quantitative Model of Feature Combinations

Subjects perceive words defined by either type of dot feature (motion or luminance), and both types of dot features evoke a VWFA response. Motion-dot and luminance-dot features

quarter circle (Graham, 1989; Graham et al., 1978; Quick, 1974). All of these behaviors can be captured by a single equation with one free parameter, n , where $p(c)$ is the probability correct, and m is the motion (m_1) or luminance (m_2) coherence in the stimulus:

$$p(c) = \left(\sum_{i=1}^{i=2} m_i^n \right)^{\frac{1}{n}} \tag{1}$$

The psychophysical data from the five subjects cluster around a probability summation model with an exponent of around $n = 1.7$.

BOLD responses to the same feature-mixture stimuli were measured in several cortical regions of interest. The points in Figure 7B show nine VWFA BOLD responses (± 1 SEM across

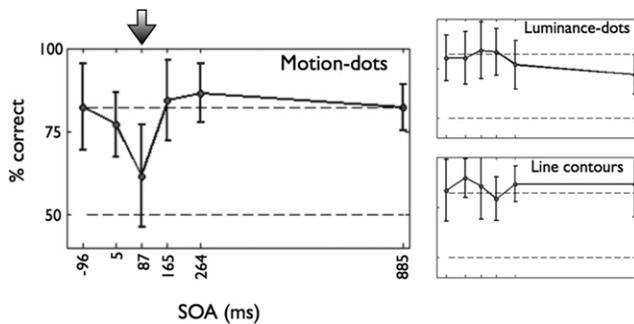


Figure 5. TMS to Left hMT+ Disrupts Lexical Decision Performance Only for Motion-Dot Words

The average performance (% correct) is shown as a function of stimulus-pulse onset asynchrony (SOA). Subjects were consistently and significantly impaired at the lexical decision task for motion-dot words at an SOA of 87 ms (indicated by the arrow; 2nd pulse at 132 ms). There was no significant difference in performance for luminance-dot and line-contour words at any SOA (right panels). Chance performance is 50% (bottom dashed line), and the expected (no TMS effect) performance is 82% based on psychophysical visibility thresholds set prior to each subject's TMS session (top dashed line).

six subjects) at different luminance-dot coherence levels, as a function of motion-dot coherence. Generally, at the lowest luminance-dot coherence (black points), adding motion-dot coherence increases the response. Meanwhile, when the luminance-dot coherence is high (light gray points), adding motion-dot coherence has either no effect or perhaps a slight negative effect.

We fit curves through these BOLD data using a probability summation model that parallels the model used to fit the behavioral thresholds (Figure 7B). This model predicts the BOLD response (B) as arising from two separate neural circuits, one driven by luminance-dot coherence (l) and a second by motion-dot coherence (m). We assume that these signals converge at the VWFA where they are combined with a conventional probability summation rule, with an exponent of $n = 1.7$. This value of n is selected to match the model fit to the behavioral data. The equation for this probability summation model is given by:

$$B = (L^n + M^n)^{\frac{1}{n}} + k, \text{ where } L = \frac{l}{l + \sigma_1} \text{ and } M = \frac{m}{m + \sigma_2} \quad (2)$$

The values l and m are the luminance and motion dot coherence, and k is a constant.

There is good qualitative agreement between the predicted and measured BOLD responses. The predicted and observed responses increase at $l = 0$ with increasing motion-dot coherence, and the predicted and observed responses increase at $m = 0$ with increasing luminance-dot coherence. The responses at relatively high luminance or motion-dot coherence converge. The differential VWFA sensitivity to luminance- and motion-dots using these parameters is captured by the different values of the semi-saturation values, σ_i . The measurements and model are one approach to connecting behavioral judgments to a quantitative model of the BOLD response in the VWFA. Future studies should refine this model

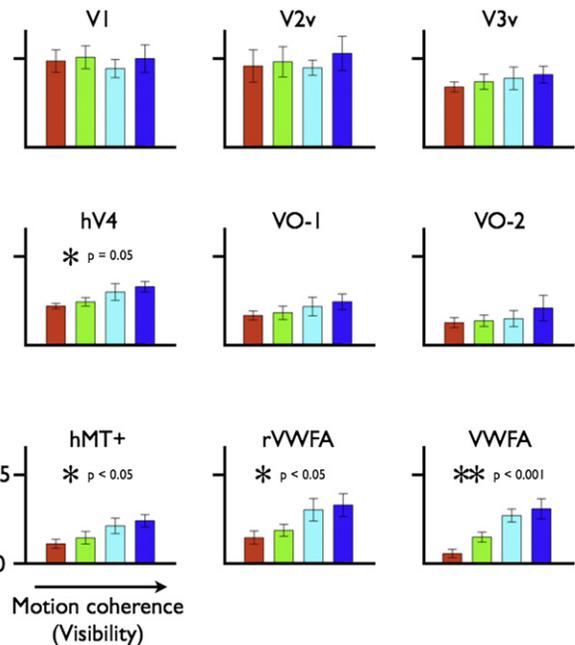


Figure 6. BOLD Response Amplitudes for Increasing Levels of Motion-Dot Word Visibility in Multiple Visual Field Maps and ROIs

The responses are shown for several left visual field maps (V1, the ventral portions of V2 and V3, hV4, VO-1/2), left hMT+, the VWFA and the right-hemisphere homolog of the VWFA (rVWFA). Responses for hMT+ and VWFA are as shown in Figures 2 and 4, respectively, and are included here for comparison. Response amplitude increases with motion coherence in hV4, hMT+, rVWFA, and VWFA. Other details as in Figure 2.

and test competing quantitative models to link behavioral and fMRI responses.

DISCUSSION

Neurological accounts of reading have a long history of emphasizing the importance of localized language regions (Broca, 1861; Dejerine, 1892; Wernicke, 1874) and efficient communication between these regions (Geschwind, 1965). However, there remains much to be learned about the sequence of transformations that occur between the initial visual word representation in primary visual cortex and specialized language areas (Dehaene et al., 2005).

