



Published in final edited form as:

Brain Struct Funct. 2021 December ; 226(9) : 3051–3065. doi:10.1007/s00429-021-02384-8.

Anatomy and physiology of word-selective visual cortex: From visual features to lexical processing

Sendy Caffarra^{*,a,b,c,d}, Iliana I. Karipidis^{*,e}, Maya Yablonski^{a,b}, Jason D. Yeatman^{a,b}

^aStanford University School of Medicine, Division of Developmental-Behavioral Pediatrics, 291 Campus Drive, Li Ka Shing Building Stanford, CA 94305-5101, USA

^bStanford University Graduate School of Education, 485 Lasuen Mall, Stanford, CA 94305, USA

^cBasque Center on Cognition, Brain and Language, Mikeletegi 69, 20009, San Sebastian, Spain

^dUniversity of Modena and Reggio Emilia, Via Campi 287, 41125, Modena, Italy

^eCenter for Interdisciplinary Brain Sciences Research, Department of Psychiatry and Behavioral Sciences, School of Medicine, Stanford University, 401 Quarry Road, Stanford, CA 94305-5717, USA

Abstract

Over the past two decades, researchers have tried to uncover how the human brain can extract linguistic information from a sequence of visual symbols. The description of how the brain's visual system processes words and enables reading has improved with the progressive refinement of experimental methodologies and neuroimaging techniques. This review provides a brief overview of this research journey. We start by describing classical models of object recognition in non-human primates, which represent the foundation for many of the early models of visual word recognition in humans. We then review functional neuroimaging studies investigating the word-selective regions in visual cortex. This research led to the differentiation of highly specialized areas, which are involved in the analysis of different aspects of written language. We then consider the corresponding anatomical measurements and provide a description of the main white matter pathways carrying neural signals crucial to word recognition. Finally, in an attempt to integrate structural, functional, and electrophysiological findings, we propose a view of visual word recognition, accounting for spatial and temporal facets of word-selective neural processes. This multi-modal perspective on the neural circuitry of literacy highlights the relevance of a posterior-anterior differentiation in ventral occipitotemporal cortex for visual processing of written language and lexical features. It also highlights unanswered questions that can guide us towards

Corresponding author: Iliana I. Karipidis, Stanford University School of Medicine, Department of Psychiatry and Behavioral Sciences, 401 Quarry Road, Stanford, CA 94305-5717; ilianak@stanford.edu.

* authors contributed equally to the manuscript

Authors' contributions

All authors drafted and/or critically revised the work, SC and IK performed the literature and data search for figure preparation.

Conflicts of interest/Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Availability of data and material

Data to reproduce figure 4 are available at https://github.com/SendyCaffarra/VWFA_coordinates/tree/main

Code availability

Code to reproduce figure 4 is available at https://github.com/SendyCaffarra/VWFA_coordinates/tree/main

future research directions. Bridging measures of brain structure and function will help us reach a more precise understanding of the transformation from vision to language.

Keywords

reading; ventral occipitotemporal cortex; word recognition; fMRI; VWFA; diffusion MRI

Introduction

Literacy presents a fascinating challenge for neuroscientists and cognitive scientists alike: text reaches the retina as patterns of light and, within a couple hundred milliseconds, these visual signals are transformed into sound and meaning. Beginning in the 1990s, early neuroimaging research detailed the core parts of the brain's reading circuitry revealing: (a) the cortical regions involved in different aspects of reading, (b) the major white-matter tracts that carry signals between these regions and (c) how the development of this circuit is related to the process of learning to read (Figure 1). For example, by the early 2000s it was widely accepted that a region of left ventral occipitotemporal cortex (VOTC), termed the visual word form area (VWFA), is involved in rapid and automatic recognition of written words (Cohen et al. 2002; McCandliss et al. 2003), and that the posterior portion of the left superior temporal cortex is involved in encoding the sounds of language (Pugh et al. 1996; Binder et al. 1997). Following the overall trend of mapping the brain during this era, and using available imaging techniques, this level of description provided first answers to the question of *where*, as opposed to *how*, reading is implemented in the cortex. Despite progress in detailing the parts of the reading circuitry, the fundamental challenge of understanding how written text is transformed into language remained elusive.

A recent surge of papers combining detailed anatomical methods, high-resolution fMRI measurements, and intracranial electrophysiology recordings in the human brain, supports new insights into the mechanisms that underlie the transition from parallel processing of visual features to the extraction of lexical properties from text (White et al. 2019a). These new findings confirm some predictions of classic models, but also reveal unforeseen properties of the neural architecture underlying word recognition. In this paper, we first provide a retrospective on the research that shaped the way we think about the visual word form circuitry and laid the foundation for the most influential neurobiological models of written word recognition. Second, we review recent functional neuroimaging findings that moved the field from the premise of a single VWFA to a more granular understanding of multiple distinct visual regions that encode various aspects of written language. This transition in the field is supported by spatially distinct neural responses to different properties of written language within the section of VOTC that has classically been termed the VWFA. Third, we summarize the anatomical literature detailing the white matter anatomy of VOTC and its structural connectivity to other regions of the reading circuit. Finally, we bring the functional and structural literatures together by summarizing the flurry of recent publications that detail the computations of word recognition at a new level of precision. These new observations raise questions that foreshadow future efforts to understand how the brain's visual system enables reading.

The ventral visual stream and theories of word recognition

The notion of a visual word form area, or region of visual cortex that develops with literacy to instantiate expertise in the patterns of written language, dates back to neurological case studies in the late 1800s (Déjerine 1891; Bub et al. 1993; Graves 1997). Many modern theories describing the series of operations performed by the human visual system to process written language are grounded in the neurophysiology of object recognition in the macaque ventral visual stream. In macaque, a series of regions beginning in early visual cortex (e.g., V1), progressing through intermediate stages (e.g., V4), and continuing to inferior temporal (IT) regions (TE, TEO) compute increasingly abstract representations of visual images supporting object recognition (Gross et al. 1967, 1972; DiCarlo et al. 2012). The macaque ventral visual stream is often modeled as a hierarchy where each region inherits the representation computed by the previous region, pools over a larger portion of visual space, and constructs a new representation with greater invariance to properties such as size and location on the retina that are not important for recognition (Riesenhuber and Poggio 1999; Hung et al. 2005; Khaligh-Razavi and Kriegeskorte 2014). The goal of this architecture is to achieve neural populations that are selective for specific objects (or categories of objects) and invariant to the low-level properties of those objects. For example, while neurons in V1 selectively respond to edges at a specific retinal location and orientation (Hubel and Wiesel 1962; Hubel 1995; Heeger et al. 1996), neurons in V4 selectively respond to line junctions at specific angles over a larger portion of retinal space (Nandy et al. 2013; Kim et al. 2019). Neurons in IT respond to specific shapes over a much larger portion of the visual field (Tanaka et al. 1991; Rust and Dicarlo 2010; DiCarlo et al. 2012). Thus, by integrating adjacent features at multiple stages, recordings from IT neurons can distinguish different types of objects irrespective of their exact location, orientation or luminance. It is interesting to note that when young macaques are trained to recognize letters they develop a region in IT cortex that selectively responds to letter shapes and might be homologous to the word-selective regions in the human occipitotemporal sulcus (OTS) that instantiate expertise with written language (Srihasam et al. 2012, 2014).

Word recognition was hypothesized to depend on a similar hierarchy in the human brain whereby neurons in early visual cortex detect line segments at specific locations, neurons in intermediate visual stages (e.g., hV4) combine these features to detect individual letters at specific locations, and neurons in VOTC assemble combinations of letters to detect individual words (Dehaene et al. 2005). This late stage of processing images for visual word recognition is believed to occur in the “visual word form area”, a region of VOTC that selectively responds to written words (Dehaene et al. 2002; Hasson et al. 2002; Dehaene and Cohen 2011). First evidence for this hierarchy came from studies finding a spatial gradient of sensitivity to orthographic properties of text (Kronbichler et al. 2004; Binder et al. 2006; Brem et al. 2006; Vinckier et al. 2007). This work laid the foundation for the idea of a hierarchical system supporting word recognition in the fusiform gyrus that then led to studies probing anterior and posterior parts of this system for print sensitivity, orthography effects, and functional connectivity to other brain regions (van der Mark et al. 2009, 2011).

