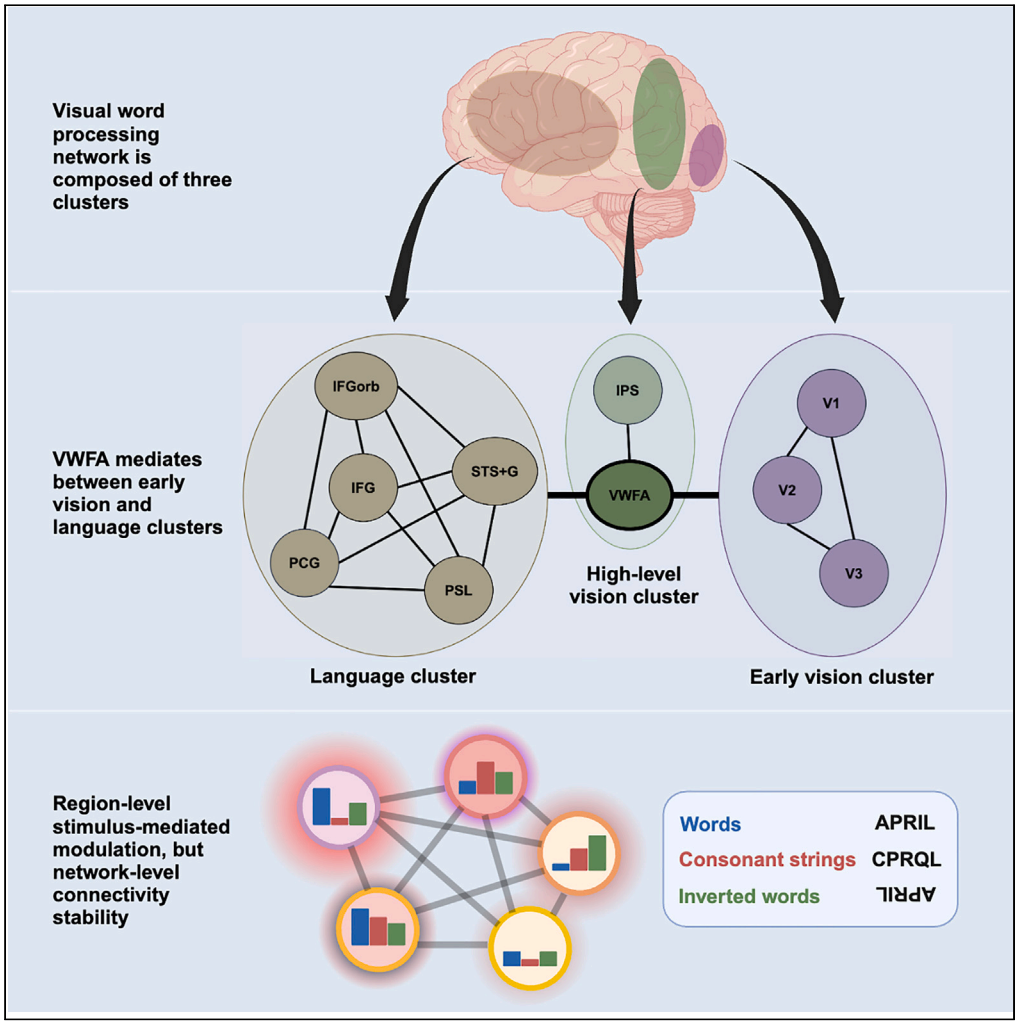
iScience

CellPress
OPEN ACCESS

Article

Visual word processing engages a hierarchical, distributed, and bilateral cortical network



Raina Vin, Nicholas M. Blauch, David C. Plaut, Marlene Behrmann

raina.vin@yale.edu (R.V.)
mbehrmann@pitt.edu (M.B.)

Highlights

The visual word processing network is modulated bilaterally by textual stimulus type

The network is composed of three clusters arrayed on the posterior-anterior brain axis

VWFA emerges as a mediator between the early vision and language clusters

Region-level stimulus-mediated modulation, but network-level stability, is observed

Vin et al., iScience 27, 108809
February 16, 2024
https://doi.org/10.1016/j.isci.2024.108809

iScience

CellPress
OPEN ACCESS

Article

Visual word processing engages a hierarchical, distributed, and bilateral cortical network

Raina Vin,[1,2,5,*] Nicholas M. Blauch,[1,3] David C. Plaut,[1,2] and Marlene Behrmann[1,2,4,6,*]

SUMMARY

Although the Visual Word Form Area (VWFA) in left temporal cortex is considered the pre-eminent region in visual word processing, other regions are also implicated. We examined the entire text-selective circuit, using functional MRI. Ten regions of interest (ROIs) per hemisphere were defined, which, based on clustering, grouped into early vision, high-level vision, and language clusters. We analyzed the responses of the ROIs and clusters to words, inverted words, and consonant strings using univariate, multivariate, and functional connectivity measures. Bilateral modulation by stimulus condition was evident, with a stronger effect in left hemisphere regions. Last, using graph theory, we observed that the VWFA was equivalently connected with early visual and language clusters in both hemispheres, reflecting its role as a mediator in the circuit. Although the individual ROIs and clusters bilaterally were flexibly altered by the nature of the input, stability held at the level of global circuit connectivity, reflecting the complex hierarchical distributed system serving visual text perception.

INTRODUCTION

Long-standing data from premorbidly literate patients with acquired reading deficits following brain insult, along with more recent electrophysiological findings from intracranial depth electrodes in humans, have uncovered causal evidence critically implicating the Visual Word Form Area (VWFA)—a small region in the ventral temporal cortex (VTC) of the left hemisphere (LH)—in the processing of text.[4,15,20,50,66,79,81] These findings have been complemented by functional magnetic resonance imaging (fMRI) investigations demonstrating disproportionate selectivity in this region for written words over non-words, faces, and objects.[2,70] Notwithstanding the identification of the VWFA as a key structure in text processing, a comprehensive account of the neural substrate underlying the perception of text and its modulation as a function of task and stimulus parameters is still lacking. The primary goal of the current investigation is to contribute to such an account by adopting a data-driven wide lens on the neural basis of text processing, examining a swath of cortex in both the left and right hemispheres.

Approaches to uncovering the neural circuitry of text processing

To date, many approaches have emerged as part of the effort to map out the neural circuitry underlying text processing and its flexibility in response to the input. Below, we roughly group studies into three approaches that have somewhat different foci but are not strictly mutually exclusive.

The first approach views the VWFA as the pre-eminent region of text processing, and has analyzed, in greater detail, its functional response properties. Some work in this vein has challenged the traditional notion of the VWFA as computing pure visual word forms, pointing out that the response properties of VWFA are more complex[14,61,63] than originally considered. For example, several functional imaging studies have demonstrated that the VWFA responds not only to visual stimuli but also to auditory input, specifically when participants hear auditory words and then repeat or think about the meaning of the words or make rhyming judgements,[19,47,62] and even to tactile input as revealed by activation when congenitally blind subjects read tactile words in Braille (Buchel et al., 1998). Other studies have retained a focus on the role of the VWFA as the major hub of reading. Close scrutiny of the VWFA has indicated that most subjects show a subdivision of multiple cortical patches, rather than a single word-selective area in VTC.[29,46,67,75,84] This research has suggested that the posterior subregion (VWFA-1) mediates visual perceptual representations, whereas the anterior subregion (VWFA-2) mediates language representations.[12,34,82] The activation in the anterior subregion has been shown to predict reading behavior outside of the scanner.[46] Consistent with the functional distinctions, posterior and anterior subregions of the VWFA differ in their micro-architectonic properties[29,46,67,75] and in their connectivity profiles.

[1]Neuroscience Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[2]Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[3]Program in Neural Computation, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[4]Department of Ophthalmology, University of Pittsburgh, Pittsburgh, PA 15219, USA
[5]Interdepartmental Neuroscience Program, Yale University, New Haven, CT 06520, USA
[6]Lead contact
*Correspondence: raina.vin@yale.edu (R.V.), mbehrmann@pitt.edu (M.B.)
https://doi.org/10.1016/j.isci.2024.108809

Specifically, the posterior subregion has direct structural connections with early visual regions in the occipital cortex, as well as with the intraparietal sulcus via the ventral occipital fasciculus,[83] whereas the anterior region has stronger connectivity with language regions.[8,46,64]

Although this line of research has enhanced our understanding of the response properties and subdivisions of word-selective cortex in the ventral occipital cortex, it has often left unexplored the functional interactions of a broader circuit encompassing visual and language regions, despite demonstrations of strong connectivity between VWFA and these regions. A second line of research has, by contrast, taken a broader look at cortical regions implicated in text processing and has evaluated the relationships between these regions. Lateral temporal regions in and around the superior temporal sulcus (STS), as well as regions in and around the inferior parietal sulcus (IPS) have been identified.[10,13] As determined by data from almost 1,000 readers, there are "core" regions, such as the insular and frontal opercular cortex, lateral temporal cortex, and early auditory cortex whose profiles are positively correlated with that of VWFA, and, unsurprisingly, whose responses are strong predictors of reading performance[43] (for further review see[12]). The interrelationships between the VWFA and these other regions of interest (ROIs) have been largely established by functional connectivity (FC) analyses; significant FC is observed between the VWFA and Broca's area, LH Precentral Gyrus (PCG) and Wernicke's area even during resting-state scanning. Furthermore, one study using perisylvian subdural electrodes reported that the pars opercularis, the supramarginal gyrus, and the superior temporal gyrus all showed broadband high gamma responses in response to text input.[39]

Beyond this, the VWFA and, in fact, most occipitotemporal regions of interest are connected to the dorsal frontoparietal attention network, either structurally or functionally,[9,51,77,83] and the strength of the connectivity is sufficiently strong that the direct top-down modulation to the VWFA[13,64] is even greater than that observed in any other region of visual cortex.[48] The bilateral enhancement of the VWFA response from the dorsal frontoparietal areas can modulate reading performance,[6] and altered FC between visual association and prefrontal attention areas[27]—specifically, between the inferior parietal lobule and the VWFA[58]—appears to play a causal role in some types of dyslexia. Relatedly, the FC between VWFA and other regions is heightened for slow compared with fast readers, perhaps indicative of the need of the former to constrain responses further with top-down modulation,[25] and the VWFA can be upregulated using neurofeedback, hinting at a potential intervention for those with difficulties in text processing.[35] Last, poor readers show atypical topology compared with good readers, with a less functionally specialized cluster around the left inferior frontal gyrus and surprisingly stronger functional links between bilateral STGs and other regions[52] and between the left IFG and OTC.[76]

The third and final approach to understanding the neural basis of text processing has, as its major aim, the evaluation of the modifiability of the nodes and the weights between regions and clusters as a function of task demands and properties of the input. This approach has received the least attention in the literature to date. Relevant task-specific findings include results showing that the strength of the VWFA-Wernicke's FC predicts performance on a semantic classification task with words but not with other categories of visual stimuli[70] and that the FC between VWFA and Broca's area is twice as strong during a lexical task than during a task that required focused attention on the fixation point.[77] Of interest too is that the engagement of these areas can also be modulated by the specific task demands, with enhanced connectivity during lexical and semantic tasks.[59,70,77] In a causal study, using TMS and EEG jointly, Planton et al.[59] observed modification of cortical sensitivity: during TMS stimulation to the left occipitotemporal area, effective connectivity between the VWFA and other brain regions varied in both a task-dependent and stimulus-dependent fashion. Specifically, they showed that, following TMS, signal propagation to the right visual regions as well as to parietal and temporal areas was significantly slowed during text processing versus non-text processing tasks. Moreover, the strengths of the functional interactions between the different regions of this distributed neural network were modulated by the type of stimulus (e.g., words vs. pseudowords). Last, in a 7T study, Chinese-English bilinguals had Chinese-specific regions of activation, as well as regions activated to script in general, but this was not so for French-English bilinguals, whose left occipitotemporal region was activated by inputs from both languages.[84] Taken together, these findings suggest that visual text processing is subserved by a widely distributed network of LH regions with a complex system of functional and structural interactions and that perturbation of this network can adversely impact reading performance. Given the different foci of these studies, however, it is difficult to obtain a comprehensive view of all the regions and their interrelations and the reconfiguration of the network by task and stimulus type.

The current study

In the current work, we attempt to characterize the organization of the entire circuit that is functionally implicated in text processing. We go beyond existing studies by initially demarcating a large number of left hemisphere (LH) ROIs previously implicated in text processing as well as their homologous right hemisphere (RH) counterparts. We then ascertain whether there is internal structure within this set of ROIs, as determined by a clustering algorithm on the basis of their response profiles. We show that three clusters emerge from the ROIs, one covering early visual areas, a second including mid-level regions, and a third containing language (semantics and phonological) areas, indicating organizational structure within the distributed circuit. We characterize the profile of each ROI (or node in the network) and assess pairwise connectivity between the ROIs as well as between the clusters. Last, for a more holistic view of the entire circuit, we measure the node strength of each ROI and each cluster, to provide a broad-scale characterization of text-selective cortex bilaterally. Additionally, throughout, we evaluate the modifiability of the nodes, their connections, and the circuit as a whole by contrasting activation profile to words, nonwords, and inverted words (see Figure 1).

Our analyses provide evidence for the role of the VWFA as a hub and a mediator between the early visual and language clusters. We also demonstrate that the VWFA is well suited to synthesize bottom-up visuospatial input features from occipital cortex with top-down conceptual representations involved in language processing. Thus, whereas previous research has identified the VWFA as being important for visual text processing, our results provide graph-theoretic evidence for a possible theory of why the VWFA is important and how it functions within a

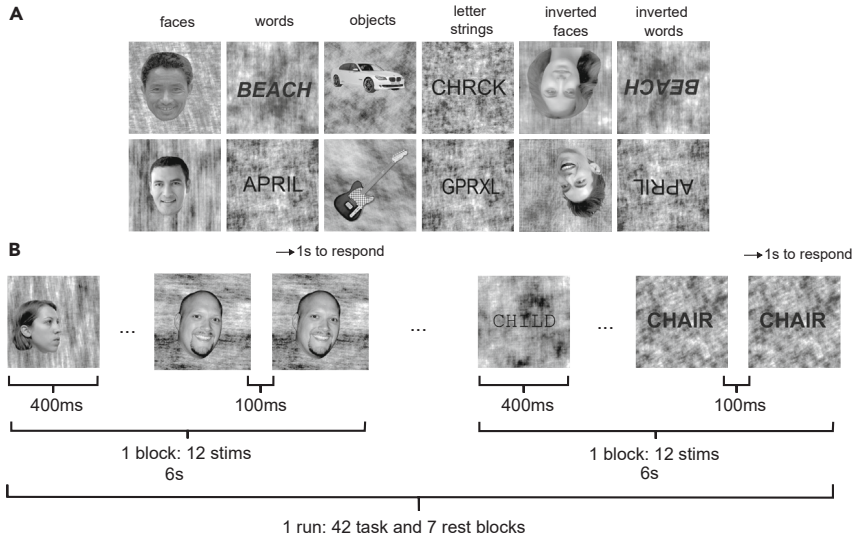


Figure 1. Experimental design

(A) Example stimuli from each of the six categories.