The location of the VWFA, adjacent to several visual field maps (Figure 8) and object-selective regions, suggests that this part of the reading network is closely integrated with the visual hierarchy. However, many questions remain. Does the VWFA provide a feature-independent link between visual and language cortex? Is the VWFA specialized for recognizing particular combinations of line contours, such as T or Y junctions, or some more abstract shape representation that does not depend on line contours? Using word stimuli with unconventional stimulus features, we measured whether the VWFA responds to words using a feature-tolerant representation or whether words defined by features other than line contours communicate to language cortex via other routes (Figure 1).

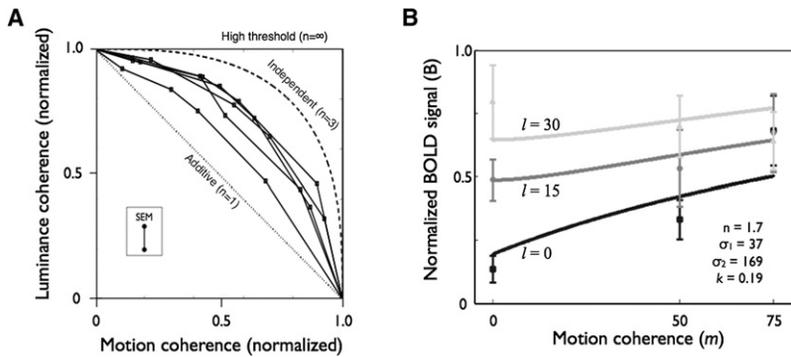


Figure 7. A Model of Responses to Combinations of Motion- and Luminance-Dot Features

(A) Psychophysical thresholds on a lexical decision task to combinations of luminance- and motion-dot features. The dotted line is the predicted performance if features combine additively. The dashed curve is the predicted performance from a probability summation model with an exponent of $n = 3$, which was the across-subject average value fit to the psychometric functions for motion-dot coherence and luminance-dot coherence separately. The outer boundary of the box is the predicted performance from a high-threshold model in which signals are completely independent. The features combine according to a rule that is sub-additive ($n = 1.7$) but more effective than pure probability summation. The inset shows ± 1 SEM across all subjects ($N = 5$) and mixture conditions.

(B) VWFA BOLD response amplitudes with increasing motion-dot coherence at different fixed luminance-dot coherence levels. The curves are predictions from a probability summation model (see main text). The black, dark gray, and light gray are measured response levels (points) and model predictions (curves) for the three luminance-dot coherence levels. The normalized BOLD signal is the VWFA response divided by the response in left V1. The model parameters are shown in the inset; the exponent ($n = 1.7$) is derived from the psychophysical data (A); the other parameters are fit to the data. See text for model details. Error bars are ± 1 SEM between subjects.

Flexible Systems-Level Circuitry Produces VWFA Feature-Tolerance

In the VWFA, word form responses are feature-independent; responses are virtually unchanged when word forms are defined by very different features (Figure 2). These results suggest that the signal transformations from visual cortex to the VWFA compute a shape representation that is abstracted from the specific stimulus features. When relating VWFA

BOLD responses to behavior, the VWFA is necessary but not sufficient for good reading performance (Figure 3). High VWFA activity does not guarantee good reading performance on a lexical decision task, but when VWFA activity is weak, reading performance is poor. This dissociation is true for all feature types, suggesting that the VWFA is a common bottleneck for information flow from visual to language cortex.

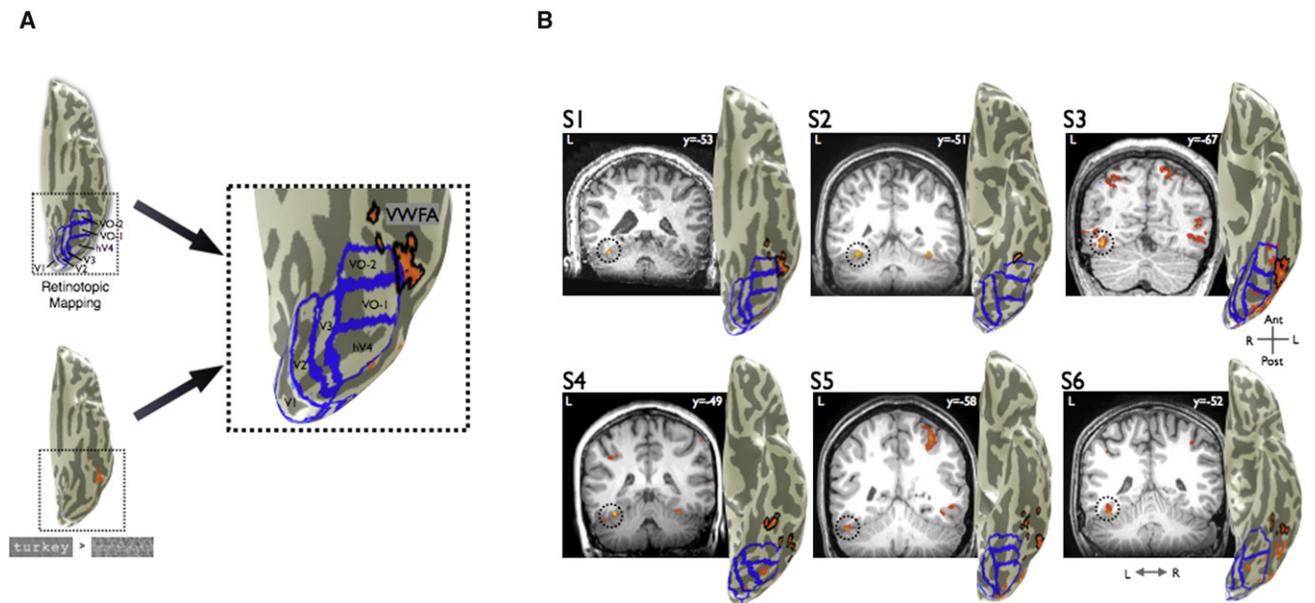


Figure 8. Location of the VWFA and Visual Field Maps

(A) In individual subjects, we performed retinotopic mapping to define the boundaries of multiple visual areas (V1, V2, V3, hV4, VO-1, VO-2). The map boundaries are shown by the blue lines. The VWFA localizer contrasted words with phase-scrambled words ($p < 0.001$, uncorrected). All significantly responsive gray matter voxels on the ventral occipito-temporal cortex anterior to hV4 and falling outside of known retinotopic areas were included in the VWFA ROI (outlined in black). (B) Coronal slices showing the position of the VWFA ROI for each subject; the MNI y-coordinate is shown in the inset. VWFA activation is outlined by dotted circles. Left hemisphere cortical surface renderings adjacent to each slice show a ventral view with all identifiable retinotopic areas outlined in blue, contrast maps in orange, and the VWFA outlined in black. We could not identify retinotopic area VO-2 in S4 and S6. A parietal activation seen in several slices, which is not studied in this paper, was also present routinely. See also Figures S2 and S3 and Table S1.