The idea that visual word form processing involves a series of cortical regions with specific functional properties reflects what we know about the visual system: a set of regions

process specific aspects of visual input, such as motion, color, shape and retinal location. Even within specific categories of visual objects such as faces, different features seem to be encoded in separate regions (Freiwald and Tsao 2010). Objects that are behaviorally relevant for a species are likely recognized by combining this information that is encoded in different visual regions. However, the details of the computations performed by specific VOTC regions within this posterior-anterior axis remained hypothetical and many questions remained unanswered: Does activation in the left VOTC flow through a gradient or reflect distinct computations performed by specific subregions? Are VOTC regions organized in a hierarchy? At what point are visual features combined to compute an invariant representation of a word? How does the visual system interface with language regions to transition from visual processing of image features to the extraction of lexical properties from text? How does feedback from higher order cognitive processes facilitate specialization within visual cortex?

Functional responses to written words in ventral occipitotemporal cortex

Starting from the early 2000s, a major endeavor in neuroimaging research was the spatial localization of functional responses to written language in the human visual system (Figure 2). To this aim, different experimental designs were used to localize the parts of VOTC that preferentially respond to written language. fMRI responses to text have been compared to a variety of visual control stimuli e.g., symbols or false fonts presented in isolation or as a string of characters, as well as other types of non-linguistic visual objects such as faces, tools, houses (Hasson et al. 2002; Cohen et al. 2002; Rossion et al. 2003; Gaillard et al. 2006; Baker et al. 2007; Ben-Shachar et al. 2007; Dehaene et al. 2010; Stigliani et al. 2015). In addition, a wide range of orthographic stimuli have been examined and compared with each other e.g., isolated letters, consonant strings, pseudowords, and words (Binder et al. 2006; Vinckier et al. 2007; Glezer et al. 2009; Thesen et al. 2012). These neuroimaging studies consistently reported responses to written language in a portion of left VOTC including the fusiform gyrus and OTS, which can be differentiated from other VOTC areas showing sensitivity to non-linguistic visual object categories such as faces (Malach et al. 2002; Hasson et al. 2002; Grill-Spector and Weiner 2014). The reason for this consistent left lateralization and its purpose are still unclear. However, mounting evidence points to the proximity and connectivity with left-dominant brain areas involved in spoken language production and comprehension (e.g., Broca's area; Van der Haegen et al. 2012; Gerrits et al. 2019), as likely reasons for the left hemispheric lateralization of word-selective regions in visual cortex.

The majority of the abovementioned studies described regions within left VOTC as the source of word-selective responses and supported the notion that visual regions specialized for text are located at a high level of the visual system. It is worth noting that the term “word-selective cortex” is often used to broadly refer to the portion of VOTC that preferentially responds to visual words (and text more broadly) without taking a stance on the function of this region. The term “visual word form area” is generally used interchangeably with the term “word-selective cortex” and attributes a specific role in encoding orthographic patterns (Figure 3). Different experimental contrasts have been employed to isolate brain areas that are sensitive to text: word-selective responses have

been investigated by probing either perceptual features of words with comparisons of words against non-orthographical stimuli, such as symbols (Hasson et al. 2002; Cohen et al. 2002; Gaillard et al. 2006; Dehaene et al. 2010), or by manipulating lexical properties, with more fine-grained comparisons of words against letter strings or pseudowords (Cohen et al. 2002; Dehaene et al. 2004; Binder et al. 2006; Gaillard et al. 2006; Vinckier et al. 2007).

The type of response selectivity examined in these different comparisons is heterogeneous. While comparing words with non-orthographic material identifies portions of the cortex sensitive to both perceptual and linguistic aspects of text (text-selective responses), the comparison between word and word-like stimuli isolates areas that are sensitive to the lexical properties of the visual input (lexical-selective responses). An examination of the spatial coordinates reported in previous studies shows substantial spatial variability spanning many centimeters of cortex (Figure 4). This variability may be the result of differences in the sampled populations and stimuli, with most studies adopting cross-sectional designs to compare adults with children (Brem et al. 2009; Martin et al. 2015; Centanni et al. 2017), literates with illiterates adults (Dehaene et al. 2010; Skeide et al. 2017), and children with different levels of reading expertise (Pleisch et al. 2019; Brem et al. 2020), while studies following children longitudinally throughout the stages of reading acquisition are still scarce (Ben-Shachar et al. 2011; Chyl et al. 2021). In addition, it is also likely that this spatial variability stems from the fact that different stimuli and tasks are, in fact, tapping into different sub-regions within the word-selective cortex. Here, we use the term "word-selective" to broadly refer to all parts of the cortex that respond stronger to written words compared to other categories of visual stimuli including not only "lexical-selective" but also "text-selective" regions.

Investigations of the spatial layout of face-, object- and limb-selective regions in humans and macaques have demonstrated a sequence of at least three discrete regions that are selective for each category and span centimeters of the posterior-anterior axis of VOTC (Downing et al. 2001; Weiner and Grill-Spector 2010; Freiwald and Tsao 2010; Grill-Spector and Weiner 2014; Stigliani et al. 2015; Weiner et al. 2017a; Park et al. 2017; Bao et al. 2020). Assuming that the cerebral architecture for reading follows organizational principles that are similar to those observed for face and object recognition (Liu et al. 2010; Grill-Spector and Weiner 2014; Stigliani et al. 2015), we should expect to see multiple, distinct word-selective regions that are distributed along the posterior-anterior axis of left VOTC (Lerma-Usabiaga et al. 2018; White et al. 2019a; Poeppel et al. 2020). The variability in the localization of word-selective responses across different studies might, therefore, reflect the presence of such a complex spatial organization of functional responses to written language. In line with this proposal, when previously reported VWFA coordinates are organized based on the type of stimulus contrast, a posterior-anterior differentiation seems to emerge (Figure 4). Specifically, studies that isolated responses to lexical properties of text (e.g., orthographic regularity, frequency) reported activation peaks that cluster in the anterior portion of VOTC ($y = -40$ to -68). On the other hand, in studies that employed comparisons, where the stimuli differed in both lexical and perceptual properties (e.g., words vs checkerboards), the range of reported VWFA coordinates includes the posterior portions of VOTC ($y = -48$ to -85).

This finer-grained differentiation of brain responses to written words has been further confirmed by neuroimaging studies that employed multiple experimental contrasts (isolating more perceptual or more lexical features) within the same design (Dehaene et al. 2004; Vinckier et al. 2007; Taylor et al. 2019). For example, (Dehaene et al. 2004) conducted an fMRI priming study where the location and the orthographic similarity between the prime and the target were manipulated. The results showed that, when moving from posterior to anterior coordinates in left VOTC, there was increased sensitivity to orthographic similarity and a concomitant decreased sensitivity to location similarity. Similar results were observed in Vinckier et al. (2007), where orthographic regularity was gradually manipulated to make the visual inputs progressively more similar to real words (i.e., infrequent letters, frequent letters, frequent bigrams, frequent quadrigrams). Moving towards anterior locations in VOTC was associated with an increased selectivity for word-like stimuli as compared to nonwords. Overall, these findings suggest that while posterior portions of left VOTC represent the perceptual features of written language, anterior portions of left VOTC are sensitive to linguistic aspects of the written input (e.g., frequency, semantic similarity, rhyming properties; Dehaene et al. 2004; Vinckier et al. 2007; Taylor et al. 2019; Wang et al. 2021b; Figure 4). Anterior portions of VOTC might also be involved in higher order stages of visual word recognition where orthography is related to phonology (e.g., audiovisual integration effects of letters and speech sounds have been localized in a group of brain areas that includes anterior portions of VOTC; Macaluso et al. 2004; Blau et al. 2008).

It is worth noting that some recent work has made a binary distinction between posterior and anterior word-selective regions in VOTC i.e., VWFA-1 and VWFA-2 (White et al. 2019b), or pOTS and mOTS (Lerma-Usabiaga et al. 2018), respectively (see Figure 3; for earlier evidence see Hasson et al. 2002; Stigliani et al. 2015). These two discrete word-selective regions have been differentiated not only based on their function but also based on (a) cytoarchitecture and (b) structural connectivity. Specifically, VWFA-1 responds to visual features that define written language and it analyzes this information in parallel when multiple stimuli are presented simultaneously at different locations in the visual field (Lerma-Usabiaga et al. 2018; White et al. 2019b). In contrast, VWFA-2 is sensitive to orthographic and lexical properties of the linguistic input and it appears to process information in a sequential fashion, one word at a time (Lerma-Usabiaga et al. 2018; White et al. 2019b). In addition, VWFA-2 responses seem to be more left-lateralized as compared to VWFA-1, likely reflecting the fact that VWFA-2 is involved in analyzing the linguistic aspects of written language (Hasson et al. 2002; Vinckier et al. 2007; Weiner et al. 2017b; White et al. 2019b).