(B) A schematic of one run of the experiment. Note that the repeat condition could be anywhere within the block, not only at the end.

distributed, bilateral and hierarchical network of regions. Moreover, although the response of each individual region in the visual word processing network is flexibly altered by the textual stimulus type, stability is observed at the level of widespread network connectivity. This illuminates the complex dynamics of the network as a whole, in which region- and cluster-level activity fluctuations in response to varying textual stimuli occur in the context of global network-level functional homeostasis.

RESULTS

Regions of interest and network clusters involved in visual text processing

Subjects performed a 1-back stimulus identity task with six stimulus categories: words, inverted words, consonant strings, faces, inverted faces, and objects. A schematic of the experiment and sample stimuli can be seen in Figure 1. We identified the ROIs for our analyses based on whole-brain group selectivity contrasts for text (words, inverted words, and consonant strings) against non-text (faces, inverted faces, and objects) conditions. Hence, we delineated ten ROIs in each hemisphere encompassing the VWFA, Inferior Frontal Gyrus (IFG), and the Peri-Sylvian Language (PSL) regions (Figure 2). We also included the IFG orbitofrontal region (IFGorb), the Inferior Parietal Sulcus (IPS), and visual areas V1, V2, and V3, as well as the Superior Temporal Sulcus and Gyrus (abbreviated as STS+G) and the Precentral Gyrus (PCG), as referenced by[70] (see Table 1 for the anatomical search spaces used to define each ROI).

At the outset, we examined whether there was any inherent organization among these ROIs in which subsets clustered naturally based on similarity of response profile. The results of the clustering analysis (as described in the STAR Methods) are summarized in Figure 3. The "Early vision" ROIs (V1, V2, V3) split from all other ROIs, which then split into two clusters, referred to here as the "High-level vision" and "Language" clusters. The Early vision cluster split into V1 and V2/V3, followed by the latter cluster splitting into separate left and right hemispheric clusters. The High-level vision cluster split into sets of homologous ROIs (e.g., left and right VWFA and left and right IPS). Last, the Language cluster (STS+G, PSL, IFG, IFGorb, PCG) was strongly divided into separate clusters of LH and RH ROIs, with one exception, which is that RH STS+G fell in the otherwise LH-dominant language cluster.

We take as our structure the clustering at the *highest* level (see blue lines), with three clusters—Early vision, High-level vision, and Language—probed separately in the left and right hemispheres. Given this organizational substructure, in the following sections, we report the statistical analyses at the single ROI level, pairwise ROI level using FC (both within and between hemispheres), and at the level of the three-cluster parcellation in each hemisphere (see Methods for details).

Selectivity as a function of ROI, cluster, hemisphere, and stimulus type

First, we assessed the selectivity for textual stimuli (words, inverted words, and consonant strings) in each of the anatomical ROIs separately for each hemisphere. To do so, we computed the localized selectivity of each ROI for each stimulus type against the other non-text conditions (see Methods for details). The significance was determined using one-sample t-tests against the null distribution across subjects (Figure 4A).

We then conducted a repeated measures ANOVA with the localized selectivity as the dependent measure and ROI, hemisphere, and stimulus type as within-subject factors. As evident in Figure 4A, there was a significant three-way interaction of ROI × Hemisphere × Stimulus-type ($F(18, 486) = 11.32$, $p < 0.001$, $\eta^2 = 0.30$). The decomposition of the three-way interaction revealed many pairwise differences both within and

Table 1. Anatomical search spaces for defining each ROI

ROI	Glasser parcels
VWFA_a	Area PH (PH) and Area TE2p
Inferior frontal gyrus (IFG)	Area 44, Area IFJa, Area 6r, and Area IFSp
PeriSylvian language area (PSL)	PeriSylvian language area
Precentral gyrus (PCG)	Area 55b, posterior eye fields (PEF)
Superior temporal sulcus and gyrus (STS+G)	Areas STSdp, TPOJ1, A5, STV
Intraparietal sulcus (IPS)	Area IP0, Area IP1, Area IP2, and Area IPS1
IFG orbito-frontal region (IFGorb)	Area IFSa and Area 45
Primary visual cortex (V1)	Area V1
Secondary visual cortex (V2)	Area V2
Third visual cortex (V3)	Area V3

Each ROI is composed of one or more anatomical parcels from the Glasser atlas.[28] Areas V1–V3 were not aggregated but still constitute ROIs.

between hemispheres. The set of all pairs of ROIs that differed significantly from each other is listed in Appendix A (Table S1). In Figure 4A, we include horizontal black lines only between the top 30% (a large enough number to provide a representative sample of the results while making the figure details still visible) of all significant pairwise differences across ROIs and stimulus types within each hemisphere, as well as all significant comparisons between hemispheres. As shown by the vertical black lines joining the two hemispheres in the figure, in general, there was greater selectivity for at least one text type in the LH than in the RH for all ROIs including V1–V3, with the exception of IPS and PSL. The PSL was the only ROI in the LH that did not exhibit significant selectivity to any of the three stimulus types. Across the other ROIs in the Language cluster, the LH selectivity advantage held to a greater extent for words followed by inverted words and then by consonant strings. In contrast, the Early vision and High-level vision ROIs exhibited a stronger LH selectivity advantage for consonant strings and inverted words. In the LH, the VWFA (particularly for inverted words and consonant strings) and STS+G (particularly for words and inverted words) exhibited the greatest selectivity (greater than that of the early visual ROIs—V1, V2, and V3). In the RH, in addition to V1, V2, and V3 (which were similarly selective to words, inverted words, and consonant strings), the STS+G was also significantly activated (particularly in response to words), with the highest average selectivity across the three stimulus types compared with all other ROIs in the RH.

As can be inferred from the three-way interaction, there were significant two-way interactions of ROI × Hemisphere ($F(9, 243) = 13.14$, $p < 0.001$, $\eta^2 = 0.33$), reflecting greater localized selectivity for many, albeit not all, ROIs in the LH compared with RH, and ROI × Stimulus ($F(18, 486) = 16.75$, $p < 0.001$, $\eta^2 = 0.38$), indicating greater localized selectivity for words and inverted words compared with consonant strings in STS+G in both the LH and RH. The Hemisphere × Stimulus interaction was not significant ($F(2, 54) = 1.08$, $p = 0.35$, $\eta^2 = 0.04$)). The main effects of ROI ($F(9,243) = 13.96$, $p < 0.001$, $\eta^2 = 0.34$) and Hemisphere ($F(1, 27) = 78.46$, $p < 0.001$, $\eta^2 = 0.74$) reached statistical significance, whereas that of Stimulus ($F(2, 54) = 2.91$, $p = 0.06$, $\eta^2 = 0.1$) trended toward significance.

We also sought to study selectivity at the level of the cluster. This provides a higher level view of text activation. As with the ROI analysis mentioned earlier, we began by computing the functionally localized selectivity of each of the clusters in each hemisphere for all three textual stimulus types (words, inverted words, and consonant strings) against the non-text conditions, followed by one-sample t-tests against the null distribution across subjects (Figure 4B). In the LH, all three clusters (Early vision, High-level vision, and Language) exhibited significant selectivity for all three stimulus types. In the RH, the Early vision cluster was significantly selective for words and consonant strings, the High-level vision cluster was selective for inverted words, and the Language cluster was selective for words and inverted words. The significant selectivity for the clusters, for example, in High-level vision, but not in the individual ROIs making up this cluster, suggests that there is activation that is cumulative but is not discernible at the level of the ROI.

To measure differences in selectivity between the clusters and the corresponding effect of stimulus type and hemisphere, we repeated the same ANOVA as earlier but with selectivity at the level of the cluster as the dependent variable. As shown in Figure 4B, there was a significant three-way interaction ($F(4, 108) = 20.68$, $p < 0.001$, $\eta^2 = 0.43$). Post-hoc testing revealed greater LH than RH cluster activation for all three stimulus types for all three clusters (no lines on figure as all are significant). The black horizontal lines in Figure 4B display all within-hemisphere pairwise comparisons whose difference surpassed the HSD threshold (the full set of all pairs that differed significantly from each other is listed in Appendix A, Table S2). In the LH, all three clusters exhibited significant modulation of their localized selectivity by stimulus type. In particular, the Early vision cluster displayed the highest selectivity for consonant strings, the High-level vision cluster had similar selectivity for inverted words and consonant strings (both of which were greater than the selectivity for words), and the Language cluster had substantially stronger selectivity for words and inverted words compared with consonant strings. Modulation of cluster-level selectivity by stimulus type was less prominent in the Early vision and Language clusters in the RH, even though all three clusters were significantly selective for at least one text condition. There were also several cross-cluster differences as a function of stimulus type in both the LH and RH.

Along with the three-way interaction, there was a significant two-way interaction of Cluster × Hemisphere ($F(2, 54) = 3.62$, $p = 0.03$, $\eta^2 = 0.12$), reflecting the higher localized selectivity for clusters in the LH than RH, especially for the High-level vision and Language clusters. There was also an interaction of Cluster × Stimulus ($F(4,108) = 20.80$, $p < 0.001$, $\eta^2 = 0.44$), with the High-level vision cluster exhibiting greater

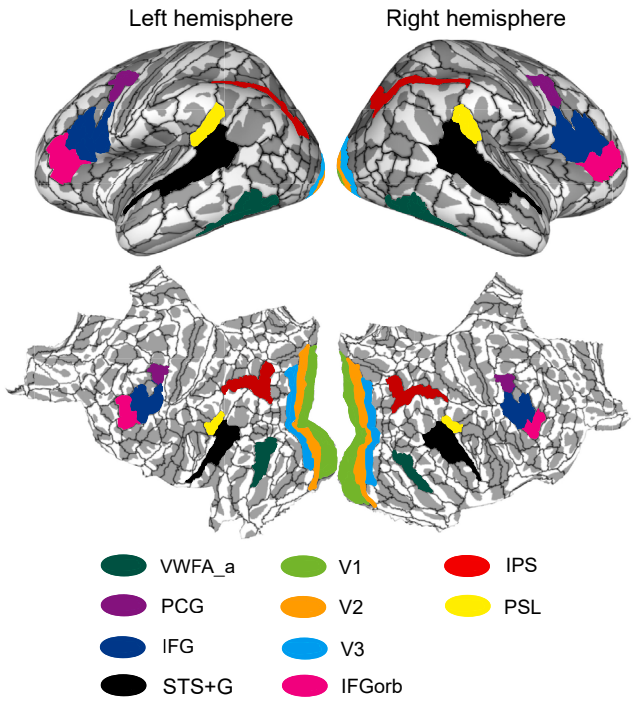


Figure 2. Defining the ROIs involved in visual text processing

ROI-level parcellation using sets of Glasser parcels, viewed on lateral and flat map views of the cortical surface of both hemispheres. See table in Methods for abbreviations.

selectivity for inverted words and consonant strings compared with words in both hemispheres. The Language cluster, on the other hand, was bilaterally more selective for words and inverted words compared with consonant strings (likely driven primarily by the selectivity of the STS+G). Surprisingly, the interaction of Hemisphere × Stimulus was not significant ($F(2,54) = 1.08$, $p = 0.35$, $\eta^2 = 0.04$), reflecting equivalent bilateral sensitivity to stimulus type. Last, there was a main effect of Hemisphere (LH > RH: $F(1, 27) = 78.46$, $p < 0.001$, $\eta^2 = 0.74$) and an effect of Stimulus ($F(2, 54) = 2.91$, $p = 0.06$, $\eta^2 = 0.09$).

In summary, the univariate selectivity analysis at both the ROI and cluster level brings out clear differences between the hemispheres. Generally higher localized selectivity was revealed as well as increased stimulus-based modulation in the LH over the RH (albeit with words and consonant strings differing to a greater extent than words and inverted words in most Language ROIs). At the ROI level, the LH VWFA and STS+G had the highest average selectivity among all other ROIs in the LH, whereas the PSL had the lowest average selectivity. In the RH, the STS+G was the most selective ROI, particularly for the word stimulus, and was closely followed by V1, V2, and V3 (which were relatively similarly selective to all three stimulus types). Notably, although V1–V3 demonstrated text selectivity in the RH, LH selectivity was typically significantly stronger, demonstrating that the LH preference for texts extends even into early visual areas. Selectivity at the cluster level was also significantly higher in the LH than RH and was modulated by stimulus type. Particularly in the LH, there was greater selectivity in the High-level vision cluster for inverted words compared with words (intermediate selectivity for consonant strings). The Language cluster, on the other hand, had similar selectivity for words and inverted words, both of which elicited significantly higher selectivity than consonant strings. Despite this hemispheric advantage, the grouping of ROIs into clusters revealed significant selectivity for at least one text condition in each of the three RH clusters. For language, this appears to be driven predominantly by the RH STS+G, which demonstrated significant word selectivity and trending significance for inverted words. For High-level vision in the RH, the trending selectivity for inverted words in IPS and VWFA became significant at the level of the cluster, implying a weak but significant involvement of the RH within High-level selective visual processing of text.

Multivariate information as a function of ROI, cluster, hemisphere, and stimulus type

To evaluate the response profiles of the ROIs and clusters (in both hemispheres) at a finer grain, and irrespective of whether there is selectivity for any of the text conditions over other categories, we attempted to decode, pairwise, the text condition from the multivariate pattern in each ROI and then in each cluster (i.e., decoding accuracy for words against consonant strings, inverted words against consonant strings, and words against inverted words) (Figure 5). We also derived the decoding accuracy (at the level of the ROI and cluster) for words, inverted words, and consonant strings, against the objects condition (see Appendix B, Figure S1).

A repeated measures ANOVA with decoding accuracy as the dependent measure and ROI, hemisphere, and textual stimulus pair (henceforth referred to as stimulus pair) as within-subject factors did not reveal a statistically significant three-way interaction

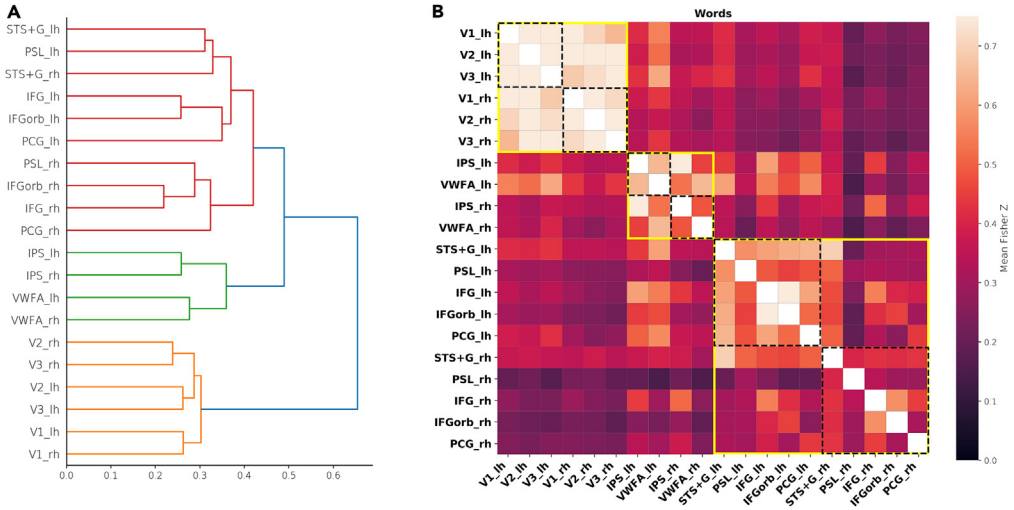


Figure 3. Constructing clusters from patterns of functional connectivity

(A) Dendrogram showing agglomerative hierarchical clustering result using the Paris algorithm. The three colored clusters correspond to the three clusters analyzed throughout the paper (orange: Early vision, green: High-level vision, red: Language).