In other cortical areas, word form responses are feature-dependent (hMT+, Figure 4). The earliest visual processing stages segregate visual information into different channels that are optimized for different types of features, such as motion, color, or luminance (Livingstone and Hubel, 1988; Zeki, 1978). Changing the features of a given stimulus from luminance-contrast to motion-contrast evokes a response in a different set of retinal ganglion cells. These responses project to largely separate cortical streams (Ungerleider and Mishkin, 1982; Zeki et al., 1991). The BOLD responses in hMT+ suggest that motion-dot words were indeed processed by hMT+ (Figure 4). TMS experiments that disrupt hMT+ activity and thereby cause lower lexical decision task performance demonstrate that hMT+ signals are necessary for seeing motion-dot words (Figure 5).

Thus, despite early feature-specific divergence of signals into the dorsal and ventral streams, the word information reconverges from feature-specialized areas at or before the level of the VWFA. Depending on the stimulus features, signals are carried through different parts of cortex to the VWFA. Hence, future computational models of seeing words should not assume a fixed pathway through visual cortex, but they should allow for flexible connectivity of the VWFA.

VWFA as a Gateway between Vision and Language

Upon convergence of visual signals in the VWFA, outputs are sent to language areas. The VWFA may have a privileged position in human VOT cortex by virtue of its connections to language areas, such as the posterior superior temporal and inferior frontal gyri (Ben-Shachar et al., 2007c; Bokde et al., 2001). The language system probably requires word form signals to be represented in a specific format. It is possible that learning to see words and then representing the results in a format appropriate for language systems takes place in parallel cortical circuits, but it would seem inefficient to expect that the same complex learning takes place in multiple circuits. A conservative position to explain the current data is that the VWFA has uniquely evolved the capability of providing properly formatted sensory information to language areas (Devlin et al., 2006; Jobard et al., 2003). Another recent report supports this view, showing that the VWFA circuitry is useful in communicating even somatosensory data to language systems in congenitally blind subjects (Reich et al., 2011). Nevertheless, it remains possible that circuits not identified in this study are capable of both recognizing the sensory information and communicating the information to language (Richardson et al., 2011). If so, the circumstances in which these alternative routes are utilized should be further explored.

The format of word representations required by the language system is probably independent of most basic visual features, such as letter case and font (Dehaene et al., 2001; Polk and Farah, 2002; Qiao et al., 2010). Our results provide evidence that even when stimulus features initiate activation in different parts of early visual cortex, the VWFA can use the pattern of activity to recognize the presence of a word form. Yet this feature-tolerance cannot be based on learning, because our experience with words is specific to line contours and junctions. Learning in the VWFA and VOT related to word forms may instead be about the statistical regularities between abstract

shape representations (Binder et al., 2006; Dehaene et al., 2005; Glezer et al., 2009; Vinckier et al., 2007), independent of the specific visual features that define these shapes.

Origins of Feature-Tolerance in the Reading System

Feature-independent word form responses in the VWFA parallel feature-independent object responses in the nearby lateral occipital complex (Ferber et al., 2003; Grill-Spector et al., 1998; Kourtzi and Kanwisher, 2001). In the object recognition literature this feature-tolerance is thought to help recognize objects whose detailed properties (e.g., spectral radiance) can vary depending on viewing conditions (e.g., ambient lighting). The need for feature-tolerance is reduced in reading because words are typically differentiated by line-contours, but the capability may exist because the same cortical circuits produce the shape representations used for seeing words and objects. Rather than the VWFA specifically learning feature-tolerance for word shapes, feature-tolerance may be present throughout VOT for all shape recognition tasks, including word form recognition.

If feature-tolerant responses for words in humans are a consequence of general visual processing, then one might expect that these representations also exist in homologous regions of non-human primates. Feature-tolerant single unit responses in monkey inferotemporal cortex have been reviewed elsewhere (Logothetis and Sheinberg, 1996; Rolls, 2000). Briefly, electrophysiological studies of neural responses have shown feature-independent shape responses in single neurons in inferotemporal (IT) cortex of the nonhuman primate (Sáry et al., 1993), and lesions of IT cortex impair form-from-motion discrimination performance (Britten et al., 1992). Similarly feature-tolerant single cell responses are present in V4 (Logothetis and Charles, 1990; Mysore et al., 2006). The parallels between feature-tolerant responses in nonhuman primate IT cortex and human VOT cortex support the hypothesis that VWFA representations are derived from the same visual circuitry that creates all feature-tolerant shape responses.

Area hMT+ Is Not Necessary for Seeing Standard Words

The necessity of hMT+ for seeing motion-dot words (Figure 5) might have been surmised based on many human lesion studies (Blanke et al., 2007; Marcar et al., 1997; Regan et al., 1992; Vaina et al., 1990). Damage in the anatomical region around hMT+ can reduce shape-from-motion perception performance (Blanke et al., 2007; Marcar et al., 1997; Regan et al., 1992; Vaina et al., 1990), although not in all cases (Vaina, 1989; Vaina et al., 1990). Experiments in nonhuman primates have also shown that MT lesions produce shape discrimination deficits when forms are defined by motion but not luminance (Marcar and Cowey, 1992; Schiller, 1993).