Two additional pieces of evidence further support the functional differentiation between regions of anterior and posterior word-selective cortex. First, intracranial electrophysiological recordings have shown a temporal dissociation between early orthographic effects (~150-250 ms) and late prelexical/lexical responses (~300-500 ms; Hirshorn et al. 2016; Woolnough et al. 2020), which seems to follow an posterior-anterior spatial topographic distribution (Nobre et al. 1994; McCarthy et al. 1995; Nobre and McCarthy 1995; Lochy et al. 2018; Boring et al. 2020). For example, Thesen and colleagues found that a posterior VOTC region differentiated letters from false-fonts 60ms before a more anterior VOTC region differentiated words from orthographically implausible strings

of consonants (Thesen et al. 2012). This is consistent with classic electrophysiological models of visual word recognition, where evoked responses to low-level perceptual features are temporally localized at an early stage as compared to evoked responses to lexical/semantic aspects of the linguistic stimulus (although the spatial localization of these electrophysiological effects is underspecified; Hauk et al. 2006; Holcomb and Grainger 2006; Barber and Kutas 2007; Grainger and Holcomb 2009). Second, neuroimaging studies examining bottom-up and top-down effects on VOTC have further highlighted the relevance of a posterior-anterior distinction. While bottom-up effects were most likely localized in posterior portions of left VOTC (Kay and Yeatman 2017), top-down effects initially involved more anterior brain regions and moved posteriorly at later stages of word recognition (Heilbron et al. 2020; Woolnough et al. 2020).

Rethinking the functional organization of ventral occipitotemporal cortex

The recent functional characterization of different word-selective patches in VOTC has advanced our understanding of visual word recognition. However, it should be noted that there has been a lack of clarity on the extent to which this posterior-anterior organization corresponds to (a) a gradual change or gradient of functional responses versus (b) a sequence of functionally distinct brain regions. In other words, it is still unclear whether the VWFA should be considered a unitary “area” showing a continuous range of sensitivity from perceptual to linguistic properties (Dehaene et al. 2004; Vinckier et al. 2007), or rather a “collection” of multiple discrete subregions specialized for different types of computations on words (White et al. 2019b, a). It is worth noting that these two possibilities are not mutually exclusive. At the millimeter resolution, fMRI studies on visual object recognition have been able to localize functionally distinct VOTC areas that do not overlap (Grill-Spector and Weiner 2014). However, at the resolution of individual neurons, some regions appear to have relatively homogenous populations of cells while others show substantial heterogeneity (Tsao et al. 2006; Park et al. 2017; Bao et al. 2020). Even within a VOTC region that is highly selective for a visual category (e.g., places) the proportion of neurons showing a category-specific response can vary between 30% and 90% (Tsao et al. 2006; Bell et al. 2011). Hence, while in neuroimaging studies distinguishing between a gradient and discrete areas is still highly informative, when we consider neurons within a single voxel, we might still expect to find gradients or intermingled populations of neurons. One thing that is clear is that distinct regions will appear as a gradient if the spatial resolution of the measurements is not sufficient to distinguish the regions (discussed in more detail below). Progressive improvements in measurement techniques and theories will help us reconcile these different levels of description. Just as models of word recognition have found improved precision over the last century (Figure 1), we can expect more precise models of the VOTC reading circuitry in the years and decades to come.

As reviewed above, recent evidence supports the idea that the VOTC has (at least) two functionally distinct areas that selectively respond to words (Lerma-Usabiaga et al. 2018; White et al. 2019b, a). At first, this new proposition of two separate VWFAs (or VWFA subregions), with a posterior subregion being responsible for processing perceptual aspects of written text and a more anterior subregion enabling automated lexical processing of words, seems to contradict the literature of the past two decades which was largely driven by

the assumption of a single VWFA. However, the wide spatial distribution of word-selective responses has been acknowledged for decades (Cohen et al. 2000, 2002; Hasson et al. 2002), with some suggesting that spatially distinct portions of the VWFA might be attributed to separate functions (Cohen and Dehaene 2004; Dehaene et al. 2004). This proposal is conceptually in line with the studies showing a gradual sensitivity to lexical features on the posterior-anterior axis (Vinckier et al. 2007; Taylor et al. 2019). A closer look at the reported coordinates of the VWFA may suggest that many of the early studies only described one of the two regions, the anterior/middle (classical VWFA; VWFA proper; Cohen et al. 2000) or the posterior VWFA (Ben-Shachar et al. 2011; Rauschecker et al. 2012; Yeatman et al. 2013; Kay and Yeatman 2017; see Figure 4).

If we work from the assumption that there are separate, discrete word-selective regions in VOTC, the appearance of a single region or gradient could arise from the reduced anatomical precision of commonly used analysis approaches. For example, none of the previous studies reporting a spatial gradient of VOTC responses, to our knowledge, performed analysis accounting for individual anatomy (Kronbichler et al. 2004; Binder et al. 2006; Brem et al. 2006; Vinckier et al. 2007). The practice of defining the location of the VWFA based on group analyses in template space, or based on previous studies/samples found in the literature, may have “smeared” the signal coming from distinct areas (Glezer and Riesenhuber 2013) and made it appear as one large word-selective patch in VOTC with gradually increasing sensitivity to lexical features (Figure 3). Moreover, in patient studies it is difficult to achieve a detailed characterization of VOTC segregation, because lesions are rarely as focal and usually affect extensive parts of cortex. Hence, the methodological choices made to localize the VWFA have a great impact on the level of precision with which its spatial organization can be described (Wandell et al. 2012; Glezer and Riesenhuber 2013). Using localizer scans and defining the VWFA within individual anatomy will help us reach a more precise understanding of the sequence of computations that underlie word recognition. As our understanding of VOTC anatomy improves, and the structural pathways that connect to it are described in more detail, we will be able to develop a more complete model that not only describes differences in functional responses but also accounts for anatomical boundaries within VOTC. Additionally, greater emphasis on linking the function and anatomical organization of VOTC (in both humans and non-human primates) will help us better characterize the nature of sub-divisions within word-selective cortex.

White matter connections of ventral occipitotemporal cortex indicate a posterior-anterior distinction

In the last two decades advances in neuroimaging methods have opened the possibility to explore white matter anatomy with a high degree of precision (Wandell 2016; Jeurissen et al. 2019), and relate it to functional responses at the level of individual subjects. Modern diffusion MRI (dMRI) measurements in combination with tractography algorithms make it possible to reliably track structural connections between brain areas (Kruper et al. in press; Yeatman et al. 2012b, 2014; Pestilli et al. 2014; Takemura et al. 2016; Wandell 2016; Jeurissen et al. 2019), providing the basis for a deeper understanding of the nature of functional brain responses and their possible interactions with other structurally

connected neural sources. In the case of VOTC, a detailed structural description of the human ventral stream has revealed major white matter connections to both attention- and language-related brain areas (Yeatman et al. 2013, 2014; Takemura et al. 2015; Weiner et al. 2017b; Lerma-Usabiaga et al. 2018; Chen et al. 2019). The posterior-anterior organization observed in the functional domain (see previous section) is also evident in terms of anatomical connectivity: posterior versus anterior word-selective regions have distinct structural connections. Specifically, recent tractography studies have revealed that posterior portions of VOTC primarily connect to parietal regions involved in the allocation of attention through vertical tracts, namely the left vertical occipital fasciculus (Yeatman et al. 2014; Takemura et al. 2015). On the other hand, mid/anterior portions of the VOTC primarily connect to frontotemporal regions that are essential for language analysis (e.g., Broca's area) through the left arcuate fasciculus (Yeatman et al. 2014; Weiner et al. 2017b; Lerma-Usabiaga et al. 2018). In addition, anterior and posterior portions of the VOTC are interconnected through the inferior longitudinal fasciculus (Catani et al. 2003), whose function is essential for visual object recognition and reading (Yeatman et al. 2012a; Herbet et al. 2018).