(B) Pairwise functional connectivity matrix. Each cell plots the mean Fisher *Z* score across subjects for functional connectivity (correlation of mean parcel time-series) between the corresponding ROI pair during viewing of Words (see Methods for further details). Yellow boxes indicate the clusters revealed through the clustering algorithm [see (A), and text for further details]. Dashed black boxes show each hemisphere within each of the clusters. See table in Methods for abbreviations.

($F(18, 486) = 0.9$, $p = 0.51$, $\eta^2 = 0.03$). However, all two-way interactions were significant: ROI × Hemisphere ($F(9, 243) = 4.9$, $p < 0.001$, $\eta^2 = 0.16$); ROI × Stimulus-pair ($F(18, 486) = 6.2$, $p < 0.001$, $\eta^2 = 0.19$) and Hemisphere × Stimulus-pair ($F(2, 54) = 15.8$, $p < 0.001$, $\eta^2 = 0.37$). Two of the three main effects were also significant: ROI ($F(9,243) = 6.48$, $p < 0.001$, $\eta^2 = 0.19$) and Hemisphere ($F(1, 27) = 11.8$, $p < 0.005$, $\eta^2 = 0.3$). The main effect of Stimulus-pair was not significant ($F(2, 54) = 2.2$, $p > 0.10$, $\eta^2 = 0.07$).

All ROIs that exhibited a significantly greater decoding accuracy than chance (0.5) for specific stimulus-pairs, as well as all pairwise comparisons within each of the hemispheres that survived post-hoc Tukey correction, are shown in Figure 5A. In the LH, all ROIs except the IPS were able to decode at least one pair of stimulus types with an accuracy significantly above chance. All Early vision ROIs could decode words from inverted words, whereas all Language ROIs decoded words and inverted words from consonant strings, but most could not distinguish between words and inverted words. The LH VWFA exhibited significant decoding accuracy for all three stimulus-pairs.

In contrast, in the RH, V1, V2, V3, VWFA, and STS+G were the only ROIs whose decoding accuracy was significantly greater than chance (for at least one stimulus-pair). Similar to the LH, the voxel activation pattern in V1, V2, V3, and VWFA, but not STS+G, could decode Words from Inverted Words.

To assess decoding accuracy at the level of the cluster, we repeated the same ANOVA as earlier but with Cluster as a factor rather than ROI. There was neither a statistically significant three-way interaction ($F(4, 108) = 0.4$, $p = 0.8$, $\eta^2 = 0.01$) nor a significant two-way interaction of Stimulus-pair × Hemisphere ($F(2, 54) = 2.2$, $p = 0.1$, $\eta^2 = 0.07$). However, the two-way interactions of Cluster × Hemisphere ($F(2, 54) = 8.0$, $p < 0.001$, $\eta^2 = 0.23$) and Cluster × Stimulus-pair ($F(4, 108) = 13.0$, $p < 0.001$, $\eta^2 = 0.33$) were significant. There was also a significant main effect of Hemisphere ($F(1, 27) = 12.4$, $p < 0.005$, $\eta^2 = 0.32$) but not of Stimulus-pair ($F(2, 54) = 0.3$, $p > 0.50$, $\eta^2 = 0.01$) or Cluster ($F(2, 54) = 0.6$, $p > 0.50$, $\eta^2 = 0.02$).

Figure 5B represents the specific stimulus-pairs for which each of the clusters exhibited a significantly high decoding accuracy (i.e., significantly greater than chance) and also pairwise comparisons within each hemisphere that survived post-hoc Tukey correction. As evident, the Early vision and High-level vision clusters in both the LH and RH could decode words from inverted words. This is in contrast to the Language cluster in both hemispheres, which exhibited significant decoding accuracy for words and inverted words against consonant strings but failed to decode words from inverted words. None of the between-hemisphere pairwise comparisons (across homologous cluster-stimulus combinations) survived post-hoc correction, suggesting a very similar pattern of decoding accuracy across the three clusters in both hemispheres (for all three stimulus types), in strong contrast to the pattern found for selectivity.

In summary, evaluating decoding accuracy provided information over and above that gleaned from the univariate analysis. Specifically, we observed similar overall decoding accuracy in the LH and RH (both for decoding textual stimuli from each other and textual stimuli against objects), even though selectivity of the RH was significantly lower than that of the LH for ROIs as well as clusters (Figures 4A and 4B). As evident in Figure 5, in both hemispheres, the accuracy of decoding words from inverted words decreased from the posterior (primarily ROIs in the Early vision cluster) to more anterior regions (ROIs in the Language cluster). The only notable exception to this trend was the VWFA, which, despite being more anterior to the ROIs of the Early vision cluster, exhibited similar, if not greater, decoding accuracy for words against inverted words

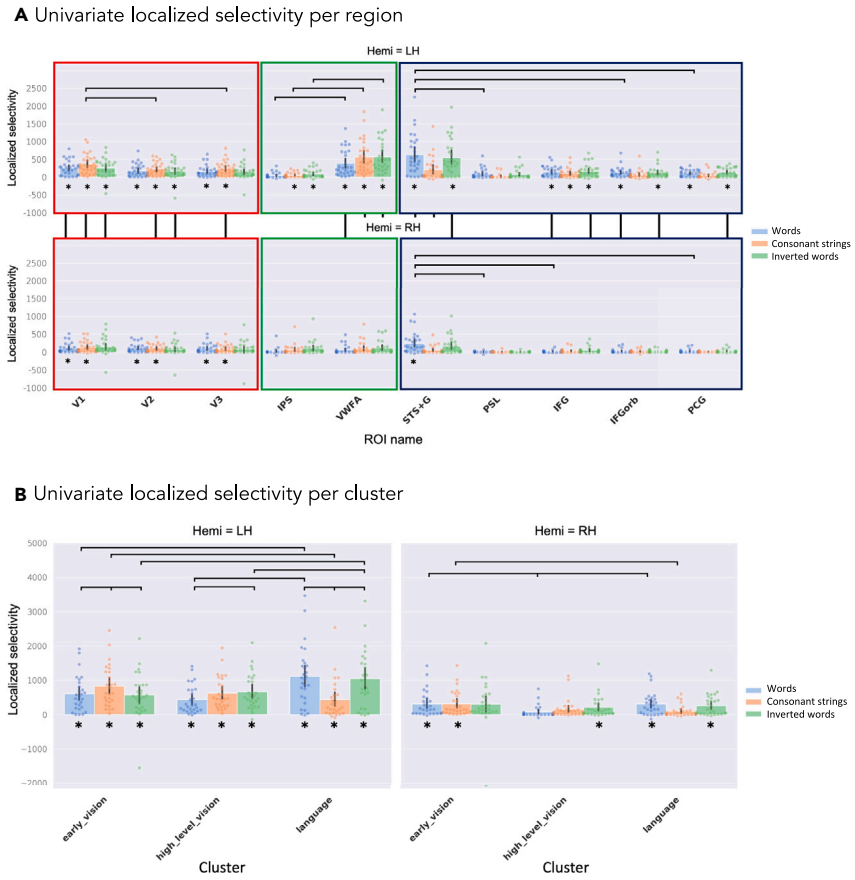


Figure 4. Effects of stimulus type and hemisphere on univariate localized selectivity across ROIs and clusters

(A) Localized selectivity of each ROI for each of the three stimulus types in each hemisphere (for convenience, we group the ROIs according to their cluster membership so this can be appreciated as well). Horizontal black lines indicate the top 30% of all significant pairwise differences across ROIs and stimulus types in each hemisphere. Vertical black lines indicate all significant between-hemisphere pairwise differences across homologous ROI-stimulus combinations that survived correction.

(B) Localized selectivity per cluster for each stimulus condition. Horizontal black lines indicate all significant within-hemisphere pairwise differences (across clusters and stimulus types). All homologous (i.e., for the same cluster-stimulus combination) between-hemisphere comparisons are significant (see text). In both (A) and (B), asterisks indicate ROI/cluster × stimulus combinations in the LH and RH that have significantly greater selectivity than the null (i.e., 0) (computed using one-sample t-tests and corrected for multiple comparisons). Dots show scores for individual subjects, bars show the average across subjects, and black error lines indicate the 95% confidence interval of scores across subjects. Significance of all pairwise differences was computed with Tukey's Honest Significant Difference (HSD) test and $p < 0.01$. See table in Methods for abbreviations.

compared with early visual regions (along with a similarly significant decoding accuracy for the other two stimulus-pairs). A similar posterior-to-anterior transition was seen at the cluster-level, such that the Early vision cluster had the highest decoding accuracy for words against inverted words, but the lowest decoding accuracy for words against consonant strings, in both hemispheres. This pattern was reversed in the Language cluster, which displayed significantly higher decoding accuracy for words versus consonant strings compared with words versus inverted words. This was likely primarily driven by the STS+G, which was the only ROI in the Language cluster to display significant differences in decoding accuracy across stimulus-pairs. Combined with the selectivity analyses, the strong multivariate information for distinguishing words and inverted words from consonant strings in the RH STS+G provides converging evidence for a role of the RH in visual word processing.

These first univariate and multivariate analyses elucidate the selectivity and stimulus-based modulation of fMRI responses for all ROIs and the clusters they constitute. Having identified the nodes, we now examine the functional connectivity between these nodes (ROIs and clusters) to map out the extent of their associations.

Functional connectivity of the visual word processing network

These second analyses probe the associations between the nodes. Here, we examine the interrelationships among ROIs and among clusters by documenting their functional connectivity profiles and the modulation of these interrelationships by stimulus type.

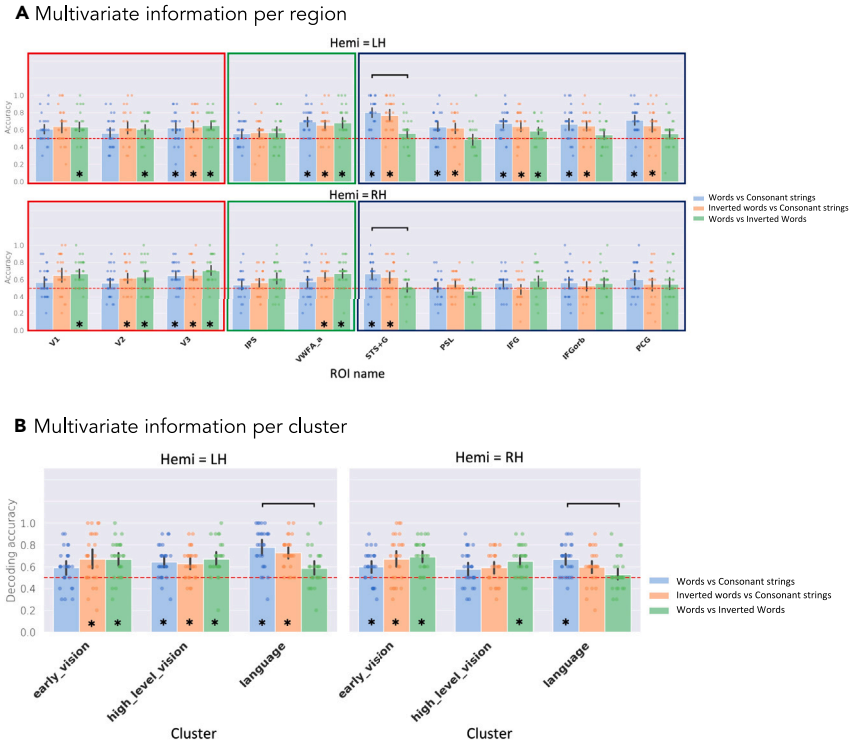


Figure 5. Multivariate decoding between the three text stimuli in the ROIs and clusters in both hemispheres

(A) Multivariate decoding accuracy for each pair of text conditions (i.e., words versus consonant strings, inverted words versus consonant strings, and words versus inverted words) in each ROI of each hemisphere. Black lines indicate all significant within-hemisphere and within-cluster pairwise differences across stimulus types (horizontal lines). No between-ROI (within the same cluster) or between-hemisphere pairwise differences (across homologous ROI-stimulus combinations) survived post-hoc correction.

(B) Multivariate decoding accuracy within each of the three clusters in each hemisphere. Horizontal black lines indicate all significant within-hemisphere pairwise differences (across stimulus types). None of the between-cluster (within the same hemisphere) comparisons or between-hemisphere comparisons survived correction. In each plot, dots show scores for individual subjects, bars show the average across subjects, and black error lines indicate the 95% confidence interval of scores across subjects. Bars with asterisks are significantly greater than chance, computed using a one-sample t-test against 0.5, and corrected for multiple comparisons. Significance of all pairwise differences were computed with Tukey's Honest Significant Difference (HSD) test and $p < 0.01$. See table in Methods for abbreviations.

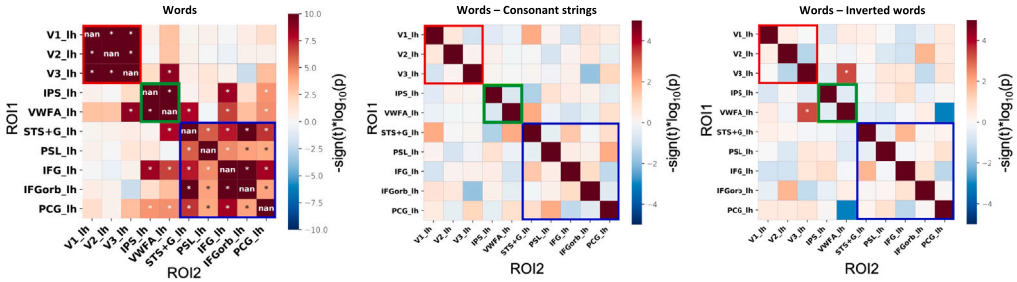
We derived the FC of all ROI pairs within each hemisphere, as well as all homologous ROI pairs across hemispheres. Because the within-LH, within-RH, and between hemisphere comparisons could not be captured in an orthogonal factorial analysis, we performed three separate ANOVAs, one for all ROI pairs within the LH ("*LH-comparisons*"), one for the ROI pairs within the RH ("*RH-comparisons*"), and the last between all homologous (but not non-homologous) ROI pairs in the LH and RH ("*between-hemisphere*"). For the factorial analysis of variance, we refer to the first ROI in a pair as ROI1 and the second as ROI2. As the partial functional correlation is symmetric, the pair index label is arbitrary; thus, where the effect of interest involves only a single ROI as in the cross-hemisphere homotopic regions (e.g., ROI × hemisphere), we do not use the ROI pair index label. Below, we present the heat maps of words against the null distribution and then the partial functional correlation between ROI pairs for words compared with inverted words and consonant strings.