More surprising is that TMS of hMT+ does not affect reading words defined by luminance-dots or line-contours (Figure 5). This lack of a disruptive effect by TMS suggests that hMT+ responses are not necessary for seeing standard words. These results are surprising because a large body of literature has shown correlations between reading skill and hMT+ BOLD responses to motion stimuli (Ben-Shachar et al., 2007a; Demb et al., 1997, 1998) with decreased hMT+ responses in dyslexics

(Eden et al., 1996). There are at least three possible explanations for why hMT+ responses are correlated with reading ability without assigning hMT+ a causal role in reading.

First, the development of rapid-processing pathways, including the magnocellular pathway, may be a prerequisite for the healthy development of other essential reading pathways (Witton et al., 1998). Between-subject differences in the development of the magnocellular pathway would be reflected in measurements of hMT+ responses, which primarily receive magnocellular input (Maunsell et al., 1990). These pathways may carry signals that coordinate development, but the signals may not be important for reading line-contour stimuli in the adult.

Second, hMT+ processing may be necessary for certain reading tasks, but not others. For example, hMT+ may be important for directing fixation and for passage reading, but not for single-word lexical decisions (Stein, 2003). The theory that hMT+ is necessary for correct saccadic eye movements could be tested in future TMS experiments that disrupt hMT+ neural processing at different latencies in relation to reading saccades.

Third, the experiments here show that signals from hMT+ can contribute to the VWFA responses. In normal adult reading this connection may not provide useful signals, but the connection is nevertheless present. Improper hMT+ development may produce noise that is transmitted to the VWFA through this connection and such noise may limit skilled reading.

Two previous TMS studies analyzed the necessity of hMT+ during reading. One study used several tasks and found a very small TMS influence only on a non-word reading task (Liederman et al., 2003); a second group found an effect of TMS on a visual word identification task (Laycock et al., 2009), while we used a lexical decision task. Another methodological difference between our study and previous studies is that we localized hMT+ using fMRI to ensure target specificity during TMS sessions. Liederman et al. used a TMS-based procedure and Laycock et al. used skull markers. The targeting method is important given the close proximity of area hMT+ to other visual areas (Wandell et al., 2007), as well as individual subject variability in hMT+ location in relation to skull (Sack et al., 2006) and even sulcal landmarks (Dumoulin et al., 2000). We took great care to direct TMS pulse trajectories to the center of individually defined hMT+ regions of interest in each subject. The TMS pulses are unlikely to have disrupted neural processing in nearby cortical areas (such as the VWFA) because the effect was limited to motion-dot words, while disruption of VWFA or early visual cortex would be expected to be detrimental to seeing all word stimuli.

Clinical Applications of Understanding Cortical Information Flow for Words

Understanding how information flow changes with stimulus features may be helpful in designing novel compensation strategies for people with reading difficulties (i.e., alexia or dyslexia). If we understand the flow of word information, it may be possible to change word stimulus properties in ways that force a re-routing of information through specific pathways (e.g., through hMT+). For instance, a patient reported by Epelbaum et al. (2008) showed alexia after damage to input pathways (inferior longitudinal fasciculus) to the VWFA. Conceivably, in such a patient one might access the anatomically intact VWFA using words

defined by unconventional features that can be communicated to the VWFA via preserved pathways.

This speculation is supported by the feature mixture experiments, which show that different stimulus features combine in a partially additive manner to boost performance over either feature alone (Figure 7A). A combination of stimulus features could benefit patients who have difficulty reading words drawn with line contours alone. In at least some patients with reading difficulties, rerouting word information through the magnocellular pathways may be beneficial (McCloskey and Rapp, 2000). The benefits may depend on the cortical location at which features combine in relation to the specific neural abnormalities, and future experiments in different types of dyslexic readers can test these hypotheses.

Conclusions

The early divergence of signals from early visual cortex into feature-specialized areas, followed by convergence in the VWFA, creates feature-tolerant representations of words. Depending on visual stimulus features, information about words is routed to different specialized areas. For example, words defined by motion features necessarily rely on hMT+ processing. In contrast, standard line contour words do not rely on hMT+. This result constrains the possible causal role of hMT+ in reading and suggests that hMT+ processing is not necessary for successful single word decoding under normal circumstances. After early specialized processing, signals reconverge in VOT cortex. The VWFA is well positioned to serve as a common gateway between orthographic and language processing. Such a gateway would benefit from a feature-tolerant, abstract shape representation. This type of abstract representation for words, a word form area, is advantageous for simplifying communication between early visual areas and the language system.

EXPERIMENTAL PROCEDURES

Subjects

Six subjects (3 females; ages 27–30, median age 28) participated in the main fMRI study. The study was approved by the institutional review board at Stanford University, and all subjects gave informed consent to participate in the study. Eight subjects (4 females; ages 19–58, median age 28.5) participated in the TMS experiments. Four subjects (1 female; 2 of the same subjects as main fMRI study, 2 different subjects; ages 24–29, median age 28) participated in the supplemental block-design fMRI experiment. All subjects were native English speakers and had normal or corrected-to-normal vision.

fMRI

Scanning Parameters

Anatomical and functional imaging data were acquired on a 3T General Electrical scanner using an 8-channel head coil. Subject head motion was minimized by placing padding around the head. Functional MR data were acquired using a spiral pulse sequence (Glover, 1999). Thirty 2.5-mm-thick coronal oblique slices oriented approximately perpendicular to the calcarine sulcus were prescribed. These slices covered the whole occipital lobe and parts of the temporal and parietal lobes. Data were acquired using the following parameters: acquisition matrix size = 64 × 64, FOV = 180 mm, voxel size of 2.8 × 2.8 × 2.5 mm, TR = 2000 ms, TE = 30 ms, flip angle = 77°. Some retinotopy scans were acquired with 24 similarly oriented slices at a different resolution (1.25 × 1.25 × 2 mm, TR = 2000 ms, TE = 30 ms). Using a back-bore projector, stimuli were projected onto a screen that the subject viewed through a mirror fixed above the head. The screen subtended a radius of

12 degrees along the vertical dimension. A custom MR-compatible eye tracker mounted to the mirror continuously recorded (software: ViewPoint, Arrington Research, Arizona, USA) eye movements to ensure good fixation performance during scanning sessions.