This structural distinction is also aligned with new insights coming from studies focused on the cellular architecture of the cortex. Recent cytoarchitectonic measurements have revealed a sequence of regions in VOTC with distinct cellular structure. Specifically, four cytoarchitectonic areas have been identified in the fusiform gyrus and nearby sulci (FG1, FG2, FG3, and FG4, respectively; Caspers et al. 2013; Lorenz et al. 2017). Critically, VWFA-1 and VWFA-2 seem to be localized to cortical areas that have qualitatively different cellular architectures: While functionally-defined VWFA-1 lies within cytoarchitectonically-defined FG2, VWFA-2 is within FG4 (Weiner et al. 2017a). Other similar relationships between functionally defined regions and cytoarchitectonic structures have been observed in human and non-human primates (Zangenehpour and Chaudhuri 2005; Borra et al. 2010; Weiner et al. 2017a).

Bridging functional and structural evidence of the posterior to anterior distinction in VOTC

The distinction between posterior and anterior VOTC has been observed based on multiple types of measurements. Posterior-anterior differences have emerged in: (1) the functional sensitivity of VOTC responses to written language, (2) the time course of VOTC responses, (3) the structural connectivity between VOTC and other brain regions, and (4) the cytoarchitecture of VOTC. In this section we try to bridge the results from these distinct domains and highlight important future research directions.

Based on the findings reviewed above, we propose how information is processed in a sequence of regions within the visual word recognition system (Figure 5). When a written stimulus is presented, posterior portions of VOTC show an early sensitivity to perceptual properties of the orthographic input, and this early visual response can be flexibly modulated based on structural connections with attentional networks (e.g., vertical occipital fasciculus connecting posterior VOTC to the intraparietal sulcus; Yeatman et al. 2014; Takemura et

al. 2015; Weiner et al. 2017a; Lerma-Usabiaga et al. 2018; Chen et al. 2019). At this early temporal stage (~150-250 ms after stimulus onset), the processing of low-level visual features that compose written language is carried out in a parallel fashion (White et al. 2019b). On the other hand, anterior portions of the VOTC show a later sensitivity to higher-order linguistic properties of the visual input (~300-500 ms after stimulus onset), which can be prioritized and further processed through connections with language areas (e.g., arcuate fasciculus connecting anterior VOTC with Broca's area; Yeatman et al. 2014; Weiner et al. 2017a; Lerma-Usabiaga et al. 2018; Chen et al. 2019). At this stage, the visual word form processing seems to proceed in a serial fashion: one word at a time (White et al. 2019b). Hence, this posterior-anterior distinction supports the presence of a functional and structural transition from vision to language. This proposal highlights the need for further supporting evidence and it sheds light on new research directions.

Additional studies showing a double dissociation between anterior and posterior VOTC functional responses are needed to further confirm the transformation from vision to language. In line with what the literature suggests so far, anterior portions of VOTC (i.e., VWFA-2) should be sensitive to linguistic features (e.g., frequency) but not to low-level perceptual properties (e.g., location) of written words, while posterior portions of VOTC (i.e., VWFA-1) should show the reversed response pattern. Anatomical and functional evidence on the role of the ventral reading pathways in extracting information about the structure of written words, suggests that VWFA-2 may also be involved in morphological processing and lexical access (Yablonski et al. 2019). Similarly, the allocation of spatial attention should only affect responses in posterior portions of VOTC, gating the information that arrives at more anterior areas. Despite the recent advancement of our knowledge about how our brain processes text, there are still many questions that remain unanswered. Here, we propose four main points that need further investigation.

First, the degree of alignment between structural and functional properties of VOTC subregions is still unclear and highly dependent on the level of precision of our methods. New cutting edge diffusion-weighted imaging techniques are reaching sub millimeter resolution - a level of precision that was unimaginable a decade ago (Setsompop et al. 2012, 2013; Johansen-Berg and Behrens 2013; Wang et al. 2021a). We expect that a precise anatomical description of white matter tracts that pass through vOTC will enable a more accurate localization of functionally distinct VOTC subregions. Moreover, increasing anatomical precision might result in increased predictive power for functional properties of VWFA-1 and VWFA-2. However, this remains a challenge for future research and the correspondence between structure and function is likely to depend on the spatial scale of the measurements.

Second, functional connectivity patterns of different VOTC subregions have not been defined yet. Evidence based on a single VWFA ROI that is drawn from group analyses suggests that connectivity to frontoparietal, auditory, and visual networks depends on literacy skills (López-Barroso et al. 2020). Defining a single VWFA at the individual level, as the subregion that shows the strongest activation to words compared to other visual stimuli categories, revealed connectivity patterns with Wernicke's area which correlated with literacy skills (Stevens et al. 2017). This effect was not evident when the VWFA was defined

on a group level. This highlights the need for a more precise definition of seed regions that account for individual variability and the involvement of more than one vOTC subregion in reading. A segregated description of how VWFA-1 and VWFA-2 interact with other brain areas could provide new insights on the role of structural white matter connections in reading. Therefore, associating structural and functional connectivity of VOTC subregions represents an important future research direction.

Third, it is still unclear whether the model proposed here can be generalized to different languages. Although key brain areas of the reading circuit are shared across different languages (Rueckl et al. 2015), cross-linguistic differences due to orthographic depth and writing system may impact how text is processed in VOTC (Kumar et al. 2010; Kim et al. 2017; Oliver et al. 2017). An fMRI study suggested that VOTC regions that are sensitive to lexicality show different co-activation patterns depending on language transparency. Second language (L2) readers of a transparent language showed stronger co-activation with regions of the dorsal pathway, i.e. frontoparietal and superior temporal regions, while L2 readers of an opaque language showed higher co-activation of the vOTC seed and regions of the ventral pathway, i.e. more anterior parts of the ventral visual stream and the pars triangularis (Oliver et al. 2017). A similar pattern occurred when word reading was compared in a phonological and logographic script and showed a higher engagement of regions in the dorsal and ventral pathways respectively (Kim et al. 2017). There is need for further investigation to determine whether two (or more) anatomically and functionally distinct VWFAs are universal in expert readers. VWFA-1 is likely to show a similar development across languages with potentially small differences associated with the visual complexity of different scripts. VWFA-2 seems to be more susceptible to effects of orthographic depth and it remains unknown whether readers of transparent languages rely less on the sequential processing from VWFA-1 to VWFA-2 or exhibit the same propagation along VOTC but, potentially, on a faster time scale.

Fourth, the structural and functional developmental trajectories of VOTC subregions are still underspecified. Longitudinal studies have described that VOTC becomes progressively responsive to written words during reading acquisition (Ben-Shachar et al. 2011; Dehaene-Lambertz et al. 2018; Nordt et al. 2021; for earlier cross-sectional evidence see Brem et al. 2009). In addition, functional connectivity patterns may indicate which parts of VOTC become word-selective during development (Saygin et al. 2016). However, longitudinal studies with higher anatomical precision regarding this functional specialization are sorely lacking and the emergence of VWFA-1 and VWFA-2 in normal development has not been described yet. Given the structural and cytoarchitectonic differences observed in VOTC, we can expect that during reading acquisition VWFA-1 and VWFA-2 show different developmental trajectories. Examining prereaders' white-matter pathways that are important for the functional specialization of VWFAs might help us predict individual variability in the location of word-selective VOTC areas and its importance for future reading skills (Grotheer et al. 2021). We expect that while the endpoints of the left vertical occipital fasciculus guide the future location of VWFA-1, the endpoints of the left arcuate fasciculus guide the emergence of VWFA-2. Similarly, functional connectivity patterns observed in the developing brain might help predict the functional specialization of VWFA subareas (O'Rawe et al. 2019; Li et al. 2020).

Conclusion

The progressive refinement of our models of the brain's reading circuitry have resulted in a better understanding of how the brain can extract linguistic information from a visual input (Figure 1). In some ways, our understanding of the brain's reading circuitry has remained relatively consistent across the decades. For example, it was over a century ago when (Déjerine 1891) originally proposed the notion of a visual word form area and early fMRI work was able to document many fundamental properties of VOTC for reading. However, in other ways, the level of precision in current models of the brain's reading circuitry would have been unimaginable in the early days of fMRI. For example, a sequence of specific regions that perform distinct operations can now be reliably localized in an individual's brain, and relative to the white matter connections that communicate information throughout the reading circuitry. Continued progress, and incremental refinement of our models, will depend on integrating structural and functional measurements at finer spatial and temporal scales, and building theories that bridge vision science, psycholinguistics, and cognitive neuroscience.

Acknowledgements

We thank Michal Ben-Shachar for her helpful comments during revision.