For the *LH-comparisons* analysis, with ROI1, ROI2, and stimulus type as within-subject factors and FC as the dependent measure, we observed a significant three-way interaction ($F(162,4374) = 2.07$, $p < 0.001$, $\eta^2 = 0.07$). The two-way interaction of ROI1 × ROI2 ($F(81,2187) = 720.41$, $p < 0.001$, $\eta^2 = 0.97$) was also significant but not ROI1/ROI2 × Stimulus ($F(18,486) = 0.98$, $p = 0.48$, $\eta^2 = 0.03$). There were also significant main effects of ROI ($F(9,243) = 14.9$, $p < 0.001$, $\eta^2 = 0.36$) but not of Stimulus ($F(2,54) = 1.51$, $p = 0.23$, $\eta^2 = 0.05$). IFG was the ROI with the largest overall partial functional correlation with other regions, followed by STS+G, VWFA, and then IFGorb.

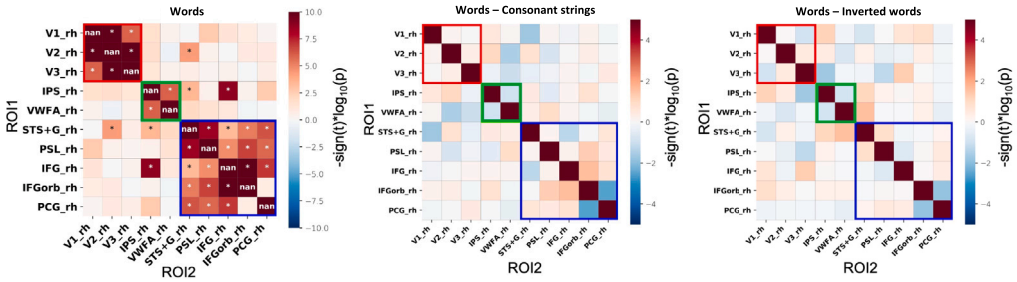
As with the *LH-comparisons* analysis, we conducted the *RH-comparisons* analysis with ROI1, ROI2, and stimulus type as within-subject factors and FC as the dependent measure. We observed a significant three-way interaction ($F(162,4374) = 1.26$, $p = 0.01$, $\eta^2 = 0.04$). There was also a two-way interaction of ROI1 × ROI2 ($F(81,2187) = 751.3$, $p < 0.001$, $\eta^2 = 0.97$) but not ROI × Stimulus ($F(18,486) = 1.06$, $p = 0.38$, $\eta^2 = 0.04$). There was also a significant main effect of ROI ($F(9,243) = 13.9$, $p < 0.001$, $\eta^2 = 0.34$) but not of Stimulus ($F(2,54) = 0.51$, $p = 0.60$, $\eta^2 = 0.02$). IFG was the region most strongly connected to others followed by V2 and then STS+G.

To identify the ROI pairs that were significantly functionally correlated in the words stimulus condition, we computed one-sample post-hoc t-tests of the partial functional correlation of all ROI pairs in the LH and RH for words against a null distribution (leftmost panel; Figures 6A and

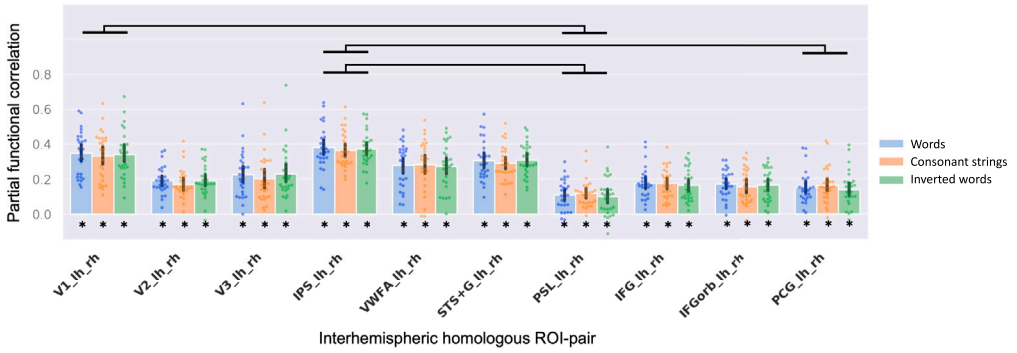
A Left hemisphere



B Right hemisphere



C Interhemispheric FC for homologous ROI-pairs



D FC per cluster

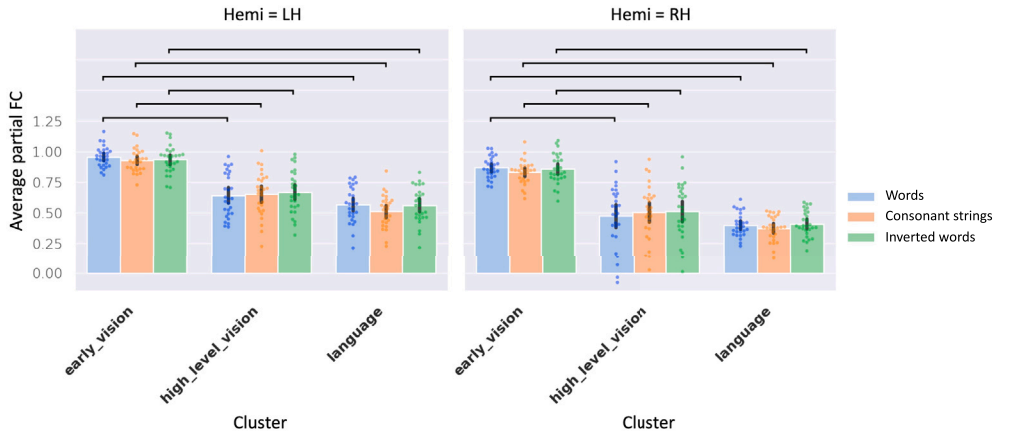


Figure 6. Effect of stimulus type on intra- and inter-hemispheric functional connectivity between ROIs and within clusters

(A) Functional correlation (connectivity) of all ROI pairs in the left hemisphere for the three stimulus conditions.

(B) Functional connectivity of all ROI pairs in the right hemisphere for the three stimulus conditions. The leftmost panels in (A) and (B) are one-sample t-tests of the partial functional correlation between ROI pairs for Words against the null distribution. The middle and rightmost panels in (A) and (B) are two-tailed t-tests of the partial functional correlation between ROI pairs for Words compared with Consonant Strings and Inverted Words, respectively. Solid red, green, and blue boxes surround ROI pairs in the Early vision, High-level vision, and Language clusters, respectively. Cells with asterisks represent partial functional correlations that survived Bonferroni correction ($p < 0.0005$).

(C) Partial FC of each of the interhemispheric homologous ROI-pairs for each of the three stimulus types. Horizontal black lines indicate all ROI pairs that have significantly different interhemispheric partial FC, regardless of stimulus type. Bars with asterisks are significantly greater than 0, computed using a one-sample t-test against the null distribution (i.e., 0) and corrected for multiple comparisons.

(D) Average within-cluster partial FC for each of the three stimulus types in both hemispheres. Horizontal black lines indicate the top 70% of all significant within-hemisphere pairwise differences (across clusters and stimulus types). None of the between-hemisphere comparisons survived post-hoc correction. In both (C) and (D), dots show scores for individual subjects, bars show the average across subjects, and black error lines indicate the 95% confidence interval of scores across subjects. Significance of all pairwise differences were computed with Tukey's Honest Significant Difference (HSD) test and $p < 0.01$. See table in Methods for abbreviations.

6B). To further illustrate the findings of the above-described ANOVAs and highlight the key interrelationships (between ROI pairs and across stimulus types), we also computed two-sample t-tests comparing the partial functional correlation of ROI pairs for words with that for consonant strings and inverted words (Figures 6A and 6B; middle panel for consonant strings, rightmost panel for inverted words). Asterisks denote ROI pairs that survived post-hoc Bonferroni correction ($p < 0.0005$). As evident from the heatmaps, there is considerable similarity across stimulus types, with significant connectivity between V1, V2, and V3, between the IPS and VWFA, and also between STS+G, PSL, IFG, IFGorb, and PCG, albeit all to a slightly stronger degree in LH than RH. Although it is not surprising to see this connectivity pattern for words (as the clustering in Figure 3 was derived from the pairwise correlations of activations in ROIs in response to words), highly similar pairwise correlations are seen across all stimulus types. Notably, the only significant stimulus difference that was seen was in the FC between VWFA and V3 comparing words and inverted words, with this connection being weaker under word inversion. Given the non-independence of the clusters and the FC based on words as inputs, we repeated the clustering based just on the inverted word condition and then just based on the consonant string condition (see Appendix C), and the clustering into the three subnetworks is identical for all three stimulus conditions. Hence, the FC reported here for words is not unduly skewed by the specific clustering performed on the word condition, as the same outcome holds for all the three stimulus types.

The *between-hemisphere* analysis revealed no interaction between ROI pair (i.e., pairs of homologous ROIs, for example, the FC value for LH-RH VWFA, labeled "VWFA pair") and stimulus type ($F(18,486) = 1.3$, $p > 0.1$, $\eta^2 = 0.05$), nor a main effect of stimulus ($F(2,54) = 1.8$, $p > 0.1$, $\eta^2 = 0.06$). There was, however, an effect of ROI pair ($F(9,243) = 22.7$, $p < 0.001$, $\eta^2 = 0.46$). A one-sample t-test revealed that all ten between-hemisphere ROI pairs had significant partial functional correlation compared with the null distribution for all three stimulus types and corrected for multiple comparisons ($p < 0.001$). The IPS emerged as having the strongest interhemispheric partial FC, followed by V1 and STS+G, based on a post-hoc Tukey HSD test (Figure 6C). These results highlight the existence of significant interhemispheric communication in ROIs of each of the three network clusters, regardless of stimulus type.

We also determined whether certain clusters showed higher FC than others and whether this differed as a function of stimulus type. Note that we used FC initially to calculate the clustering based just on the words condition (see Figure 3) so the analysis of the activation for words is not independent from the definition of clusters; however, we revealed the exact same three-cluster division when the aggregation was done with inverted words and consonant strings, suggesting that there is not a strong bias in the clustering that might affect the FC outcome. In a three-way ANOVA with cluster, hemisphere, and stimulus type as within-subject factors, there was neither a significant three-way interaction ($F(4, 108) = 1.1$, $p > 0.3$, $\eta^2 = 0.04$) nor two-way interactions of either Hemisphere × Cluster ($F(2, 54) = 2.6$, $p > 0.05$, $\eta^2 = 0.09$) or Hemisphere × Stimulus ($F(2, 54) = 1.3$, $p > 0.2$, $\eta^2 = 0.05$). The only significant interaction was between Cluster × Stimulus ($F(4, 108) = 5.1$, $p < 0.001$, $\eta^2 = 0.16$). For all three stimulus types, there was greater FC within the Early vision, then High-level vision, and last, Language clusters. All of these comparisons survived correction using Tukey's HSD test and $p < 0.01$ (Figure 6D; for the full list of all significant cluster functional connectivity comparisons that survived the post-hoc Tukey HSD correction, see Appendix C, Table S4). None of the clusters individually exhibited significant difference in FC between words, consonant strings, and inverted words. Although the main effect of Stimulus was not significant ($F(2, 54) = 2.1$, $p > 0.10$, $= 0.07$), there was a main effect of Hemisphere ($F(1, 27) = 43.1$, $p < 0.001$, $\eta^2 = 0.62$), reflecting the greater FC in the LH than RH, and of Cluster ($F(2, 54) = 102.8$, $p < 0.001$, $\eta^2 = 0.79$), with significantly greater FC in Early vision, followed by High-level vision, and with lowest FC in the Language cluster.

In summary, FC was highly modulated by hemisphere and ROI pair but less by stimulus type. Pairwise connectivity was strong between V1, V2, and V3, between IPS and VWFA, and between the remaining language regions, although almost always to a greater degree in the LH than RH. The LH VWFA exhibited significant FC with the largest number of ROIs, spanning across all three clusters. Although some ROI pairs (both within and between hemispheres) exhibited a small difference in partial functional correlation for inverted words and consonant strings compared with words, few comparisons survived post-hoc correction. All ROIs exhibited significant interhemispheric connectivity to their contralateral homologue, for all three stimulus types. The connectivity at the cluster level demonstrated slightly greater LH than RH connectivity and stronger connectivity of the Early vision, then High-level vision clusters followed by the Language cluster for all three stimulus types. Although this reliable pattern of FC across stimulus types (at both the level of the ROI-pair and cluster) may seem surprising, we confirmed this pattern using only the peak of the hemodynamic response function (HRF) for the middle of each stimulus block (see Appendix C, Figure S3)

and using the beta coefficients associated with each stimulus block obtained from the GLM, across all subjects (see Appendix C, Figure S4). Hence, it appears that although the activity of individual ROIs (in their univariate selectivity and multivariate decoding accuracy) is modulated by stimulus type, the interactions between pairs of ROIs, as well as average FC of the three clusters of ROIs remains largely stable across words, inverted words, and consonant strings.

Node strength across stimulus types and hemispheric differences

Following the comparisons of individual ROIs/clusters against each other and then pairs of ROIs/clusters against each other, in this final set of analyses, we stepped back to assess the general organization of the bilateral visual word processing circuit. To do so, we adopted a graph-theoretic approach and calculated node strength of each ROI and then, separately, node-cluster strength of each ROI (defined below).

For a given ROI, node strength was computed as the total sum of the absolute partial functional correlation between that ROI and all other ipsilateral ROIs as well as its contralateral homologue (i.e., the sum of the absolute weights of all edges attached to that ROI, other than the edges between the ROI and non-homologous contralateral ROIs). The contralateral homologue was included, as our FC analyses revealed significant interhemispheric homologous functional connectivity for all ROIs (Figure 6C). Moreover, previous research has provided evidence of significant structural connectivity between contralateral homologous ROIs through the corpus callosum. Connectivity between non-homologous ROIs across hemispheres is less clear.

The node-cluster strength for a given ROI and a given cluster was calculated as the sum of the absolute partial functional correlation between the ROI and all other ipsilateral ROIs within the cluster (in this case, summation was restricted to only the *ipsilateral* edges between the given ROI and cluster).

A repeated measures ANOVA with ROI, hemisphere, and stimulus type as within-subject factors and node strength as the dependent measure revealed neither a significant three-way interaction ($F(18, 486) = 0.71$, $p = 0.80$, $\eta^2 = 0.03$) nor a significant two-way interaction of Stimulus × Hemisphere ($F(2, 54) = 2.05$, $p = 0.14$, $\eta^2 = 0.07$) or ROI × Stimulus ($F(18, 486) = 1.30$, $p = 0.18$, $\eta^2 = 0.05$). However, the two-way interaction of ROI × Hemisphere ($F(9, 243) = 8.59$, $p < 0.001$, $\eta^2 = 0.24$) was significant. Figure 7A shows the node strength of individual ROIs as a function of stimulus type and hemisphere. Post-hoc Tukey HSD testing revealed a host of significant differences in node strength as a function of ROI and hemisphere. As shown by the vertical black lines between the LH and RH graphs, node strength for most ROI-stimulus pairs was higher in the LH than RH. The horizontal black lines show the top 10% of all within-hemisphere pairwise differences. In the LH, the PSL had the lowest average node strength across all three stimulus types; thus, the strongest differences in node strength were between the PSL and most other ROIs (particularly V3, VWFA, STS+G, and IFG), which were similar to each other. The STS+G exhibited the highest average node strength across all stimulus types, closely followed by IFG and VWFA. In the RH, IPS and IFG had higher average node strength compared with most other ROIs, but especially PSL and PCG. In contrast to the LH, the RH VWFA had the third lowest average node strength across stimulus types after PSL and PCG. In both hemispheres, none of the ROIs exhibited significant modulation of their node strength by stimulus type. For the full list of all significant node strength comparisons that survived the post-hoc Tukey HSD correction, see Appendix D, Table S5.