T1-weighted anatomical images were acquired separately for each subject, as described in the [Supplemental Experimental Procedures](#). General analysis procedures are also described therein.

Region of Interest Identification

We identified several functional brain areas (early visual areas [V1, V2, V3, hV4, VO-1, VO-2], hMT+, and VWFA) using separate localizer scans conducted within a single session (multiple runs) for each subject individually. The BOLD activation was measured within these regions of interest. The VWFA localizer is described below. Please see [Supplemental Experimental Procedures](#) for hMT+ localizer and retinotopy descriptions. Retinotopic mapping was performed following previously published methods (Dumoulin and Wandell, 2008).

VWFA Localizer. The visual word form area (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 per block. Words were six-letter nouns with a minimum word frequency of seven per million (Medler and Binder, 2005). The size of all stimuli was 14.2×4.3 degrees. 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 as 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.

The VWFA was defined in each subject as the activation on the ventral cortical surface from a contrast between words and phase-scrambled words ($p < 0.001$, uncorrected, [Figure 8](#)). The region was restricted to responsive voxels outside retinotopic areas and anterior to hV4. The Montreal Neurological Institute (MNI) coordinates of the peak voxel within the region of interest (ROI) was identified by finding the best-fitting transform between the individual T1-weighted anatomy with the average MNI T1-weighted anatomy and then applying that transform to the peak voxel within the VWFA for the same contrast.

The VWFA ROIs are located near the left lateral occipitotemporal sulcus ([Figure 8B](#), MNI coordinates in [Table S1](#), mean MNI coordinates: $-41 -57 -23$) and within ~ 5 mm of previous reports (Ben-Shachar et al., 2007b; Cohen et al., 2000, 2002, 2003). In 5 out of 6 subjects activations were bilateral, while in the remaining subject the activation was left-lateralized. In this manuscript, unless otherwise specified, VWFA refers to the left-hemisphere ROI. In all subjects a contrast of words versus checkerboards produces regions of interest in virtually identical locations and of similar size ([Figure S3](#)).

The ability to identify regions of interest in ventral occipital temporal cortex is limited by measurement artifacts caused by (1) the large transverse sinus (Winawer et al., 2010) and (2) susceptibility introduced by the auditory canals. The locations of these artifacts can be estimated in each subject and they are summarized in [Figure S4](#). These artifacts limit our ability to measure a portion of the VWFA in some subjects.

Experimental Design

The main experiment consisted of separate sessions (on separate days) for each feature type (line contours, motion-dot, luminance-dot, and mixture). Each subject completed six runs (312 s per run) for each feature type. The order of feature types was counterbalanced across subjects. Subjects were asked to keep fixation on a central fixation dot while reading the stimuli and to indicate by button press whether each stimulus was a word or pseudoword (i.e., lexical decision task). Eye movements were monitored (see above). We measured the BOLD response to words and pseudowords at four different visibility levels for each feature type. In analyzing the data, we grouped words and pseudowords together because they showed similar responses in all regions of interest that we examined.

Stimuli

All stimuli used for the main experimental runs were four-letter words or pseudowords (Medler and Binder, 2005). Words were nouns with a frequency of at least four per million (median: 28 per million). All words ($n = 480$) and pseudo-

words ($n = 480$) were unique within each subject, with five words and five pseudowords ($\times 4$ visibility levels $\times 6$ runs/feature $\times 4$ feature types) being assigned randomly to each of four visibility levels within each run (40 stimuli per run). All stimuli were shown for 2 s. Stimulus presentation and response collection, both for fMRI and TMS (see below), were created using custom Matlab (The MathWorks, Inc.) scripts and controlled using the Psychtoolbox (Brainard, 1997). The stimuli were created as follows:

Line Contours. The procedure used for rendering standard words at different visibility levels was similar to that used by Ben-Shachar and colleagues (2007b). We rendered words in black using the Monospaced (Sans Serif) font within a gray rectangular frame (24 degrees horizontal, 7 degrees vertical). The horizontal and vertical spans of the word within the frame were approximately 7.5 and 2.5 degrees, respectively (height of an x character was approximately 2°). To obtain different degrees of visibility, we computed the 2D Fourier transform of the word image, randomized the phase, and then applied the inverse Fourier transform. Visibility could be controlled by the degree of offset between the old and new phase. Resulting images ranged from noise (fully phase-scrambled) that contained the same amplitude spectrum as the original images, to highly visible words.

Luminance-Dot. To create words defined by dots of spatially varying luminance, we replaced the word image with a field of dots (dot density = 0.3; dot size = 1 pixel, total image size = 600×180 pixels), keeping the background color a uniform gray. The luminance of the dots was set separately for dots that fall inside (black) or outside (white) the nominal borders of the word form. Word visibility was controlled by the luminance coherence. For example, at a luminance coherence of 100%, all dots falling within the word form were black, and all dots outside the word form were white. For a luminance coherence of 50%, half the dots within the word form would be set to black (and half the dots outside the word form to white), while the rest of the dots (noise dots) were set randomly to black or white. Similarly, at 0% luminance coherence, all dots were randomly set to black or white, and thus no information about the original word form was present in the image. The values of luminance coherence used in this study were 0%, 15%, 30%, and 45%. The dots moved either left or right over successive frames (dot life = 4 frames, frame rate 60 Hz). For luminance-dot words, the motion of each dot was set randomly to left or right (0% motion coherence). The motion direction of each dot remained unchanged for 4 frames, at which point this dot disappeared and a new dot appeared in a random location to replace it.

Motion-Dot. For motion-dot words, word form was encoded by the direction of dot motion. The procedure for making these stimuli was identical to that used for making luminance-dot words, except that visibility was controlled by motion coherence, and dot luminance was randomly set to black or white. Signal dots moved to the right if they fell within the word form and to the left if they fell outside it (dot life = 4 frames). All other dots were noise dots and were therefore randomly assigned a leftward or rightward direction. Motion coherence, like luminance coherence, controlled the percentage of signal dots. The values of motion coherence were 0%, 50%, 75%, and 100%. The actual values of luminance and motion coherence are not meaningful in that their precise relationship to visibility depends critically on many other stimulus parameters, such as dot size, stimulus size, and dot density. Therefore we chose values that produced approximately similar visibility levels, from complete noise to fully visible, based on initial psychophysical piloting with our stimulus parameters.