Funding

This work was supported by European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 837228 and Rita Levi Montalcini fellowship to SC, NICHD R01HD09586101 and Jacobs Foundation Research Fellowship to JDY, Stanford Maternal and Child Health Research Institute award to IK, and the Zuckerman-CHE STEM Leadership Program to MY.

References

- Baker CI, Liu J, Wald LL, et al. (2007) Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proc Natl Acad Sci U S A* 104:9087–9092. 10.1073/pnas.0703300104 [PubMed: 17502592]
- Bao P, She L, McGill M, Tsao DY (2020) A map of object space in primate inferotemporal cortex. *Nature*, 10.1038/s41586-020-2350-5
- Barber HA, Kutas M (2007) Interplay between computational models and cognitive electrophysiology in visual word recognition. *Brain Res Rev* 53:98–123. 10.1016/j.brainresrev.2006.07.002 [PubMed: 16905196]
- Bell AH, Malecek NJ, Morin EL, et al. (2011) Relationship between functional magnetic resonance imaging-identified regions and neuronal category selectivity. *J Neurosci* 31:12229–12240. 10.1523/JNEUROSCI.5865-10.2011 [PubMed: 21865466]
- Ben-Shachar M, Dougherty RF, Deutsch GK, Wandell BA (2007) Differential sensitivity to words and shapes in ventral occipito-temporal cortex. *Cereb Cortex* 17:1604–1611. 10.1093/cercor/bhl071 [PubMed: 16956978]
- Ben-Shachar M, Dougherty RF, Deutsch GK, Wandell BA (2011) The development of cortical sensitivity to visual word forms. *J Cogn Neurosci* 23:2387–2399. 10.1162/jocn.2011.21615 [PubMed: 21261451]
- Binder JR, Frost JA, Hammeke TA, et al. (1997) Human Brain Language Areas Identified by Functional Magnetic Resonance Imaging. *The Journal of Neuroscience* 17:353–362 [PubMed: 8987760]

- Binder JR, Medler DA, Westbury CF, et al. (2006) Tuning of the human left fusiform gyrus to sublexical orthographic structure. *Neuroimage* 33:739–748. [https://doi.org/S1053-8119\(06\)00707-5](https://doi.org/S1053-8119(06)00707-5)[pii]10.1016/j.neuroimage.2006.06.053 [PubMed: 16956773]
- Blau V, van Atteveldt N, Formisano E, et al. (2008) Task-irrelevant visual letters interact with the processing of speech sounds in heteromodal and unimodal cortex. *European Journal of Neuroscience* 28:500–509
- Boring MJ, Silson EH, Ward MJ, Richardson RM, Fiez JA, Baker CI, & Ghuman AS (2021). Multiple adjoining word-and face-selective regions in ventral temporal cortex exhibit distinct dynamics. *Journal of Neuroscience* 41(29):6314–6327. 10.1523/JNEUROSCI.3234-20.2021
- Borra E, Ichinohe N, Sato T, et al. (2010) Cortical connections to area TE in monkey: hybrid modular and distributed organization. *Cereb Cortex* 20:257–270. 10.1093/cercor/bhp096 [PubMed: 19443621]
- Brem S, Bucher K, Halder P, et al. (2006) Evidence for developmental changes in the visual word processing network beyond adolescence. *Neuroimage* 29:822–837. 10.1016/j.neuroimage.2005.09.023 [PubMed: 16257546]
- Brem S, Halder P, Bucher K, et al. (2009) Tuning of the visual word processing system: distinct developmental ERP and fMRI effects. *Hum Brain Mapp* 30:1833–1844. 10.1002/hbm.20751 [PubMed: 19288464]
- Brem S, Maurer U, Kronbichler M, et al. (2020) Visual word form processing deficits driven by severity of reading impairments in children with developmental dyslexia. *Sci Rep* 10:18728. 10.1038/s41598-020-75111-8 [PubMed: 33127943]
- Bub DN, Arguin M, Lecours AR (1993) Jules Dejerine and his interpretation of pure alexia. *Brain Lang* 45:531–559. 10.1006/brln.1993.1059 [PubMed: 8118672]
- Caspers J, Zilles K, Eickhoff SB, et al. (2013) Cytoarchitectonical analysis and probabilistic mapping of two extrastriate areas of the human posterior fusiform gyrus. *Brain Struct Funct* 218:511–526. 10.1007/s00429-012-0411-8 [PubMed: 22488096]
- Catani M, Howard RJ, Pajevic S, Jones DK (2002) Virtual in vivo interactive dissection of white matter fasciculi in the human brain. *Neuroimage* 17:77–94. 10.1006/nimg.2002.1136 [PubMed: 12482069]
- Catani M, Jones DK, Donato R, Ffytche DH (2003) Occipito-temporal connections in the human brain. *Brain* 126:2093–2107. 10.1093/brain/awg203 [PubMed: 12821517]
- Centanni TM, King FW, Eddy MD, et al. (2017) Development of sensitivity versus specificity for print in the visual word form area. *Brain Fang* 170:62–70. 10.1016/j.bandl.2017.03.009
- Chen L, Wassermann D, Abrams DA, et al. (2019) The visual word form area (VWFA) is part of both language and attention circuitry. *Nat Commun* 10:5601. 10.1038/s41467-019-13634-z [PubMed: 31811149]
- Chyl K, Fraga-González G, Brem S, Jednoróg K (2021) Brain dynamics of (a)typical reading development—a review of longitudinal studies, *npj Science of Learning* 6:1–9. 10.1038/s41539-020-00081-5 [PubMed: 33398003]
- Cohen L, Dehaene S (2004) Specialization within the ventral stream: the case for the visual word form area. *Neuroimage* 22:466–476. 10.1016/j.neuroimage.2003.12.049 [PubMed: 15110040]
- Cohen L, Dehaene S, Naccache L, et al. (2000) The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* 123 (Pt 2):291–307. 10.1093/brain/123.2.291 [PubMed: 10648437]
- Cohen L, Lehericy S, Chochon F, et al. (2002) Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain* 125:1054–1069. 10.1093/brain/awf094 [PubMed: 11960895]
- Dehaene-Lambertz G, Monzalvo K, Dehaene S (2018) The emergence of the visual word form: Longitudinal evolution of category-specific ventral visual areas during reading acquisition. *PLoS Biol* 16:e2004103. 10.1371/journal.pbio.2004103 [PubMed: 29509766]
- Dehaene S, Cohen L (2011) The unique role of the visual word form area in reading. *Trends Cogn Sci* 15:254–262. [https://doi.org/S1364-6613\(11\)00073-8](https://doi.org/S1364-6613(11)00073-8)[pii]10.1016/j.tics.2011.04.003 [PubMed: 21592844]