All three main effects in the ANOVA were also significant: ROI ($F(9, 243) = 27.5$, $p < 0.001$, $\eta^2 = 0.50$), Stimulus ($F(2, 54) = 5.4$, $p < 0.01$, $\eta^2 = 0.17$), and Hemisphere ($F(1, 27) = 34.6$, $p < 0.001$, $\eta^2 = 0.56$).

In summary, node strength, which measures, for each ROI, the strength of its connection with the broad cluster comprising all other ipsilateral ROIs and its contralateral homologue, is higher in the LH than RH across the board. Within each hemisphere, ROIs differed from each other in terms of the strength of their connection to the rest of the ROIs, but there was no significant modulation of the ROIs themselves by stimulus type. Hence, similar to the pairwise functional connectivity between ROIs, network-level analyses of the node strength of each ROI within the broader set of ROIs exhibits a stability across stimulus types.

Differential node-cluster strength across ROIs, with high-level visual regions mediating between early vision and language

Finally, we assessed the connectivity of each of the ROIs to each of the three clusters—Early vision, High-level vision, and Language—separately. This was computed as the sum of an ROI's absolute partial functional correlation with all other ipsilateral ROIs in a cluster. This sum was then normalized by a factor $n - 1$, where n is the total number of ROIs being considered (the subtracted 1 is to exclude self-connections). This computation is distinct from the average functional connectivity of the three clusters, which we have reported earlier (Figure 6D), as it focuses on the connectivity of nodes to clusters, rather than all connectivity within a cluster.

In Figure 7B, the boxes surrounding groups of individual ROIs are indicative of the cluster to which they belong: the Early vision cluster (red box), the High-level vision cluster (green box), and the Language cluster (blue box). Within each of these boxes, the y-axis reflects the normalized connectivity of the ROIs in a given hemisphere to their "*native*" cluster (i.e., the cluster to which they belong, see Figure 3A) and with the other two "*non-native*" clusters. For example, for the ROI V1, its native cluster would be Early vision, and its non-native clusters would be High-level vision and Language. For each ROI, the blue bar represents its connectivity to the Early vision cluster, the orange bar to the High-level vision cluster, and the green bar to the Language cluster. Figure 7B depicts the node connectivity of each of the individual ROIs only for the words stimulus condition (figures for the consonant strings and inverted words conditions are not significantly different and are included in the Supplemental Material; see Appendix D, Figure S5).

A three-way repeated measures ANOVA with hemisphere, ROI, and cluster as within-subject factors revealed a significant interaction of all three factors ($F(18, 486) = 3.7$, $p < 0.001$, $\eta^2 = 0.12$). Detailed scrutiny of the data revealed the source of the three-way interaction. All ten ROIs (in both hemispheres) exhibited significantly greater normalized connectivity to their native cluster, compared with their non-native clusters. Moreover, an interesting pattern emerged across the ROIs in terms of their relative connectivity to their non-native clusters. The ROIs in

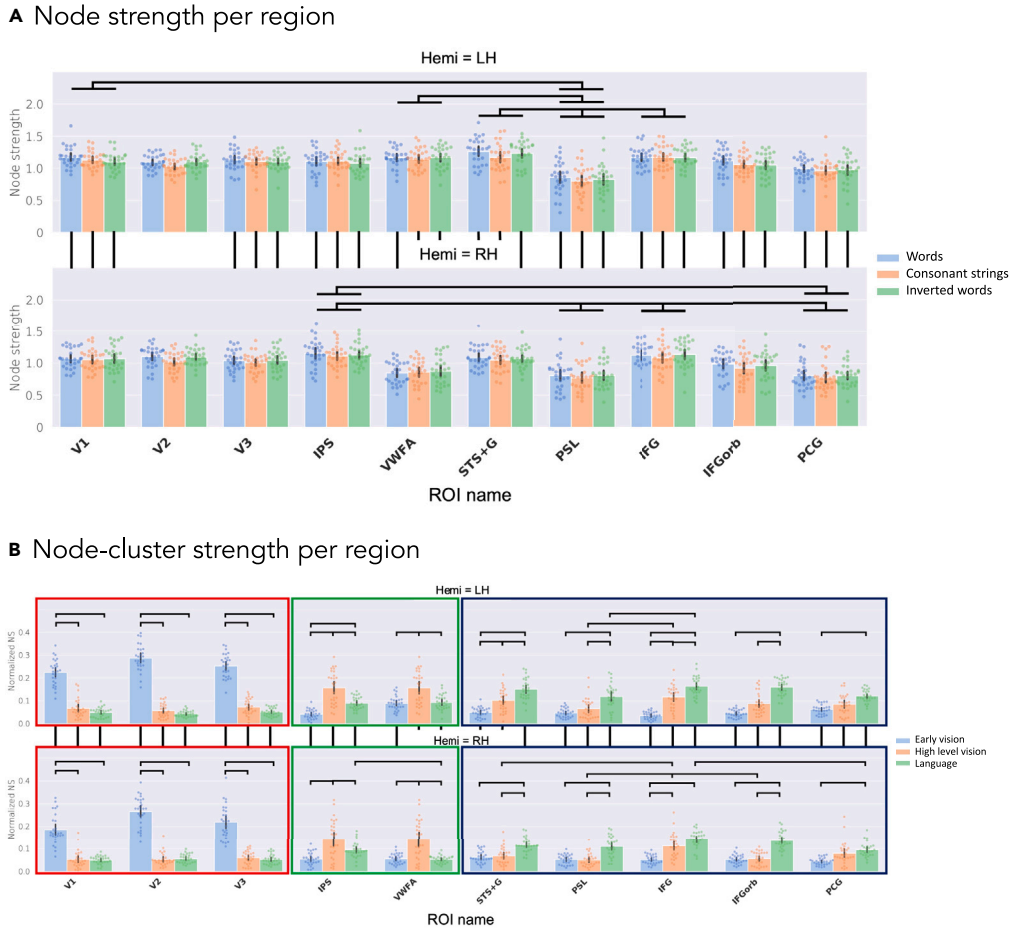


Figure 7. Effect of stimulus type and hemisphere on node strength (derived using partial FC) of each of the ROIs with the broad network of all other ipsilateral ROIs and the homologous contralateral ROI, as well as with each of the three clusters

(A) The node strength of each ROI with the broader network of all other ROIs within the same hemisphere and the homologous ROI in the contralateral hemisphere, for each of the three stimulus types and across hemispheres. Horizontal black lines indicate the top 10% of all significant within-hemisphere pairwise differences, whereas the vertical lines indicate all between-hemisphere pairwise differences across homologous ROI-stimulus combinations.

(B) Normalized connectivity of each ROI to each of the three clusters (Early vision: blue bars, High-level vision: orange bars, and Language: green bars) in the LH (top panel) and the RH (bottom panel) for the Words stimulus condition. Vertical black lines indicate all significant between-hemisphere pairwise differences across homologous ROI-cluster combinations. Horizontal black lines indicate the top 30% of all within- and across-ROI comparisons in each hemisphere for the Early vision ROIs (surrounded by solid red box) as well as for the Language ROIs (surrounded by solid blue box). For the High-level vision ROIs (surrounded by solid green box), the horizontal lines represent the top 60% of all within- and across-ROI comparisons in each hemisphere (we show the most significant subset of comparisons here to avoid overcrowding the figure; a full list of pairwise comparisons can be found in Appendix D). In both (A) and (B), dots show scores for individual subjects, bars show the average across subjects, and black error lines indicate the 95% confidence interval of the normalized node strength values across subjects. Significance of all pairwise differences was computed with Tukey's Honest Significant Difference (HSD) test and $p < 0.01$. See table in Methods for abbreviations.

the Early vision cluster (V1, V2, and V3; surrounded by the red box in Figure 7B) and the Language cluster (STS+G, PSL, IFG, IFGorb and PCG; surrounded by the blue box in Figure 7B), all shared the characteristic of having significantly higher connectivity to the High-level vision cluster than to their other non-native cluster (see Appendix D). This was true in both the LH and the RH. Among the ROIs of the High-level vision cluster (IPS and VWFA; surrounded by the green box in Figure 7B), the IPS exhibited significantly higher connectivity to the Language cluster compared with the Early vision cluster. On the contrary, unlike most other ROIs, the VWFA, in both the LH and the RH, exhibited very similar connectivity to its non-native clusters (Early vision and Language). Furthermore, VWFA had the highest connectivity to the Early vision and Language clusters compared with other ROIs for which either of those two clusters were non-native (this was also true in both hemispheres). Supplementing the findings shown in Figure 7A, most ROIs had significantly greater connectivity to the three clusters in the LH compared with the RH. For the full list of all significant node strength comparisons that survived the post-hoc Tukey HSD correction, see Appendix D, Table S6.

The two-way interactions of ROI × Cluster ($F(18, 486) = 143.9$, $p < 0.001$, $\eta^2 = 0.84$) and ROI × Hemisphere ($F(9, 243) = 3.6$, $p < 0.001$, $\eta^2 = 0.12$) were also significant, but that of Cluster × Hemisphere was not ($F(2, 54) = 1.0$, $p = 0.37$, $\eta^2 = 0.04$). Last, all three main effects were also

significant: ROI ($F(9,243) = 24.0$, $p < 0.001$, $\eta^2 = 0.47$), cluster ($F(2,54) = 19.8$, $p < 0.001$, $\eta^2 = 0.42$), and hemisphere ($F(1, 27) = 49.1$, $p < 0.001$, $\eta^2 = 0.65$).

To summarize, the results of our graph-theoretic analyses suggest that, bilaterally, the three clusters (Early vision, High-level vision, and Language) function as "subnetworks" of a broader visual word processing network, as they show greater within- than between-cluster connectivity, and each of the ROIs shows preferential connectivity with their native cluster (this is true for inverted words and consonant strings as well; see Appendix D, Figure S5). However, although a general preference for connectivity to native clusters is expected, and was generally confirmed, differential weighting of connectivity to the non-native clusters across ROIs emerged as well. All ROIs in the Early vision and Language clusters had higher connectivity to the High-level vision cluster than to their other non-native cluster. In contrast, the VWFA, in both the LH and RH, showed similar connectivity to both of its non-native clusters (Early vision and Language). This suggests that the High-level vision cluster, and in particular the VWFA, may play a role in mediating signal propagation and integration between the Early vision and Language clusters. Our collective univariate, multivariate, and functional connectivity results further reinforce the theory that the VWFA may be the more dominant of the two ROIs in the High-level vision cluster in terms of its mediating function, as discussed in greater detail in the Discussion. Moreover, similar to the generic node strength of each of the ROIs within the broad set of ROIs (Figure 7A), the normalized connectivity of each ROI to the three clusters is also relatively stable and unmodulated across stimulus types.

DISCUSSION

The goal of this study was to characterize, in a largely data-driven fashion, those regions in human cortex that show selective responses to text input. Decades of studies have attested to the key role of the Visual Word Form Area (VWFA) in the LH in the service of visual word processing. This primary focus on the VWFA and its seemingly singular role in text processing has, however, been challenged by investigations showing that text activates regions beyond the VWFA including Broca's and Wernicke's area as well as the intraparietal sulcus, precentral gyrus, and superior temporal and supramarginal gyri.[13,70]

The first aim of the present study was to define, within individual, for both hemispheres, the large-scale distributed circuit for text processing as well as the functional properties and interrelationships between the different regions. Therefore, based on whole-brain fMRI data obtained while participants viewed three different textual stimulus types (words, consonant strings, and inverted words) and non-text (faces, inverted faces and objects) conditions, we examined responses in a set of 10 regions in each hemisphere and at an intermediate level of organization derived from a clustering algorithm that grouped the ROIs into three clusters: Early vision, High-level vision, and Language. For each ROI and cluster in each hemisphere, we quantified the univariate response selectivity profile, the multivariate decoding accuracy, and the pairwise functional connectivity (FC). Last, using graph theory metrics, we measured node strength, with each ROI as a node within the broader network of all 10 ROIs and as a node within each of the three defined clusters.

The analysis of univariate ROI selectivity in the LH revealed significant selectivity in most ROIs for at least one of the three text stimulus types (see Figure 4), with ROIs of the Early vision cluster evincing similar selectivity for each of the text stimulus types and the High-level vision cluster showing more selectivity for consonant strings and inverted words than for words. Of all the ROIs, the VWFA exhibited the highest average selectivity across text stimulus types, replicating the standard profile reported ubiquitously in the literature. Last, most ROIs in the Language cluster exhibited significant selectivity to words and inverted words with reduced selectivity to consonant strings (see Figure 4). These findings were supplemented by multivariate results, revealing that at least one pair of textual stimuli could be decoded in most LH ROIs (see Figure 5; see Appendix B, Figure S1), with all three pairs of textual stimuli decodable from each other in the VWFA. Words and consonant strings were better decoded from inverted words in the Early vision cluster, suggesting representation of visual features (i.e., upright vs. inverted text). The ROIs of the Language cluster had at- or below-chance decoding accuracy for words against inverted words but exhibited very high accuracy for both stimuli against consonant strings, emphasizing the lexical aspects of the input. All of these findings were replicated in the univariate and multivariate analyses when sampling from the three clusters in the LH with, again, similar significant modulation by stimulus type.

For the RH, selectivity and decoding ability were also identified at the ROI and cluster levels in the univariate and multivariate analyses. RH univariate selectivity for text was found consistently only in early visual regions and in the STS+G for words. Although the right VWFA anatomical correlate did not contain significant univariate selectivity, multivariate analyses revealed significant decodability of text conditions from each other, similarly to RH STS+G. The early visual regions of the RH decoded textual stimuli similarly to the LH. Besides STS+G, no other language regions showed consistent text selectivity, nor consistent decodability of text conditions. These findings highlight the partial involvement of the RH in text processing, typically limited to more posterior regions. The STS+G, in particular, seems to be activated by the lexical or semantic content of word-like stimuli and may play an important role in feedback to other regions of the visual word processing network.

Analyses of pairs of ROIs revealed strong functional connectivity (FC) between ROIs within a cluster (i.e., V1–V3, between IPS and VWFA, and between the remaining language regions), as expected, although not as strongly within the high-level vision cluster, which connected strongly with both Early vision and Language ROIs. A similar pattern was observed at the level of the cluster (see Figure 6). While greater within vs. between cluster connectivity was strongly expected for the words condition as the clustering solution was defined on the basis of the group averaged words FC, the same patterns of FC were also observed in the other conditions (Appendix D, Figure S5). Moreover, the other conditions generated the same intermediate clustering solutions (see Appendix C, Figure S2). Thus, the greater within vs. between cluster connectivity cannot simply be explained in terms of non-independence or "double dipping".[42] The FC profile of the RH mirrored that of the LH, although slightly weaker, as reflected both in the t-test between ROI pairs for words against the null distribution and in the two-tailed

t-tests between ROI pairs for words compared with consonant strings and inverted words, respectively. The parallel arrangements of the LH and RH were also revealed in the FC interhemispheric analysis: homologous pairs of ROIs across hemispheres were more strongly connected than non-homologous pairs, with V1, IPS, VWFA, and STS+G being the most strongly functionally connected to their contralateral homologues. Of particular interest, both within- and between-hemisphere FC between most pairs of ROIs held to an equivalent extent across stimulus types.