Mixture. This stimulus type was constructed identically to the motion-dot and luminance-dot stimulus types. Four conditions were chosen by adjusting both luminance and motion coherence of the stimuli, as described above. The luminance and motion coherence values matched the middle two coherence values for the luminance-dot and motion-dot stimuli (thus producing $2 \times 2 = 4$ total conditions).

Examples illustrating the two dynamic stimuli and the line contour stimulus are included in the [Figure S2](#) and [Movies S1](#) and [S2](#).

TMS

To examine the necessity of area hMT+ for reading words of different stimulus features, we used transcranial magnetic stimulation (TMS) and targeted the center of the functional hMT+ ROI defined for each individual subject. We used theBrainsight 2 neuronavigation system (Rogue Research, Inc; Montreal,

Canada) for precisely targeting the area of functional activation. A Magstim (The Magstim Company, UK) figure-of-eight coil was used for dual-pulse stimulation (45 ms between pulses) at 60% maximum stimulator output. The time between stimulus onset and onset of the first TMS pulse (stimulus-pulse onset asynchrony; SOA) was controlled using Matlab (The MathWorks, Inc; Massachusetts, USA). We used the following SOAs: -95, 5, 87, 165, 264, and 885 ms (± 5 ms error). Stimuli were randomly chosen from a set of 504 four-letter words and pseudowords with the same properties as those described for fMRI. As for fMRI data analysis, words and pseudowords were grouped in analyzing the TMS data. Chance performance for the task was 50%, since half the stimuli were words and half were pseudowords.

Stimuli were identical to those used for the main fMRI experiment, except that the stimulus duration was limited to one second, plus a one second response time window (total trial time = 2 s). The lexical decision task was also identical: subjects indicated via button press whether the stimulus on the screen was a word or a pseudoword. In contrast to the fMRI experiments, however, the degree of phase-scrambling, motion coherence, and luminance coherence were set according to psychophysical lexical visibility thresholds acquired directly before the main TMS experiment. For each feature type, we used standard psychophysical procedures to measure subjects' individual stimulus thresholds for visibility such that subjects achieved 82% correct on a lexical decision task at the same viewing distance as used during the TMS session. This baseline performance criterion was chosen so that disturbances in task performance caused by TMS would be reflected by a lower percent correct.

After setting psychophysical thresholds, the TMS sessions consisted of 3 runs of 72 trials each (3 stimulus feature types \times 6 SOAs \times 2 lexical classes \times 2 exemplars per run). Trials were spaced on average 4 s apart (jitter based on a Poisson distribution with mean of 4000 ms, adjusted to have a minimum of 2 s between trials). Thus, each run was approximately 430 s long. The order and exact timing of stimuli within each run was randomized across subjects. Subjects were asked to fixate on a central fixation dot throughout the duration of the run. The fixation dot was present during and between stimulus presentations. Fixation performance was monitored by the experimenters in the room, and all subjects maintained excellent fixation. Head position was maintained using a forehead rest. Subjects received short (~ 5 min) breaks between runs.

Psychophysics

In the behavioral mixture experiments, subjects were presented four-letter words and pseudowords defined by a combination of luminance- and motion-dots set to one of five different coherence ratios. The feature coherence of both features was scaled by a common factor across trials, preserving the ratio of coherences. The probability of making a correct lexical decision was measured using a staircase procedure. A threshold level (82%) was estimated from the responses to multiple coherence levels. The thresholds at each of the different ratios are shown in Figure 7A.

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures, one table, Supplemental Experimental Procedures, and two movies and can be found with this article online at doi:10.1016/j.neuron.2011.06.036.

ACKNOWLEDGMENTS

We thank Michal Ben-Shachar and Jason Yeatman for their advice. This work was supported by the Medical Scientist Training Program, the Bio-X Graduate Student Fellowship Program, and NIH RO1 Grant EY015000.

Accepted: June 16, 2011

Published: September 7, 2011

REFERENCES

Ben-Shachar, M., Dougherty, R.F., Deutsch, G.K., and Wandell, B.A. (2007a). Contrast responsivity in MT+ correlates with phonological awareness and reading measures in children. *Neuroimage* 37, 1396–1406.

Ben-Shachar, M., Dougherty, R.F., Deutsch, G.K., and Wandell, B.A. (2007b). Differential sensitivity to words and shapes in ventral occipito-temporal cortex. *Cereb. Cortex* 17, 1604–1611.

Ben-Shachar, M., Dougherty, R.F., and Wandell, B.A. (2007c). White matter pathways in reading. *Curr. Opin. Neurobiol.* 17, 258–270.

Binder, J.R., Medler, D.A., Westbury, C.F., Liebenthal, E., and Buchanan, L. (2006). Tuning of the human left fusiform gyrus to sublexical orthographic structure. *Neuroimage* 33, 739–748.

Blanke, O., Brooks, A., Mercier, M., Spinelli, L., Adriani, M., Lavanchy, L., Safran, A.B., and Landis, T. (2007). Distinct mechanisms of form-from-motion perception in human extrastriate cortex. *Neuropsychologia* 45, 644–653.

Bokde, A.L., Tagamets, M.A., Friedman, R.B., and Horwitz, B. (2001). Functional interactions of the inferior frontal cortex during the processing of words and word-like stimuli. *Neuron* 30, 609–617.

Braddick, O.J., O'Brien, J.M., Wattam-Bell, J., Atkinson, J., Hartley, T., and Turner, R. (2001). Brain areas sensitive to coherent visual motion. *Perception* 30, 61–72.

Brainard, D.H. (1997). The Psychophysics Toolbox. *Spat. Vis.* 10, 433–436.

Brem, S., Bucher, K., Halder, P., Summers, P., Dietrich, T., Martin, E., and Brandeis, D. (2006). Evidence for developmental changes in the visual word processing network beyond adolescence. *Neuroimage* 29, 822–837.