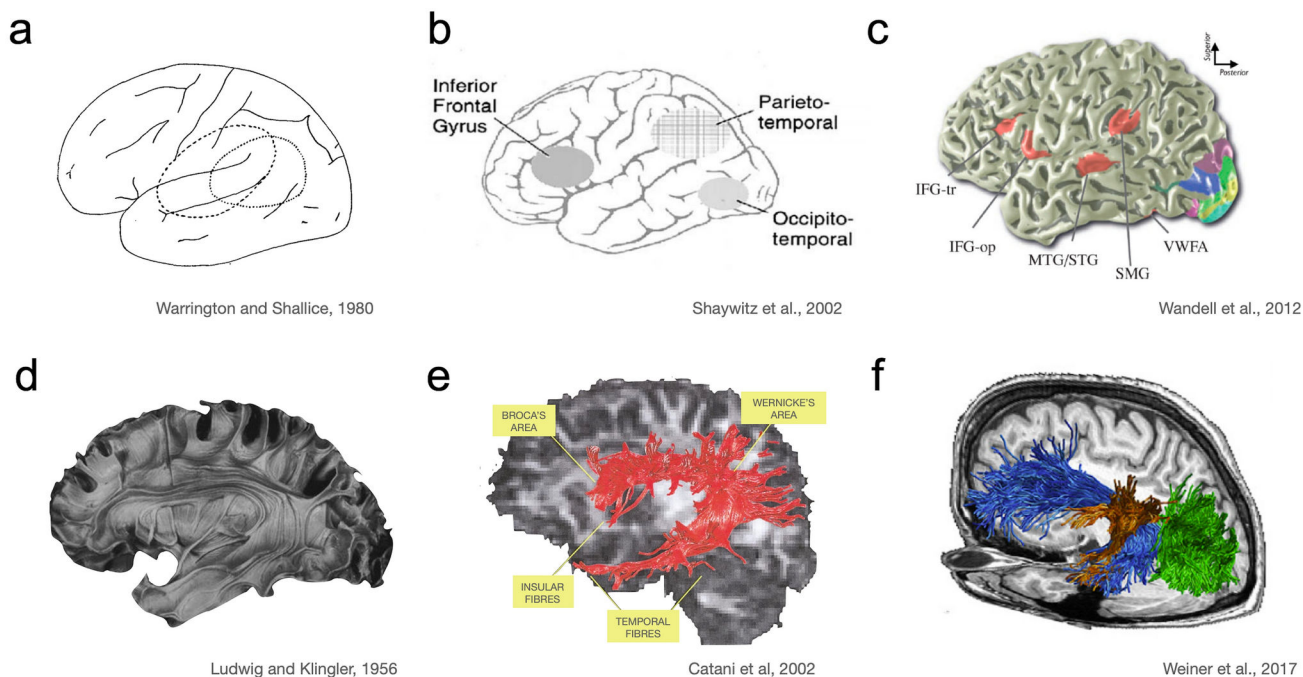
- Dehaene S, Cohen L, Sigman M, Vinckier F (2005) The neural code for written words: a proposal. *Trends Cogn Sci* 9:335–341. 10.1016/j.tics.2005.05.004 [PubMed: 15951224]
- Dehaene S, Jobert A, Naccache L, et al. (2004) Letter binding and invariant recognition of masked words: behavioral and neuroimaging evidence. *Psychol Sci* 15:307–313. 10.1111/j.0956-7976.2004.00674.x [PubMed: 15102139]
- Dehaene S, Le Clecq HG, Poline JB, et al. (2002) The visual word form area: a prelexical representation of visual words in the fusiform gyrus. *Neuroreport* 13:321–325 [PubMed: 11930131]
- Dehaene S, Pegado F, Braga LW, et al. (2010) How learning to read changes the cortical networks for vision and language. *Science* 330:1359–1364. 10.1126/science.1194140 [PubMed: 21071632]
- Déjerine J (1891) Sur un cas de cécité verbale avec agraphie suivi d'autopsie. *Mémoires de la Société de Biologie* 3:197–201
- DiCarlo JJ, Zoccolan D, Rust NC (2012) How Does the Brain Solve Visual Object Recognition? *Neuron Perspective* 73:415–434. 10.1016/j.neuron.2012.01.010
- Downing PE, Jiang Y, Shuman M, Kanwisher N (2001) A cortical area selective for visual processing of the human body. *Science* 293:2470–2473. 10.1126/science.1063414 [PubMed: 11577239]
- Fischl B (2012) FreeSurfer. *Neuroimage* 62:774–781. 10.1016/j.neuroimage.2012.01.021 [PubMed: 22248573]
- Freiwald WA, Tsao DY (2010) Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science* 330:845–851. 10.1126/science.1194908 [PubMed: 21051642]
- Gaillard R, Naccache L, Pinel P, et al. (2006) Direct intracranial, FMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron* 50:191–204. 10.1016/j.neuron.2006.03.031 [PubMed: 16630832]
- Gerrits R, Van der Haegen L, Brysbaert M, Vingerhoets G (2019) Laterality for recognizing written words and faces in the fusiform gyrus covaries with language dominance. *Cortex* 117:196–204 [PubMed: 30986634]
- Glezer LS, Jiang X, Riesenhuber M (2009) Evidence for Highly Selective Neuronal Tuning to Whole Words in the “Visual Word Form Area.” *Neuron* 62:199–204. 10.1016/j.neuron.2009.03.017 [PubMed: 19409265]
- Glezer LS, Riesenhuber M (2013) Individual Variability in Location Impacts Orthographic Selectivity in the “Visual Word Form Area.” *J Neurosci* 33:11221–11226. 10.1523/JNEUROSCI.5002-12.2013 [PubMed: 23825425]
- Grainger J, Holcomb PJ (2009) Watching the Word Go by: On the Time-course of Component Processes in Visual Word Recognition. *Language and Linguistics Compass* 3:128–156 [PubMed: 19750025]
- Graves RE (1997) The legacy of the Wernicke-Lichtheim model. *J Hist Neurosci* 6:3–20. 10.1080/09647049709525682 [PubMed: 11619195]
- Grill-Spector K, Weiner KS (2014) The functional architecture of the ventral temporal cortex and its role in categorization. *Nat Rev Neurosci* 15:536–548. 10.1038/nrn3747 [PubMed: 24962370]
- Gross CG, Rocha-Miranda CE, Bender DB (1972) Visual properties of neurons in inferotemporal cortex of the Macaque. *J Neurophysiol* 35:96–111. 10.1152/jn.1972.35.1.96 [PubMed: 4621506]
- Gross CG, Schiller PH, Wells C, Gerstein GL (1967) Single-unit activity in temporal association cortex of the monkey. *J Neurophysiol* 30:833–843. 10.1152/jn.1967.30.4.833 [PubMed: 4962340]
- Grotheer M, Yeatman J, Grill-Spector K (2021) White matter fascicles and cortical microstructure predict reading-related responses in human ventral temporal cortex. *Neuroimage* 227:117669. 10.1016/j.neuroimage.2020.117669 [PubMed: 33359351]
- Hasson U, Levy I, Behrmann M, et al. (2002) Eccentricity bias as an organizing principle for human high-order object areas. *Neuron* 34:479–490. 10.1016/s0896-6273(02)00662-1 [PubMed: 11988177]
- Hauk O, Davis MH, Ford M, et al. (2006) The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage* 30:1383–1400. 10.1016/j.neuroimage.2005.11.048 [PubMed: 16460964]
- Heeger DJ, Simoncelli EP, Movshon JA (1996) Computational models of cortical visual processing. *Proc Natl Acad Sci U S A* 93:623–627. 10.1073/pnas.93.2.623 [PubMed: 8570605]