The final analysis explored circuit wide connectivity by calculating the node connectivity of each ROI with all other ROIs or with each of the three clusters separately. Node strength of most ROIs was consistently greater in the LH than RH and similar to FC, exhibited little modulation by stimulus type (see Figure 7). In the LH, across all stimulus types, STS+G and VWFA exhibited the highest average node strength, whereas PSL had the lowest average node strength, but, in the RH, compared with most other ROIs, the VWFA had significantly lower node strength, and IFG, IPS, and STS+G had higher node strength. In both hemispheres, most ROIs did not exhibit significant modulation of their node strength by stimulus type. Bilaterally, the three clusters appeared to serve as "subnetworks" of a broader visual word processing network, with greater within- than between-cluster connectivity and with each of the ROIs exhibiting preferential connectivity to its native cluster. Of note, though, there was also differential weighting preferences of the non-native clusters for each ROI—all ROIs in the Early vision and Language clusters exhibited higher connectivity to the High-level vision cluster than to their other non-native cluster, implicating the High-level vision cluster in mediating information flow and integration between the Early vision and Language clusters. Unlike any of the other ROIs, the VWFA in both hemispheres showed highly similar node connectivity to its non-native clusters (Early vision and Language).

One noteworthy outcome of this study is the identification of several regions involved in visual word processing in the LH, with IFG, the Precentral Gyrus, and the Superior Temporal Sulcus and Gyrus showing significant univariate activation (greater selectivity for words > inverted words > consonant strings) and decoding accuracy almost as high as VWFA.[73] That this result has not been reported widely is likely a function of the fact that most studies focus specifically on the VWFA. Along similar lines, several regions in the RH also exhibited significant selectivity and decoding accuracy. The RH STS+G showed the strongest sensitivity to text as evident in univariate selectivity, decoding accuracy, and even FC profile, particularly compared with other RH regions such as IPS and VWFA. All ROIs displayed strong interhemispheric connectivity with their homologous contralateral counterpart. Hence, together, these various analyses bring out subtleties in the interactions among this broad network of ROIs that spans both hemispheres, as well as the selective modulation of this network by differing stimulus types (as described below).

It is interesting that many ROIs exhibit stimulus-dependent differences in their intrinsic selectivity and decoding ability among text conditions. One notable instance of this is the differences in modulation by text type for the VWFA compared with STS+G. Whereas the responses of the former differ for upright versus inverted words, the responses of the latter have greater differentiation between words (both upright and inverted) versus consonant strings. Although these regional differences are useful for highlighting the role of these two regions in reading, the differences need to be clarified further. Other studies have also compared visual word and nonword processing (pseudowords, Stevens et al. 2017), but the conditions are so different from those adopted in the present study, and comparisons are not easily made.

Notwithstanding the differential ROI sensitivities to stimulus type, when interactions between ROIs are analyzed (i.e., their pairwise functional connectivity or their node strength within an inter-dependent interconnected cluster of ROIs), they are less modulated by stimulus type and, instead, exhibit stability. This is reflected in the largely invariant response to different stimuli within and largely across hemispheres, as well. This stimulus-independent profile of the interconnected cluster of ROIs confers a well-defined structure to the overall visual word processing network, which comprises widespread cortical areas associated with early and higher-level visual cortex, and language regions. On the other hand, the stimulus-dependent selectivity and multivariate voxel-wise pattern of the individual ROI entities demonstrate the perceptual sensitivity of the network's responses to stimulus types. The simultaneous presence of response flexibility and connectivity stability is consistent with artificial neural networks, in which different stimuli engage different representations to differing degrees even with a single set of weights, and with studies of physiology in which population level activity maintains consistent properties so that the network exhibits homeostasis.[21,22]

We have found several indications of involvement of the RH, particularly the STS+G ROI. The obvious question, though, is whether RH engagement is epiphenomenal or whether the RH does, indeed, play a functional role. One possibility is that the RH has no involvement in visual word processing and that through tight coupling of homologous areas via the callosum, activation might merely cascade interhemispherically from the LH. If this were the case, then we might expect all RH regions to "inherit" the activation from the LH but this is not the case at all, and early visual regions, VWFA and STS+G are more prominently responsive than the other ROIs. Thus, a more parsimonious explanation is that the particular RH activation we discovered is functionally relevant.

Recent findings confirm that the RH homologue of Wernicke's area (STS+G) retains some language capabilities even if as a "weak shadow" to the LH prowess in the mature brain.[53,54] These language capabilities appear sufficient for visual word processing following left occipito-temporal cortex resection, especially in childhood, with the RH VWFA activated by text in many cases.[16,49] Relatedly, children with a single hemisphere (post-hemispherectomy for the management of medically resistant epilepsy) score, on average, 80% accuracy on a word discrimination task, and this is so even with just a preserved RH.[32] Also, in adult patients with prosopagnosia following focal damage to the RH, a moderate reading deficit, less severe than after LH damage, can be detected,[3] and in an adult patient with (partially recovered) pure alexia, stimulation of the RH, but not the LH, disrupted oral reading.[17] Together, these data also reveal meaningful and functional involvement of RH circuitry in visual word processing. Further investigation of the contribution of the RH to text processing is warranted, especially in the context of causal, rather than correlational, approaches.[59]

Another outcome of this research is that, using a data-driven clustering approach, we found that the ROIs aggregated into clusters easily labeled as Early vision, High-level vision, and Language. Moreover, within the LH High-level vision subnetwork, the well-situated VWFA appears to play a more prominent mediating role than IPS as evident from the consistently stronger results across all analytic approaches. The LH VWFA exhibits significantly greater selectivity compared with the IPS for all three stimulus types, can differentiate between the three textual stimulus types more accurately compared with the IPS, and displays significant functional connectivity with the largest number of ROIs across all three clusters (Figure 6). These characteristics of the VWFA allow it to play a more "privileged" role than IPS in mediating information between the Early vision subnetwork and the Language subnetwork. Moreover, the geographic location of the LH VWFA is well suited to its synthesis of bottom-up visuospatial input features from occipital cortex with more conceptual representations such as those involved in language processing.[62] The hierarchical arrangement among the three subnetworks in the RH is less clear, and the mediation role of the VWFA is present but less obvious presumably because of the RH's 'minor', non-dominant language status.

A final consideration concerns the role of dorsal cortex in text processing and, specifically, here, the role of the IPS. It is interesting to note that IPS clusters with VWFA irrespective of whether the clustering is based on responses to words, consonant strings, or inverted words. Whether IPS has a specific functional contribution in the domain of text processing or plays a more general role in the processing of all visual stimuli is still debatable. A host of possible roles has been attributed to the IPS—one such role concerns the attentional processing of the input. In a systematic manipulation of spatial attention in a letter processing task, endogenous valid cues, to a greater degree than exogenous cues, reduced errors due to interference from proximal letters, especially for the most crowded letter positions.[65] Whether such an effect of endogenous attention operates on all visual inputs in which crowding of elements is present remains to be determined. Support for the generality of this function, however, comes from the finding that almost all VTC fROIs are modulated by attention and not just the VWFA[47] (for related work on top-down modulation by the frontoparietal control network, see[64]). But several other functional roles have also been considered. IPS is well known to subserve saccadic eye movements,[41] and even though we used only five-letter words, generally regarded as too short to warrant a saccade, crowding of the letters may require more careful decoding of the letters. IPS also plays a role in perceptual-motor coordination and may also be invoked in mnemonic functions such as working memory.[11] We have no evidence to ascribe the IPS activation observed here to any one (or more than one) specific function and, clearly, the function of the IPS, and its apparently robust relationship to VWFA remains to be adjudicated.

Conclusions

There has been a growing realization that regions of the brain are deeply interconnected and that behavior is the emergent property of distributed networks rather than the outcome of computations in a single, circumscribed region of cortex.[72] Taken together, the strong evidence for engagement of the LH early visual, high-level vision, and language areas, along with the provocative findings of the contribution of many of the same regions in the RH, suggest that the whole-brain organization of the visual word processing network is not simply confined to one hemisphere nor to one region. Rather, there are multiple regions engaged by textual input, and this is so in both the right and left hemispheres. There is flexibility in individual ROIs, as reflected by the modulation of activation by input type, and yet, at the level of the network, there is less modulation by input type and, hence, greater stability. Despite the bilateral engagement, the LH appears primary, consistent with increasing representational distinctiveness in the left VTC and increasing reading performance in children, beyond the effect of age.[56]

Limitations of the study

We have provided a detailed characterization of the structure (the multiplicity of nodes and edges) and behavior (stimulus-driven modulation) of the visual word processing network. Nonetheless, our study has some limitations that should be acknowledged. First, it focused solely on group-level analyses and identified commonalities in network dynamics during visual word processing across subjects. However, delving into the extensive individual differences in localized selectivity, functional connectivity and graph theoretic properties that emerged within the network were beyond the scope of the current analyses. A better characterization of these individual differences both within- and between-hemispheres, the structural and functional factors that influence them, and their association with variations in behavior across individuals warrants future research.[36,44,52,55,69] Second, the present study revealed significant engagement of the right hemisphere (particularly the STS+G) during visual word processing. Yet, the precise role played by the right hemisphere during visual word processing remains elusive, providing an exciting avenue for exploration in future research. One possibility is that the activation in the RH is merely the propagation across the corpus callosum from the LH homotopic region but this seems unlikely in that the RH STS+G is more activated by words than other regions that also have LH homologues. Third, the visual text processing network defined in this study is based on the analysis of data collected from native English speakers and using textual stimuli that were presented in English. Therefore, it is unclear whether the current network characterization is generalizable to broader populations and in the context of word processing in other languages, especially those with different orthographies and sound-to-spelling mappings. Finally, the adoption of a task of single-word sequential discrimination is unlike the fluent reading of the mature reader. Approximating a more naturalistic scenario of word reading, with all of its complexity that extends beyond perceptual processing, is needed to advance a deeper understanding of the process of reading. A thorough behavioral assessment in parallel with elucidating the neural circuitry will inform a deeper understanding of hierarchical organization and flexibility and stability associated with different task conditions.[77] Moreover, temporal dynamics that unfold during the process of reading await further explication.[74,80]

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - fMRI experiment
  - Functional data pre-processing
  - fMRI general linear model (GLM)
  - Functional selectivity
  - Region of Interest (ROI) definition
  - Network cluster definition
  - ROI and cluster-level selectivity
  - Multi-voxel pattern analysis (MVPA)
  - Stimulus-based functional connectivity
  - Graph theory
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.108809.

ACKNOWLEDGMENTS

This research was supported by grants from NSF to M.B. and D.C.P. (BCS2123069), Undergraduate Research Fellowship to R.V., and a Neuroscience Institute Presidential Fellowship to N.B. M.B. acknowledges support from P30 CORE award EY08098 from the National Eye Institute, NIH, and unrestricted supporting funds from The Research to Prevent Blindness Inc, NY, and the Eye & Ear Foundation of Pittsburgh. This work was supported by CMU-Pitt BRIDGE Center development funds (https://coremarketplace.org/RRID:SCR_023356). We thank Dr. Anne Margarette S. Maallo and Roshni Nischal for their assistance with data collection.

AUTHOR CONTRIBUTIONS

Conceptualization: M.B., D.C.P., N.M.B., and R.V.; methodology: N.M.B., R.V., M.B., and D.C.P; software: R.V. and N.M.B.; formal analysis: R.V., N.M.B., and M.B.; investigation: N.M.B.; data curation: N.M.B. and R.V.; writing original draft: M.B., R.V., and N.M.B.; writing review and editing: M.B., R.V., N.M.B., and D.C.P.; visualization: R.V. and N.M.B.; supervision: M.B., D.C.P., and N.M.B.; project administration: M.B. and D.C.P.; funding acquisition: M.B. and D.C.P.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 8, 2023
Revised: December 20, 2023
Accepted: January 2, 2024
Published: January 6, 2024

REFERENCES

1. Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., and Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. Front. Neuroinf. *8*.
2. Baker, C.I., Liu, J., Wald, L.L., Kwong, K.K., Benner, T., and Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. Proc. Natl. Acad. Sci. USA *104*, 9087–9092.
3. Behrmann, M., and Plaut, D.C. (2014). Bilateral hemispheric processing of words and faces: Evidence from word impairments in prosopagnosia and face impairments in pure alexia. Cerebr. Cortex *24*, 1102–1118.
4. Behrmann, M., Plaut, D.C., and Nelson, J. (1998). A literature review and new data supporting an interactive account of letter-by-letter reading. Cogn. Neuropsychol. *15*, 7–51.
5. Behzadi, Y., Restom, K., Liau, J., and Liu, T.T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fmri. Neuroimage *37*, 90–101.
6. Bertoni, S., Franceschini, S., Campana, G., and Facoetti, A. (2023). The effects of bilateral posterior parietal cortex trns on reading performance. Cerebr. Cortex *33*, 5538–5546.
7. Bonald, T., Charpentier, B., Galland, A., and Hollocou, A. (2018). Hierarchical graph