Britten, K.H., Newsome, W.T., and Saunders, R.C. (1992). Effects of inferotemporal cortex lesions on form-from-motion discrimination in monkeys. *Exp. Brain Res.* 88, 292–302.

Broca, P. (1861). Nouvelle observation d'aphémie produite par une lésion de la troisième circonvolution frontale. *Bulletins de la Société d'anatomie (Paris)*, 2e serie 6, 398–407.

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.A., and 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., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., and Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain* 125, 1054–1069.

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

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

Dehaene, S., Naccache, L., Cohen, L., Bihan, D.L., Mangin, J.F., Poline, J.B., and Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nat. Neurosci.* 4, 752–758.

Dehaene, S., Cohen, L., Sigman, M., and Vinckier, F. (2005). The neural code for written words: a proposal. *Trends Cogn. Sci. (Regul. Ed.)* 9, 335–341.

Dejerine, J. (1892). Contribution à l'étude anatomo-pathologique et clinique des différentes variétés de cécité verbale. *Mémoires de la Société de Biologie* 4, 61–90.

Demb, J.B., Boynton, G.M., and Heeger, D.J. (1997). Brain activity in visual cortex predicts individual differences in reading performance. *Proc. Natl. Acad. Sci. USA* 94, 13363–13366.

Demb, J.B., Boynton, G.M., and Heeger, D.J. (1998). Functional magnetic resonance imaging of early visual pathways in dyslexia. *J. Neurosci.* 18, 6939–6951.

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

DiCarlo, J.J., and Cox, D.D. (2007). Untangling invariant object recognition. *Trends Cogn. Sci. (Regul. Ed.)* 11, 333–341.

Dumoulin, S.O., and Wandell, B.A. (2008). Population receptive field estimates in human visual cortex. *Neuroimage* 39, 647–660.

Dumoulin, S.O., Bittar, R.G., Kabani, N.J., Baker, C.L., Jr., Le Goualher, G., Bruce Pike, G., and Evans, A.C. (2000). A new anatomical landmark for reliable