- Heilbron M, Richter D, Ekman M, et al. (2020) Word contexts enhance the neural representation of individual letters in early visual cortex. *Nat Commun* 11:321. 10.1038/s41467-019-13996-4 [PubMed: 31949153]
- Herbet G, Zemmoura I, Duffau H (2018) Functional Anatomy of the Inferior Longitudinal Fasciculus: From Historical Reports to Current Hypotheses. *Front Neuroanat* 12:77. 10.3389/fnana.2018.00077 [PubMed: 30283306]
- Hirshorn EA, Li Y, Ward MJ, et al. (2016) Decoding and disrupting left midfusiform gyrus activity during word reading. *Proc Natl Acad Sci U S A* 113:8162–8167. 10.1073/pnas.1604126113 [PubMed: 27325763]
- Holcomb PJ, Grainger J (2006) On the time course of visual word recognition: an event-related potential investigation using masked repetition priming. *J Cogn Neurosci* 18:1631–1643. 10.1162/jocn.2006.18.10.1631 [PubMed: 17014368]
- Hubel DH (1995) Eye, brain, and vision. Scientific American library series, No 22 242:
- Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160:106–154. 10.1113/jphysiol.1962.sp006837 [PubMed: 14449617]
- Hung CP, Kreiman G, Poggio T, DiCarlo JJ (2005) Fast readout of object identity from macaque inferior temporal cortex. *Science* 310:863–866. 10.1126/science.1117593 [PubMed: 16272124]
- Jeurissen B, Descoteaux M, Mori S, Leemans A (2019) Diffusion MRI fiber tractography of the brain. *NMR Biomed* 32:e3785. 10.1002/nbm.3785 [PubMed: 28945294]
- Johansen-Berg H, Behrens TEJ (2013) Diffusion MRI: From Quantitative Measurement to In vivo Neuroanatomy. Academic Press
- Kay KN, Yeatman JD (2017) Bottom-up and top-down computations in word- and face-selective cortex. *Elife* 6.: 10.7554/eLife.22341
- Khaligh-Razavi S-M, Kriegeskorte N (2014) Deep Supervised, but Not Unsupervised, Models May Explain IT Cortical Representation. *PLoS Comput Biol* 10:e1003915. 10.1371/journal.pcbi.1003915 [PubMed: 25375136]
- Kim N, Kim J, Kang C-K, et al. (2017) Human Brain Mapping of Visual Script Familiarity between Phonological and Logographic Language: 3 T Functional MRI Study. *Biomed Res Int* 2017:5732642. 10.1155/2017/5732642 [PubMed: 28758118]
- Kim T, Bair W, Pasupathy A (2019) Neural Coding for Shape and Texture in Macaque Area V4. *J Neurosci* 39:4760–4774. 10.1523/JNEUROSCI.3073-18.2019 [PubMed: 30948478]
- Kronbichler M, Hutzler F, Wimmer H, et al. (2004) The visual word form area and the frequency with which words are encountered: evidence from a parametric fMRI study. *Neuroimage* 21:946–953. 10.1016/j.neuroimage.2003.10.021 [PubMed: 15006661]
- Krupar J, Yeatman JD, Richie-Halford A, Bloom D, Grotheer M, Caffarra S, Kiar G, Karipidis II, Roy E, Chandio BQ, Garyfallidis E, and Rokem A (in press). Evaluating the reliability of human brain white matter tractometry. *Aperture*.
- Kumar U, Das T, Bapi RS, et al. (2010) Reading different orthographies: an fMRI study of phrase reading in Hindi-English bilinguals. *Read Writ* 23:239–255. 10.1007/s11145-009-9176-8
- Jerma-Usabiaga G, Carreiras M, Paz-Alonso PM (2018) Converging evidence for functional and structural segregation within the left ventral occipitotemporal cortex in reading. *Proc Natl Acad Sci U S A* 115:E9981–E9990. 10.1073/pnas.1803003115 [PubMed: 30224475]
- Li J, Osher DE, Hansen HA et al. Innate connectivity patterns drive the development of the visual word form area. *Sci Rep* 10, 18039 (2020). 10.1038/s41598-020-75015-7 [PubMed: 33093478]
- Liu J, Harris A, Kanwisher N (2010) Perception of face parts and face configurations: an FMRI study. *J Cogn Neurosci* 22:203–211. 10.1162/jocn.2009.21203 [PubMed: 19302006]
- Lochy A, Jacques C, Maillard L, et al. (2018) Selective visual representation of letters and words in the left ventral occipito-temporal cortex with intracerebral recordings. *Proc Natl Acad Sci U S A* 115:E7595–E7604. 10.1073/pnas.1718987115 [PubMed: 30038000]
- López-Barroso D, Thiebaut de Schotten M, Morais J, et al. (2020) Impact of literacy on the functional connectivity of vision and language related networks. *Neuroimage* 213:116722. 10.1016/j.neuroimage.2020.116722 [PubMed: 32156625]
- Lorenz S, Weiner KS, Caspers J, et al. (2017) Two New Cytoarchitectonic Areas on the Human Mid-Fusiform Gyrus. *Cereb Cortex* 27:373–385. 10.1093/cercor/bhv225 [PubMed: 26464475]

- Ludwig E, Klingler J (1956) Atlas Cerebri Humani: Der innere Bau des Gehirns dargestellt auf Grund makroskopischer Präparate. Karger S, Basel.
- Macaluso E, George N, Dolan R, et al. (2004) Spatial and temporal factors during processing of audiovisual speech: a PET study. *Neuroimage* 21:725–732. 10.1016/j.neuroimage.2003.09.049 [PubMed: 14980575]
- Malach R, Levy I, Hasson U (2002) The topography of high-order human object areas. *Trends Cogn Sci* 6:176–184. 10.1016/s1364-6613(02)01870-3 [PubMed: 11912041]
- Martin A, Schurz M, Kronbichler M, Richlan F (2015) Reading in the brain of children and adults: a meta-analysis of 40 functional magnetic resonance imaging studies. *Hum Brain Mapp* 36:1963–1981. 10.1002/hbm.22749 [PubMed: 25628041]
- McCandliss BD, Cohen L, Dehaene S (2003) The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn Sci* 7:293–299. 10.1016/s1364-6613(03)00134-7 [PubMed: 12860187]
- McCarthy G, Nobre AC, Bentin S, Spencer DD (1995) Language-related field potentials in the anterior-medial temporal lobe: I. Intracranial distribution and neural generators. *J Neurosci* 15:1080–1089 [PubMed: 7869084]
- Nandy AS, Sharpee TO, Reynolds JH, Mitchell JF (2013) The fine structure of shape tuning in area V4. *Neuron* 78:1102–1115. 10.1016/j.neuron.2013.04.016 [PubMed: 23791199]
- Nobre AC, Allison T, McCarthy G (1994) Word recognition in the human inferior temporal lobe. *Nature* 372:260–263. 10.1038/372260a0 [PubMed: 7969469]
- Nobre AC, McCarthy G (1995) Language-related field potentials in the anterior-medial temporal lobe: II. Effects of word type and semantic priming. *J Neurosci* 15:1090–1098 [PubMed: 7869085]
- Nordt M, Gomez J, Natu VS, et al. (2021) Cortical recycling in high-level visual cortex during childhood development. *Nat Hum Behav*. 10.1038/s41562-021-01141-5
- Oliver M, Carreiras M, Paz-Alonso PM (2017) Functional Dynamics of Dorsal and Ventral Reading Networks in Bilinguals. *Cereb Cortex* 27:5431–5443. 10.1093/cercor/bhw310 [PubMed: 28122808]
- O'Rawe JF, Huang AS, Klein DN, & Leung HC (2019). Posterior parietal influences on visual network specialization during development: An fMRI study of functional connectivity in children ages 9 to 12. *Neuropsychologia*, 127, 158–170. 10.1016/j.neuropsychologia.2019.03.001 [PubMed: 30849407]
- Park SH, Russ BE, McMahon DBT, et al. (2017) Functional Subpopulations of Neurons in a Macaque Face Patch Revealed by Single-Unit fMRI Mapping. *Neuron* 95:971–981.e5. 10.1016/j.neuron.2017.07.014 [PubMed: 28757306]
- Pestilli F, Yeatman JD, Rokem A, et al. (2014) Evaluation and statistical inference for human connectomes. *Nat Methods* 11:1058–1063. 10.1038/nmeth.3098 [PubMed: 25194848]
- Pleisch G, Karipidis II, Brem A, et al. (2019) Simultaneous EEG and fMRI reveals stronger sensitivity to orthographic strings in the left occipito-temporal cortex of typical versus poor beginning readers. *Dev Cogn Neurosci* 40:100717. 10.1016/j.dcn.2019.100717 [PubMed: 31704655]
- Poeppl D, Mangun GR, Gazzaniga MS (2020) *The Cognitive Neurosciences*, sixth edition. MIT Press
- Pugh KR, Shaywitz BA, Shaywitz SE, et al. (1996) Cerebral organization of component processes in reading. *Brain* 119 (Pt 4):1221–1238 [PubMed: 8813285]
- Rauschecker AM, Bowen RF, Parvizi J, Wandell BA (2012) Position sensitivity in the visual word form area. *Proc Natl Acad Sci U S A* 109:E1568–77. 10.1073/pnas.1121304109 [PubMed: 22570498]
- Riesenhuber M, Poggio T (1999) Hierarchical models of object recognition in cortex. *Nat Neurosci* 2:1019–1025. 10.1038/14819 [PubMed: 10526343]
- Rossion B, Joyce CA, Cottrell GW, Tarr MJ (2003) Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage* 20:1609–1624. 10.1016/j.neuroimage.2003.07.010 [PubMed: 14642472]
- Rueckl JG, Paz-Alonso PM, Molfese PJ, et al. (2015) Universal brain signature of proficient reading: Evidence from four contrasting languages. *Proc Natl Acad Sci U S A* 112:15510–15515. 10.1073/pnas.1509321112 [PubMed: 26621710]