clustering using node pair sampling. Preprint at arXiv.
8. Bouhali, F., Bézagu, Z., Dehaene, S., and Cohen, L. (2019). A mesial-to-lateral dissociation for orthographic processing in the visual cortex. Proc. Natl. Acad. Sci. USA *116*, 21936–21946.
9. Bouhali, F., Thiebaut de Schotten, M., Pinel, P., Poupon, C., Mangin, J.F., Dehaene, S., and Cohen, L. (2014a). Anatomical connections of the visual word form area. J. Neurosci. *34*, 15402–15414.
10. Bouhali, F., Thiebaut de Schotten, M., Pinel, P., Poupon, C., Mangin, J.-F., Dehaene, S., and Cohen, L. (2014b). Anatomical connections of the visual word form area. J. Neurosci. *34*, 15402–15414.
11. Bray, S., Almas, R., Arnold, A.E.G.F., Iaria, G., and MacQueen, G. (2015). Intraparietal sulcus activity and functional connectivity supporting spatial working memory manipulation. Cerebr. Cortex *25*, 1252–1264.
12. Caffarra, S., Karipidis, I.I., Yablonski, M., and Yeatman, J.D. (2021). Anatomy and physiology of word-selective visual cortex: From visual features to lexical processing. Brain Struct. Funct. *226*, 3051–3065.
13. Chen, L., Wassermann, D., Abrams, D.A., Kochalka, J., Gallardo-Diez, G., and Menon, V. (2019). The visual word form area (vwfa) is part of both language and attention circuitry. Nat. Commun. *10*, 5601.
14. Cohen, L., and Dehaene, S. (2004). Specialization within the ventral stream: The case for the visual word form area. Neuroimage *22*, 466–476.
15. 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.
16. Cohen, L., Henry, C., Dehaene, S., Martinaud, O., Lehéricy, S., Lemer, C., and Ferrieux, S. (2004). The pathophysiology of letter-by-letter reading. Neuropsychologia *42*, 1768–1780.
17. Coslett, H.B., and Monsul, N. (1994). Reading with the Right-Hemisphere: Evidence from Transcranial Magnetic Stimulation. Brain Lang. *46*, 198–211.
18. Cox, R.W., and Hyde, J.S. (1997). Software tools for analysis and visualization of fmri data. NMR Biomed. *10*, 171–178.
19. Dębska, A., Wójcik, M., Chyl, K., Dzięgiel-Fivet, G., and Jednoróg, K. (2023). Beyond the visual word form area – a cognitive characterization of the left ventral occipitotemporal cortex. Front. Hum. Neurosci. *17*, 1199366.
20. Dehaene, S., and Cohen, L. (2011). The unique role of the visual word form area in reading. Trends Cognit. Sci. *15*, 254–262.
21. Driscoll, L.N., Duncker, L., and Harvey, C.D. (2022). Representational drift: Emerging theories for continual learning and experimental future directions. Curr. Opin. Neurobiol. *76*, 102609.
22. Driscoll, L.N., Pettit, N.L., Minderer, M., Chettih, S.N., and Harvey, C.D. (2017). Dynamic reorganization of neuronal activity patterns in parietal cortex. Cell *170*, 986–999.e16.
23. Esteban, O., Blair, R., Markiewicz, C.J., Berleant, S.L., Moodie, C., Ma, F., Isik, A.I., Erramuzpe, A., Kent, J.D., Goncalves, M., DuPre, E., et al. (2018a). Fmriprep (Software).
24. Esteban, O., Markiewicz, C.J., Blair, R.W., Moodie, C.A., Isik, A.I., Erramuzpe, A., Kent, J.D., Goncalves, M., DuPre, E., Snyder, M., et al. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. Nat. Methods *16*, 111–116.
25. Fan, X., Guo, Q., Zhang, X., Fei, L., He, S., and Weng, X. (2023). Top-down modulation and cortical-amg/hpc interaction in familiar face processing. Cerebr. Cortex *33*, 4677–4687.
26. Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., and Kanwisher, N. (2010). New method for fMRI investigations of language: defining ROIs functionally in individual subjects. J. Neurophysiol. *104*, 1177–1194.
27. Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., and Constable, R.T. (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat. Neurosci. *18*, 1664–1671.
28. Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., et al. (2016). A multi-modal parcellation of human cerebral cortex. Nature *536*, 171–178.
29. Gomez, J., Barnett, M.A., Natu, V., Mezer, A., Palomero-Gallagher, N., Weiner, K.S., Amunts, K., Zilles, K., and Grill-Spector, K. (2017). Microstructural proliferation in human cortex is coupled with the development of face processing. Science *355*, 68–71.
30. Gorgolewski, K., Burns, C.D., Madison, C., Clark, D., Halchenko, Y.O., Waskom, M.L., and Ghosh, S.S. (2011). Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. Front. Neuroinf. *5*, 13.
31. Gorgolewski, K.J., Esteban, O., Markiewicz, C.J., Ziegler, E., Ellis, D.G., Notter, M.P., Jarecka, D., Johnson, H., Burns, C., Manhães-Savio, A., et al. (2018). Nipype (Software).
32. Granovetter, M.C., Robert, S., Ettensohn, L., Behrmann, M., Lambon-Ralph, M., Bottini, G., Leff, A., Grover, A.P., and Chamanzar, M.B. (2022). With childhood hemispherectomy, one hemisphere can support-but is suboptimal for-word and face recognition. In Proceedings of the National Academy of Sciences of the United States of America, *119*Proceedings of the National Academy of Sciences of the United States of America.
33. Greve, D.N., and Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. Neuroimage *48*, 63–72.
34. Grotheer, M., Yeatman, J., and Grill-Spector, K. (2021). White matter fascicles and cortical microstructure predict reading-related responses in human ventral temporal cortex. Preprint at bioRxiv *227*, 117669.
35. Haugg, A., Frei, N., Menghini, M., Stutz, F., Steinegger, S., Röthlisberger, M., and Brem, S. (2023). Self-regulation of visual word form area activation with real-time fmri neurofeedback. Sci. Rep. *13*, 9195.
36. Hirshorn, E.A., and Harris, L.N. (2022). Culture is not destiny, for reading: Highlighting variable routes to literacy within writing systems. Ann. N. Y. Acad. Sci. *1513*, 31–47.
37. JASP Team (2023). JASP. [Computer software]. Version 0.18.1.
38. Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage *17*, 825–841.
39. Kaestner, E., Wu, X., Friedman, D., Dugan, P., Devinsky, O., Carlson, C., Doyle, W., Thesen, T., and Halgren, E. (2022). The precentral gyrus contributions to the early time-course of grapheme-to-phoneme conversion. Neurobiol. Lang. *3*, 18–45.
40. Kanwisher, N., McDermott, J., and Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. *17*, 4302–4311.
41. Kennard, C. (2011). Disorders of higher gaze control. Handb. Clin. Neurol. *102*, 379–402.
42. Kriegeskorte, N., Mur, M., and Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. Front. Syst. Neurosci. *2*, 4.
43. Kristanto, D., Liu, M., Liu, X., Sommer, W., and Zhou, C. (2020). Predicting reading ability from brain anatomy and function: From areas to connections. Neuroimage *218*, 116966.
44. Kubota, E., Grotheer, M., Finzi, D., Natu, V.S., Gomez, J., and Grill-Spector, K. (2022). White matter connections of high-level visual areas predict cytoarchitecture better than category-selectivity in childhood, but not adulthood. Cerebr. Cortex *33*, 2485–2506. page bhac221.
45. Lanczos, C. (1964). Evaluation of noisy data. J. Soc. Ind. Appl. Math. B Numer. Anal. *1*, 76–85.
46. Lerma-Usabiaga, G., Carreiras, M., and Paz-Alonso, P.M. (2018). Converging evidence for functional and structural segregation within the left ventral occipitotemporal cortex in reading. Proc. Natl. Acad. Sci. USA *115*, E9981–E9990.
47. Li, J., Hiersche, K.J., and Saygin, Z.M. (2023). Demystifying the visual word form area: Precision fmri of visual, linguistic, and attentional properties of ventral temporal cortex. Preprint at bioRxiv.
48. Li, J., Osher, D.E., Hansen, H.A., and Saygin, Z.M. (2020). Innate connectivity patterns drive the development of the visual word form area. Sci. Rep. *10*, 18039.
49. Liu, T., Granovetter, M., Robert, S., J.Z., F., Patterson, C., and Behrmann, M. (in prep.). Microgenesis of plasticity in human visual cortex following large cortical resection. In prep
50. Lochy, A., Jacques, C., Maillard, L., Colnat-Coulbois, S., Rossion, B., and Jonas, J. (2018). Selective visual representation of letters and words in the left ventral occipito-temporal cortex with intracerebral recordings. Proc. Natl. Acad. Sci. USA *115*, E7595–E7604.
51. López-Barroso, D., Thiebaut de Schotten, M., Morais, J., Kolinsky, R., Braga, L.W., Guerreiro-Tauil, A., Dehaene, S., and Cohen, L. (2020). Impact of literacy on the functional connectivity of vision and language related networks. Neuroimage *213*, 116722.
52. Mao, J., Liu, L., Perkins, K., and Cao, F. (2021). Poor reading is characterized by a more connected network with wrong hubs. Brain Lang. *220*, 104983.
53. Martin, K.C., Seydell-Greenwald, A., Berl, M.M., Gaillard, W.D., Turkeltaub, P.E., and Newport, E.L. (2022). A weak shadow of early life language processing persists in the right hemisphere of the mature brain. Neurobiol. Lang. *3*, 364–385.
54. Newport, E.L., Seydell-Greenwald, A., Landau, B., Turkeltaub, P.E., Chambers, C.E., Martin, K.C., Rennert, R., Giannetti, M., Dromerick, A.W., Ichord, R.N., et al. (2022). Language and developmental plasticity after perinatal stroke. In Proceedings of the National Academy of Sciences of the United

States of AmericaProceedings of the National Academy of Sciences of the United States of America.
55. Nordt, M., Gomez, J., Natu, V.S., Rezai, A.A., Finzi, D., Kular, H., and Grill-Spector, K. (2021). Cortical recycling in high-level visual cortex during childhood development. Nat. Human Behav. *5*, 1686–1697.
56. Nordt, M., Gomez, J., Natu, V.S., Rezai, A.A., Finzi, D., Kular, H., and Grill-Spector, K. (2022). Longitudinal development of category representations in ventral temporal cortex predicts word and face recognition. bioRxiv.
57. Oldfield, R.C. (1971). The assessment and analysis of handedness: the edinburgh inventory. Neuropsychologia *9*, 97–113.
58. Di Pietro, S.V., Willinger, D., Frei, N., Lutz, C., Coraj, S., Schneider, C., Stämpfli, P., and Brem, S. (2023). Disentangling influences of dyslexia, development, and reading experience on effective brain connectivity in children. Neuroimage *268*, 119869.
59. Planton, S., Wang, S., Bolger, D., Bonnard, M., and Pattamadilok, C. (2022). Effective connectivity of the left-ventral occipito-temporal cortex during visual word processing: Direct causal evidence from tms-eeg co-registration. Cortex *154*, 167–183.
60. Power, J.D., Mitra, A., Laumann, T.O., Snyder, A.Z., Schlaggar, B.L., and Petersen, S.E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fmri. Neuroimage *84* (*Supplement C*), 320–341.
61. Price, C.J., and Devlin, J.T. (2003). The myth of the visual word form area. Neuroimage *19*, 473–481.
62. Price, C.J., and Devlin, J.T. (2011). The interactive account of ventral occipitotemporal contributions to reading. Trends Cognit. Sci. *15*, 246–253.
63. Price, C.J., and Mechelli, A. (2005). Reading and reading disturbance. Curr. Opin. Neurobiol. *15*, 231–238.
64. Qin, L., Lyu, B., Shu, S., Yin, Y., Wang, X., Ge, J., Siok, W.T., and Gao, J.H. (2021). A heteromodal word-meaning binding site in the visual word form area under top-down frontoparietal control. J. Neurosci. *41*, 3854–3869.
65. Ramamurthy, M., White, A.L., Chou, C., and Yeatman, J.D. (2021). Spatial attention in encoding letter combinations. Sci. Rep. *11*, 24179.
66. Rosazza, C., Appollonio, I., Isella, V., and Shallice, T. (2007). Qualitatively different forms of pure alexia. Cogn. Neuropsychol. *24*, 393–418.
67. Rosenke, M., Weiner, K.S., Barnett, M.A., Zilles, K., Amunts, K., Goebel, R., and Grill-Spector, K. (2018). A cross-validated cytoarchitectonic atlas of the human ventral visual stream. Neuroimage *170*, 257–270.
68. Satterthwaite, T.D., Elliott, M.A., Gerraty, R.T., Ruparel, K., Loughead, J., Calkins, M.E., Eickhoff, S.B., Hakonarson, H., Gur, R.C., Gur, R.E., and Wolf, D.H. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. Neuroimage *64*, 240–256.
69. Shechter, A., Hershman, R., and Share, D.L. (2022). A pupillometric study of developmental and individual differences in cognitive effort in visual word recognition. Sci. Rep. *12*, 10764.
70. Stevens, W.D., Kravitz, D.J., Peng, C.S., Tessler, M.H., and Martin, A. (2017). Privileged functional connectivity between the visual word form area and the language system. J. Neurosci. *37*, 5288–5297.
71. Stigliani, A., Weiner, K.S., and Grill-Spector, K. (2015). Temporal processing capacity in high-level visual cortex is domain specific. J. Neurosci. *35*, 12412–12424.
72. Tooley, U.A., Park, A.T., Leonard, J.A., Boroshok, A.L., McDermott, C.L., Tisdall, M.D., Bassett, D.S., and Mackey, A.P. (2022). The age of reason: Functional brain network development during childhood. J. Neurosci. *42*, 8237–8251. JN–RM–0511–22.
73. Turkeltaub, P.E., Eden, G.F., Jones, K.M., and Zeffiro, T.A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. Neuroimage *16*, 765–780.
74. Vida, M.D., Nestor, A., Plaut, D.C., and Behrmann, M. (2017). Spatiotemporal dynamics of similarity-based neural representations of facial identity. Proc. Natl. Acad. Sci. USA *114*, 388–393.
75. Weiner, K.S., Barnett, M.A., Lorenz, S., Caspers, J., Stigliani, A., Amunts, K., Zilles, K., Fischl, B., and Grill-Spector, K. (2017). The cytoarchitecture of domain-specific regions in human high-level visual cortex. Cerebr. Cortex *27*, 146–161.
76. Weiss, K.-L., Hawelka, S., Hutzler, F., and Schuster, S. (2023). Stronger functional connectivity during reading contextually predictable words in slow readers. Sci. Rep. *13*, 5989.
77. White, A.L., Kay, K.N., Tang, K.A., and Yeatman, J.D. (2023). Engaging in word recognition elicits highly specific modulations in visual cortex. Curr. Biol. *33*, 1308–1320.e5.
78. Willenbockel, V., Sadr, J., Fiset, D., Horne, G.O., Gosselin, F., and Tanaka, J.W. (2010). Controlling low-level image properties: the SHINE toolbox. Behav. Res. Methods *42*, 671–684.
79. Woolnough, O., Donos, C., Rollo, P.S., Forseth, K.J., Lakretz, Y., Crone, N.E., Fischer-Baum, S., Dehaene, S., and Tandon, N. (2021). Spatiotemporal dynamics of orthographic and lexical processing in the ventral visual pathway. Nat. Human Behav. *5*, 389–398.
80. Woolnough, O., Forseth, K.J., Rollo, P.S., Roccaforte, Z.J., and Tandon, N. (2022a). Event-related phase synchronization propagates rapidly across human ventral visual cortex. Neuroimage *256*, 119262.
81. Woolnough, O., Snyder, K.M., Morse, C.W., Mccarty, M.J., Lhatoo, S.D., Tandon, N., and Smith, V.L. (2022b). Intraoperative localization and preservation of reading in ventral occipitotemporal cortex. J. Neurosurg. *137*, 1610–1617.
82. Yablonski, M., Rastle, K., Taylor, J.S.H., and Ben-Shachar, M. (2019). Structural properties of the ventral reading pathways are associated with morphological processing in adult english readers. Cortex *116*, 268–285.
83. Yeatman, J.D., Weiner, K.S., Pestilli, F., Rokem, A., Mezer, A., and Wandell, B.A. (2014). The vertical occipital fasciculus: a century of controversy resolved by in vivo measurements. Proc. Natl. Acad. Sci. USA *111*, E5214–E5223.
84. Zhan, M., Pallier, C., Agrawal, A., Dehaene, S., and Cohen, L. (2023). Does the visual word form area split in bilingual readers? a millimeter-scale 7-t fmri study. Sci. Adv. *9*, eadf6140.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
FMRIPREP, for functional data preprocessing	NeuroImaging PREProcessing toolS (NiPreps)	https://fmriprep.org/en/stable/
Freesurfer, for functional data preprocessing (image co-registration and surface resampling)	Laboratories for Computational Neuroimaging (LCN) at the Athinoula A. Martinos Center for Biomedical Imaging	https://surfer.nmr.mgh.harvard.edu/
SPM12, for fMRI data analysis	Wellcome Center for Human Neuroimaging	https://www.fil.ion.ucl.ac.uk/spm/software/spm12/
Paris agglomerative hierarchical clustering algorithm	Bonald et al.,[7]	https://github.com/tbonald/paris
Code to perform selectivity, multi-voxel pattern analysis (MVPA), functional connectivity and graph theory analyses	This paper	https://github.com/viscog-cmu/vin_iscience_2024
Deposited data		
fMRI data collected during performance of 1-back stimulus recognition task on words, inverted words, letter strings, faces, inverted faces and objects	This paper	https://doi.org/10.1184/R1/24659637

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Marlene Behrmann (mbehrmann@pitt.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All data reported in this paper have been deposited at KiltHub (a Carnegie Mellon University research repository) and are publicly available as of the date of publication. DOIs are listed in the key resources table.
- All original code has been deposited at GitHub and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

A sample of 28 right-handed native English speakers (mean age: 22.1 yrs, 20 female), with normal or corrected-to-normal visual acuity, participated in this study (Hispanic or Latino: 1, not Hispanic or Latino: 24; White: 17, Asian: 8, Black or African American: 2, multiracial: 4; three participants did not provide information regarding their race and ethnicity). Participants were not multilingual in childhood, and had no significant neurological or psychiatric history, claustrophobia, history of metal in body, or any other counter-indication for MRI. The group had a mean score on the Edinburgh Handedness Inventory[57] of 83.7, consistent with strong right-handedness. Participants completed a 1-hour imaging session consisting of functional, structural, and diffusion scans, and were paid for their participation. The diffusion data form part of a different study and are not discussed in this paper.