- identification of human area V5/MT: a quantitative analysis of sulcal patterning. *Cereb. Cortex* 10, 454–463.
- Eden, G.F., VanMeter, J.W., Rumsey, J.M., Maisog, J.M., Woods, R.P., and Zeffiro, T.A. (1996). Abnormal processing of visual motion in dyslexia revealed by functional brain imaging. *Nature* 382, 66–69.
- Epelbaum, S., Pinel, P., Gaillard, R., Delmaire, C., Perrin, M., Dupont, S., Dehaene, S., and Cohen, L. (2008). Pure alexia as a disconnection syndrome: new diffusion imaging evidence for an old concept. *Cortex* 44, 962–974.
- Ferber, S., Humphrey, G.K., and Vilis, T. (2003). The lateral occipital complex subserves the perceptual persistence of motion-defined groupings. *Cereb. Cortex* 13, 716–721.
- Galaburda, A., and Livingstone, M. (1993). Evidence for a magnocellular defect in developmental dyslexia. *Ann. N Y Acad. Sci.* 682, 70–82.
- Geschwind, N. (1965). Disconnexion syndromes in animals and man. I. *Brain* 88, 237–294.
- Glezer, L.S., Jiang, X., and Riesenhuber, M. (2009). Evidence for highly selective neuronal tuning to whole words in the “visual word form area”. *Neuron* 62, 199–204.
- Glover, G.H. (1999). Simple analytic spiral K-space algorithm. *Magn. Reson. Med.* 42, 412–415.
- Graham, N. (1989). *Visual Pattern Analyzers* (New York, NY: Oxford University Press).
- Graham, N., Robson, J.G., and Nachmias, J. (1978). Grating summation in fovea and periphery. *Vision Res.* 18, 815–825.
- Grill-Spector, K., Kushnir, T., Edelman, S., Itzhak, Y., and Malach, R. (1998). Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron* 21, 191–202.
- Jobard, G., Crivello, F., and Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: a meta-analysis of 35 neuroimaging studies. *Neuroimage* 20, 693–712.
- Kourtzi, Z., and Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science* 293, 1506–1509.
- Laycock, R., Crewther, D.P., Fitzgerald, P.B., and Crewther, S.G. (2009). TMS disruption of V5/MT+ indicates a role for the dorsal stream in word recognition. *Exp. Brain Res.* 197, 69–79.
- Liederman, J., McGraw Fisher, J., Schulz, M., Maxwell, C., Théoret, H., and Pascual-Leone, A. (2003). The role of motion direction selective extrastriate regions in reading: a transcranial magnetic stimulation study. *Brain Lang.* 85, 140–155.
- Livingstone, M., and Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science* 240, 740–749.
- Livingstone, M.S., Rosen, G.D., Drislane, F.W., and Galaburda, A.M. (1991). Physiological and anatomical evidence for a magnocellular defect in developmental dyslexia. *Proc. Natl. Acad. Sci. USA* 88, 7943–7947.
- Logothetis, N., and Charles, E. (1990). V4 Responses to gratings defined by random dot motion. *Invest. Ophthalmol. Vis. Sci. (Supplement)*, 30.
- Logothetis, N.K., and Sheinberg, D.L. (1996). Visual object recognition. *Annu. Rev. Neurosci.* 19, 577–621.
- Marcas, V.L., and Cowey, A. (1992). The Effect of Removing Superior Temporal Cortical Motion Areas in the Macaque Monkey: II. Motion Discrimination Using Random Dot Displays. *Eur. J. Neurosci.* 4, 1228–1238.
- Marcas, V.L., Zihl, J., and Cowey, A. (1997). Comparing the visual deficits of a motion blind patient with the visual deficits of monkeys with area MT removed. *Neuropsychologia* 35, 1459–1465.
- Martin, F., and Lovegrove, W. (1987). Flicker contrast sensitivity in normal and specifically disabled readers. *Perception* 16, 215–221.
- Maunsell, J.H., Nealey, T.A., and DePriest, D.D. (1990). Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey. *J. Neurosci.* 10, 3323–3334.
- McCloskey, M., and Rapp, B. (2000). A visually based developmental reading deficit. *J. Mem. Lang.* 43, 157–181.
- Medler, D.A., and Binder, J.R. (2005). MCWord: An on-line orthographic database of the English language.
- Mysore, S.G., Vogels, R., Raiguel, S.E., and Orban, G.A. (2006). Processing of kinetic boundaries in macaque V4. *J. Neurophysiol.* 95, 1864–1880.
- Polk, T.A., and Farah, M.J. (2002). Functional MRI evidence for an abstract, not perceptual, word-form area. *J. Exp. Psychol. Gen.* 131, 65–72.
- Price, C.J., and Devlin, J.T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends Cogn. Sci. (Regul. Ed.)* 15, 246–253.
- Prieto, E.A., Barnikol, U.B., Soler, E.P., Dolan, K., Hesselmann, G., Mohlberg, H., Amunts, K., Zilles, K., Niedeggen, M., and Tass, P.A. (2007). Timing of V1/V2 and V5+ activations during coherent motion of dots: an MEG study. *Neuroimage* 37, 1384–1395.
- Qiao, E., Vinckier, F., Szwed, M., Naccache, L., Valabrègue, R., Dehaene, S., and Cohen, L. (2010). Unconsciously deciphering handwriting: subliminal invariance for handwritten words in the visual word form area. *Neuroimage* 49, 1786–1799.
- Quick, R.F., Jr. (1974). A vector-magnitude model of contrast detection. *Kybernetik* 16, 65–67.
- Raiguel, S.E., Xiao, D.K., Marcar, V.L., and Orban, G.A. (1999). Response latency of macaque area MT/V5 neurons and its relationship to stimulus parameters. *J. Neurophysiol.* 82, 1944–1956.
- Regan, D., Giaschi, D., Sharpe, J.A., and Hong, X.H. (1992). Visual processing of motion-defined form: selective failure in patients with parietotemporal lesions. *J. Neurosci.* 12, 2198–2210.
- Reich, L., Szwed, M., Cohen, L., and Amedi, A. (2011). A ventral visual stream reading center independent of visual experience. *Curr. Biol.* 21, 363–368.
- Richardson, F.M., Seghier, M.L., Leff, A.P., Thomas, M.S., and Price, C.J. (2011). Multiple routes from occipital to temporal cortices during reading. *J. Neurosci.* 31, 8239–8247.
- Rolls, E.T. (2000). Functions of the primate temporal lobe cortical visual areas in invariant visual object and face recognition. *Neuron* 27, 205–218.
- Sack, A.T., Kohler, A., Linden, D.E., Goebel, R., and Muckli, L. (2006). The temporal characteristics of motion processing in hMT/V5+: combining fMRI and neuronavigated TMS. *Neuroimage* 29, 1326–1335.
- Sáry, G., Vogels, R., and Orban, G.A. (1993). Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science* 260, 995–997.
- Schiller, P.H. (1993). The effects of V4 and middle temporal (MT) area lesions on visual performance in the rhesus monkey. *Vis. Neurosci.* 10, 717–746.
- Stein, J. (2003). Visual motion sensitivity and reading. *Neuropsychologia* 41, 1785–1793.
- Szwed, M., Cohen, L., Qiao, E., and Dehaene, S. (2009). The role of invariant line junctions in object and visual word recognition. *Vision Res.* 49, 718–725.
- Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabrègue, R., Amadon, A., and Cohen, L. (2011). Specialization for written words over objects in the visual cortex. *Neuroimage* 56, 330–344.
- Tarkiainen, A., Helenius, P., Hansen, P.C., Cornelissen, P.L., and Salmelin, R. (1999). Dynamics of letter string perception in the human occipitotemporal cortex. *Brain* 122, 2119–2132.
- Ungerleider, L.G., and Mishkin, M. (1982). Two cortical visual systems. In *Analysis of Visual Behavior*, D.J. Ingle, M.A. Goodale, and M.R.J. W, eds. (Cambridge, MA: MIT Press), pp. 549–586.
- Vaina, L.M. (1989). Selective impairment of visual motion interpretation following lesions of the right occipito-parietal area in humans. *Biol. Cybern.* 61, 347–359.
- Vaina, L.M., Lemay, M., Bienfang, D.C., Choi, A.Y., and Nakayama, K. (1990). Intact “biological motion” and “structure from motion” perception in a patient with impaired motion mechanisms: a case study. *Vis. Neurosci.* 5, 353–369.
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J.P., Sigman, M., and Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron* 55, 143–156.
- Wandell, B.A., Dumoulin, S.O., and Brewer, A.A. (2007). Visual field maps in human cortex. *Neuron* 56, 366–383.

Wandell, B.A., Rauschecker, A.M., and Yeatman, J.D. (2010). Learning to see words. *Annu. Rev. Psychol.*, in press. Published online December 8, 2010. 10.1146/annurev-psych-120710-100434.

Warrington, E.K., and Shallice, T. (1980). Word-form dyslexia. *Brain* 103, 99–112.

Wernicke, C. (1874). *Der Aphasische Symptomencomplex: Eine Psychologische Studie auf Anatomischer Basis* (Breslau: Max Cohn & Weigert).

Winawer, J., Horiguchi, H., Sayres, R.A., Amano, K., and Wandell, B.A. (2010). Mapping hV4 and ventral occipital cortex: the venous eclipse. *J. Vis.* 10, 1.

Witton, C., Talcott, J.B., Hansen, P.C., Richardson, A.J., Griffiths, T.D., Rees, A., Stein, J.F., and Green, G.G. (1998). Sensitivity to dynamic auditory and visual stimuli predicts nonword reading ability in both dyslexic and normal readers. *Curr. Biol.* 8, 791–797.

Zeki, S.M. (1978). Functional specialisation in the visual cortex of the rhesus monkey. *Nature* 274, 423–428.

Zeki, S., Watson, J.D., Lueck, C.J., Friston, K.J., Kennard, C., and Frackowiak, R.S. (1991). A direct demonstration of functional specialization in human visual cortex. *J. Neurosci.* 11, 641–649.