The procedures used in this study were approved by the Institutional Review Board of Carnegie Mellon University. All participants gave informed consent prior to their scanning session.

METHOD DETAILS

fMRI experiment

All functional images were acquired on a Siemens Prisma 3T scanner with a 64-channel head coil at the CMU-Pitt BRIDGE Center (RRID:SCR_023356). Five runs of functional imaging data (TE=30 ms, TR=2000 ms) were collected, at isotropic resolution of $(2mm)^3$. Each

run lasted approximately 5 minutes and contained several mini-blocks of multiple images from a given category presented rapidly at fixation while the subjects performed a 1-back stimulus identity task. A schematic of the experiment and sample stimuli can be seen in Figure 1. We used a modified version of the *fLoc* category localizer,[71] with 6 categories: words, inverted words, consonant strings, faces, inverted faces and objects. Faces were ambient adult face images with natural variability in expression and viewpoint. Objects were a mixture of cars and guitars presented at a variety of orientations. Words were common 5-letter words presented in a variety of fonts. Consonant strings derived from the original words with replacement of vowels with random consonants, but were otherwise identical to the word stimuli in font and presentation. Inverted words and faces were the stimuli from the words and faces categories, with images inverted by a 180 deg rotation. Faces and objects were taken from the original stimuli from.[71] Presenting the stimuli on a scrambled background allows for the total size of each of the images to be constant, reducing low-level artifacts across stimuli and categories. In line with this, each of the images was normalized with the SHINE toolbox to have identical luminance mean and histograms. Moreover, we elected to follow the approach used by,[71] who attempted to make a standardized and publicly available localizer experiment. While we introduced some new categories (consonant strings, inverted words and faces), we followed the same format for creating these stimuli as in the original localizer. Luminance histograms of all stimuli were matched using the histMatch function of the SHINE toolbox[78] and there were 104 per category. Rest mini-blocks were interleaved with the category mini-blocks, with the same frequency and duration as stimulus mini-blocks. Each mini-block lasted 6s, and contained 12 stimuli from the category of interest (or 0 in the control/no-stimulus condition), each presented for 400ms with a 100ms inter-stimulus interval (ISI). Subjects pressed a button on an fMRI compatible glove when detecting a repeat of the previous image (1-back stimulus identity task) which appeared randomly once per run. Participants had 1s to respond from the offset of the repeated image. The order of mini-blocks was fully counterbalanced across categories, such that a mini-block of each category preceded a mini-block of each other category exactly once in a run, resulting in 7 mini-blocks per category. We collected brain scans for 10s before and 14s after the experimental run to allow the scanner to reach steady state, and for the BOLD signal to return to baseline, respectively. At the end of each run, participants received feedback about their performance. The average hit rate was 88.8%, with a standard deviation of 9.8%.

Functional data pre-processing

The functional data were preprocessed using *fMRIPrep* 1.4.1,[23,24] which is based on *Nipype* 1.2.0.[30,31] Many internal operations of *fMRIPrep* use *Nilearn* 0.5.2,[1] mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in *fMRIPrep*'s documentation. Here, we quote the methods output directly from *fMRIPrep* as applied in our experiment, a procedure encouraged by the *fMRIPrep* authors to ensure accuracy of the methods description. Slight modifications were made to remove the details regarding estimation of unused confounds.

For each of the 5 BOLD runs per participant, the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. A deformation field to correct for susceptibility distortions was estimated based on two echo-planar imaging (EPI) references with opposing phase-encoding directions, using 3dQwarp[18] (AFNI 20160207). Based on the estimated susceptibility distortion, an unwarped BOLD reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration.[33] Co-registration was configured with nine degrees of freedom to account for distortions remaining in the BOLD reference. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) were estimated before any spatiotemporal filtering using mcflirt FSL 5.0.9, [38] The BOLD time-series were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as *preprocessed BOLD in original space*, or just *preprocessed BOLD*. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three region-wise global signals. FD and DVARS were calculated for each functional run, both using their implementations in *Nipype* following the definitions by.[60] Additionally, a set of physiological regressors was extracted to allow for component-based noise correction *CompCor*,.[5] Principal components were estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the anatomical *CompCor* variants (aCompCor). This subcortical mask was obtained by heavily eroding the brain mask, which ensures that it does not include cortical GM regions. For aCompCor, components were calculated within the intersection of the aforementioned mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run (using the inverse BOLD-to-T1w transformation). Components were also calculated separately within the WM and CSF masks. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each.[68] All resamplings can be performed with *a single interpolation step* by composing all the pertinent transformations (i.e., head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels.[45] Non-gridded (surface) resamplings were performed using mri_vol2surf (FreeSurfer).

fMRI general linear model (GLM)

fMRI data were analyzed using SPM12. For each subject, a general linear model was specified with an event-related design. Specifically, each mini-block was modeled with its onset and a duration of 6 seconds. The design matrix thus contained 6 stimulus regressors, one per condition. Stimulus timecourses were convolved with a canonical hemodynamic response function. A reduced set of FMRIPREP-generated confounds

was retained for nuisance regression. Specifically, we retained 6 motion parameters (X, Y, Z motion and rotation), the top 6 principal components of the aCompCor decomposition, and the framewise-displacement, yielding a total of 13 nuisance regressors, along with a runwise mean regressor. Finally, an autoregressive-1 model was used within SPM12 to reduce the effects of serial correlations. We fit voxels only within the brain mask defined by *fMRIPrep*. Estimation of the GLM resulted in a beta-weight for each stimulus condition and run, which were used for subsequent univariate and multivariate analyses.

Functional selectivity

SPM12 was used to compute whole-brain statistical maps indexing the selectivity for each condition of interest compared to the others. Selectivity contrasts were balanced in total weight across positive and negative components, as well as across domains (text, faces, objects) in either positive or negative components. We computed selectivity for words, consonant strings, and inverted words as contrasts against other non-text conditions; or example, words > ((faces + inverted faces)/2 + objects)/2. To be clear, the division of (faces + inverted faces) by 2 normalizes the "all faces" domain against objects, and the division of ((faces + inverted faces)/2 + objects) by 2 normalizes the positive and negative components of the contrast. As words, consonant strings, and inverted words evinced similar large-scale responses in ventral regions, we also analyzed a text selectivity contrast: (words + consonant strings + inverted words)/3 > ((faces + inverted faces)/2 + objects)/2. For some analyses, statistical maps used only the even or odd runs of the experiment so as to assess the reliability of the statistical maps.

Region of Interest (ROI) definition

We used the *Glasser* atlas to divide each hemisphere into 180 anatomical parcels[28] from which we identified a set of ten regions of interest in each hemisphere for further analysis (Figure 2). We based our selection of ROIs on group selectivity maps. A surface-based alignment was performed using the *Glasser* atlas aligned to the *fsaverage* subject space; after alignment to each subject's surface, the surface-aligned subject-specific atlas was projected into the cortical ribbon in order to acquire a volumetric atlas consisting of cortical voxels, allowing for analyses in native volumetric space. Based on previous literature supplementary information,[28] we defined each of these anatomical regions of interest in each hemisphere as being composed of one or more Glasser parcels, as listed in the table below. For our multivariate analyses, we used a purely anatomical definition of all regions, as defined in the table; as the VWFA is typically defined functionally, for clarity we refer to the anatomically defined region as 'VWFA_a'. For the univariate selectivity, functional connectivity and node strength analyses, we functionally localized each of the ten ROIs in each hemisphere, including the VWFA, within the anatomical bound of their corresponding Glasser parcels from the table below (discussed in detail in the following sections). Localization was performed using the statistical map with the text selectivity contrast, as described above, and a localization threshold of $p < 0.01$ for the selectivity analyses and $p < 0.1$ for functional connectivity analyses (we used a lax threshold for FC so as to capture and measure signal even from ROIs in the RH, some of which had much weaker selectivity than their homologous counterparts in the LH).

Network cluster definition

To assess the organization of our chosen ROIs into clusters, we took the mean pairwise FC matrix consisting of the correlation values of the time course between ROI pairs, plotted in Figure 3B, and subjected this to the Paris agglomerative hierarchical clustering algorithm.[7] Note that this clustering was performed only for the Words condition, which is typically used to identify the VWFA. To ensure that this clustering was not dependent on the particular agglomerative approach, we repeated this procedure using the Louvain algorithm, with default resolution parameter ($\gamma = 1$). The output confirmed our initial Paris clustering. We take as our structure the clustering at the *highest* level, with the three clusters–Early vision, High-level vision and Language–probed separately in the left and right hemispheres. Statistical analyses at the level of the three-cluster parcellation in each hemisphere were performed by resampling the data using the parcel of the entire cluster rather than simply averaging the activation of its sub-ROIs; see details below in each section. As noted, this clustering was performed with fMRI responses from just the words stimuli. To ensure that this outcome was not specifically governed by the words condition alone, we repeated this analysis for the consonant strings and inverted words as well (see Appendix C, Figure S2). The pattern was highly similar across all stimulus conditions and the same three clusters emerged, with only minor differences at finer-grained levels of clustering below that of the three clusters of interest.

ROI and cluster-level selectivity

In the first set of analyses, we followed standard practices for computing selectivity of ROIs [26,40] by using part of the data (even or odd runs) to localize selective voxels ($p<0.01$ unc) within our anatomical search spaces, and then analyzing the mean selectivity of these voxels in independent data (odd or even runs); to increase the robustness of our results, we performed this analysis in both directions (i.e, localize with even runs and analyze with odd runs, localize with odd runs and analyze with even runs) and averaged the results. We refer to this average as the *localized selectivity*. For individual ROIs, the anatomical search spaces and contrasts were defined as in Table 1. For cluster-level selectivity, the search space consisted of all of the voxels within all of the ROIs that constitute the given cluster, based on the results of the clustering algorithm (Figure 3A).

Additionally, to capture both the strength and size of the selective region within each anatomical parcel, we used the metric of *summed selectivity* following localization of candidate voxels. This approach is similar to the localized selectivity, except rather than taking the mean of selectivity over candidate localized voxels, we take the sum.

Multi-voxel pattern analysis (MVPA)

For the MVPA decoding analyses, we used the beta weights from the GLM that modeled each stimulus condition (i.e., words, inverted words, consonant strings, objects, faces and inverted faces) over each of 5 runs. To account for voxel-level uncertainty in beta-estimates, each beta weight was compared to 0 using a one-hot *t*-test, and the *t*-statistics were retained as the voxel-level features, one per stimulus condition per run. A cross-validated decoding was then performed between pairs of text conditions (words, inverted words and consonant strings). Additionally, we performed decoding of each text condition against objects. Decoding was done at the ROI level by using the complete set of voxels in each ROI, and at the cluster level by using all voxels in the cluster. A linear classifier with a Ridge (i.e., L2) penalty was used, with the optimal regularization strength determined via internal cross-validation over the training runs. Each run served as the test run for a classifier trained on the other four runs, and the mean classification accuracy over runs was retained.

Stimulus-based functional connectivity

To move beyond individual ROIs and clusters, stimulus-based FC analyses were conducted on the cortical surface. For each participant, we used 3 out of the 5 runs to functionally localize the selective voxels within the anatomically defined ROIs from the table above (see Region of Interest definition section). The activation time series of all voxels within the constituent Glasser parcels within each of the ROIs were then averaged for each of the remaining 2 runs. The time series of the 2 runs were then concatenated into a single time series. To study the visual text processing network, we extracted the time series that corresponded to the peak of the hemodynamic response function (the time-frame of 6*s* to 10*s* after condition onset was used to extract the peak, with TR = 2*s*) elicited by each stimulus condition (i.e., the 1-back perception tasks on words, inverted words and consonant strings). Using the same confounds used in our GLM analysis, we extracted the residual denoised BOLD time series of each voxel for each stimulus condition. Finally, the mean denoised time series of pairs of ROIs were correlated using Pearson's *r*, for each stimulus condition. All correlations were further transformed using Fisher's *z* before being analyzed. At the cluster-level, the functional connectivity between all pairs of constituent ROIs was averaged for each of the three stimulus types, in each hemisphere.

Graph theory

Each of the ROIs was considered a 'node' in the visual text processing network. We used the graph-theoretic metric of node strength to analyze the circuit-level properties of the ROIs. Node strength was computed as the sum of the weights of the edges connected to a given ROI; this was performed either globally (i.e., across all edges), or within a cluster (only including ipsilateral edges and the edge to the homologous contralateral ROI). The weight of the edges between each pair of ROIs was measured as the degree of absolute FC between them.

QUANTIFICATION AND STATISTICAL ANALYSIS

In each analysis, for each dependent measure (e.g., univariate selectivity or multivariate decoding accuracy), repeated-measures ANOVAs were conducted with all relevant within-subject factors using the JASP statistical analysis software[37] (see Results section and associated figures for details). All statistical analyses were performed on the entire cohort of subjects (i.e., $n = 28$). Thereafter, to decompose any observed interactions, we used conservative post-hoc Tukey Honest Significant Difference (HSD) correction on the highest order interaction. The HSD measure is a single-step multiple comparison procedure that permits identifying those means that are significantly different from each other. We restricted the number of pairwise comparisons in the HSD calculation to the differences of interest and used $p < 0.01$ to be highly conservative. Any two means that differed from each other by a value greater than the computed critical Tukey HSD value were considered statistically significant. In some analyses, we adopted additional procedures, and these are described in the relevant results sections.