- Rust NC, Dicarlo JJ (2010) Selectivity and tolerance (“invariance”) both increase as visual information propagates from cortical area V4 to IT. *J Neurosci* 30:12978–12995. 10.1523/JNEUROSCI.0179-10.2010 [PubMed: 20881116]
- Saygin ZM, Osher DE, Koldewyn K, et al. (2012) Anatomical connectivity patterns predict face selectivity in the fusiform gyrus. *Nat Neurosci* 15:321–327. 10.1038/nn.3001
- Saygin ZM, Osher DE, Norton ES, et al. (2016) Connectivity precedes function in the development of the visual word form area. *Nature Neuroscience* 19:1250–1255 [PubMed: 27500407]
- Setsompop K, Cohen-Adad J, Gagoski BA, et al. (2012) Improving diffusion MRI using simultaneous multi-slice echo planar imaging. *Neuroimage* 63:569–580. 10.1016/j.neuroimage.2012.06.033 [PubMed: 22732564]
- Setsompop K, Kimmlingen R, Eberlein E, et al. (2013) Pushing the limits of in vivo diffusion MRI for the Human Connectome Project. *Neuroimage* 80:220–233. 10.1016/j.neuroimage.2013.05.078 [PubMed: 23707579]
- Shaywitz BA, Shaywitz SE, Pugh KR, et al. (2002) Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biol Psychiatry* 52:101–110. 10.1016/S0006-3223(02)01365-3 [PubMed: 12114001]
- Skeide MA, Kumar U, Mishra RK, et al. (2017) Learning to read alters cortico-subcortical cross-talk in the visual system of illiterates. *Sci Adv* 3:e1602612. 10.1126/sciadv.1602612 [PubMed: 28560333]
- Srihasam K, Mandeville JB, Morocz IA, et al. (2012) Behavioral and anatomical consequences of early versus late symbol training in macaques. *Neuron* 73:608–619. 10.1016/j.neuron.2011.12.022 [PubMed: 22325210]
- Srihasam K, Vincent JL, Livingstone MS (2014) Novel domain formation reveals proto-architecture in inferotemporal cortex. *Nat Neurosci* 17:1776–1783. 10.1038/nn.3855 [PubMed: 25362472]
- Stevens WD, Kravitz DJ, Peng CS, et al. (2017) Privileged Functional Connectivity between the Visual Word Form Area and the Language System. *J Neurosci* 37:5288–5297. 10.1523/JNEUROSCI.0138-17.2017 [PubMed: 28450544]
- Stigliani A, Weiner KS, Grill-Spector K (2015) Temporal Processing Capacity in High-Level Visualm Cortex Is Domain Specific. *J Neurosci* 35:12412–12424. 10.1523/JNEUROSCI.4822-14.2015 [PubMed: 26354910]
- Takemura H, Caiafa CF, Wandell BA, Pestilli F (2016) Ensemble Tractography. *PLoS Comput Biol* 12:1–22. 10.1371/journal.pcbi.1004692
- Takemura H, Rokem A, Winawer J, et al. (2015) A Major Human White Matter Pathway Between Dorsal and Ventral Visual Cortex. *Cereb Cortex* 26:2205–2214. 10.1093/cercor/bhv064 [PubMed: 25828567]
- Tanaka K, Saito H, Fukada Y, Moriya M (1991) Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J Neurophysiol* 66:170–189. 10.1152/jn.1991.66.1.170 [PubMed: 1919665]
- Taylor JSH, Davis MH, Rastle K (2019) Mapping visual symbols onto spoken language along the ventral visual stream. *Proceedings of the National Academy of Sciences* 116:201818575. 10.1073/pnas.1818575116
- Thesen T, McDonald CR, Carlson C, et al. (2012) Sequential then interactive processing of letters and words in the left fusiform gyrus. *Nat Commun* 3:1284. 10.1038/ncomms2220 [PubMed: 23250414]
- Tsao DY, Freiwald WA, Tootell RBH, Livingstone MS (2006) A cortical region consisting entirely of face-selective cells. *Science* 311:670–674. 10.1126/science.1119983 [PubMed: 16456083]
- Van der Haegen L, Cai Q, Brysbaert M (2012) Colateralization of Broca’s area and the visual word form area in left-handers: fMRI evidence. *Brain Lang* 122:171–178. 10.1016/j.bandl.2011.11.004 [PubMed: 22196742]
- van der Mark S, Bucher K, Maurer U, et al. (2009) Children with dyslexia lack multiple specializations along the visual word-form (VWF) system. *Neuroimage* 47:1940–1949. 10.1016/j.neuroimage.2009.05.021 [PubMed: 19446640]
- van der Mark S, Klaver P, Bucher K, et al. (2011) The left occipitotemporal system in reading: disruption of focal fMRI connectivity to left inferior frontal and inferior parietal language

- areas in children with dyslexia. *Neuroimage* 54:2426–2436. 10.1016/j.neuroimage.2010.10.002 [PubMed: 20934519]
- Vinckier F, Dehaene S, Jobert A, et al. (2007) Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron* 55:143–156. [https://doi.org/S0896-6273\(07\)00450-3](https://doi.org/S0896-6273(07)00450-3)[pii]10.1016/j.neuron.2007.05.031 [PubMed: 17610823]
- Wandell BA (2016) Clarifying Human White Matter. *Annual Review of Neuroscience* 39:103–128
- Wandell BA, Rauschecker AM, Yeatman JD (2012) Learning to see words. *Annu Rev Psychol* 63:31–53. 10.1146/annurev-psych-120710-100434 [PubMed: 21801018]
- Wang F, Dong Z, Tian Q, et al. (2021a) In vivo human whole-brain Connectom diffusion MRI dataset at 760 μm isotropic resolution. *Sci Data* 8:122. 10.1038/s41597-021-00904-z [PubMed: 33927203]
- Wang J, Joanisse MF, Booth JR (2021b) Letter fluency in 7-8-year-old children is related to the anterior, but not posterior, ventral occipito-temporal cortex during an auditory phonological task. *Dev Cogn Neurosci* 47:100898. 10.1016/j.dcn.2020.100898 [PubMed: 33341533]
- Warrington EK, Shallice T (1980) Word-form dyslexia. *Brain* 103:99–112. 10.1093/brain/103.1.99 [PubMed: 6244876]
- Weiner KS, Barnett MA, Lorenz S, et al. (2017a) The Cytoarchitecture of Domain-specific Regions in Human High-level Visual Cortex. *Cereb Cortex* 27:146–161. 10.1093/cercor/bhw361 [PubMed: 27909003]
- Weiner KS, Grill-Spector K (2010) Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *Neuroimage* 52:1559–1573. [https://doi.org/S1053-8119\(10\)00689-0](https://doi.org/S1053-8119(10)00689-0)[pii]10.1016/j.neuroimage.2010.04.262 [PubMed: 20457261]
- Weiner KS, Yeatman JD, Wandell BA (2017b) The posterior arcuate fasciculus and the vertical occipital fasciculus. *Cortex* 97:274–276. 10.1016/j.cortex.2016.03.012 [PubMed: 27132243]
- White AL, Boynton GM, Yeatman JD (2019a) You Can't Recognize Two Words Simultaneously. *Trends Cogn. Sci* 23:812–814 [PubMed: 31477387]
- White AL, Palmer J, Boynton GM, Yeatman JD (2019b) Parallel spatial channels converge at a bottleneck in anterior word-selective cortex. *Proc Natl Acad Sci U S A* 116:10087–10096. 10.1073/pnas.1822137116 [PubMed: 30962384]
- Woolnough O, Donos C, Rollo PS, et al. (2020) Spatiotemporal dynamics of orthographic and lexical processing in the ventral visual pathway. *Nat Hum Behav.* 10.1038/s41562-020-00982-w
- Yablonski M, Rastle K, Taylor JSH, Ben-Shachar M (2019) Structural properties of the ventral reading pathways are associated with morphological processing in adult English readers. *Cortex* 116:268–285. 10.1016/j.cortex.2018.06.011 [PubMed: 30037636]
- Yeatman JD, Dougherty RF, Ben-Shachar M, Wandell BA (2012a) Development of white matter and reading skills. *Proc Natl Acad Sci U S A* 109:E3045–53. 10.1073/pnas.1206792109 [PubMed: 23045658]
- Yeatman JD, Dougherty RF, Myall NJ, et al. (2012b) Tract profiles of white matter properties: automating fiber-tract quantification. *PLoS One* 7:e49790. 10.1371/journal.pone.0049790 [PubMed: 23166771]
- Yeatman JD, Rauschecker AM, Wandell BA (2013) Anatomy of the visual word form area: adjacent cortical circuits and long-range white matter connections. *Brain Lang* 125:146–155. 10.1016/j.bandl.2012.04.010 [PubMed: 22632810]
- Yeatman JD, Weiner KS, Pestilli F, et al. (2014) The vertical occipital fasciculus: a century of controversy resolved by in vivo measurements. *Proc Natl Acad Sci U S A* 111:E5214–23. 10.1073/pnas.1418503111 [PubMed: 25404310]
- Yeatman JD, White AL (2021) Reading: The Confluence of Vision and Language. *Annu Rev Vis Sci.* 10.1146/annurev-vision-093019-113509
- Zangenehpour S, Chaudhuri A (2005) Patchy organization and asymmetric distribution of the neural correlates of face processing in monkey inferotemporal cortex. *Curr Biol* 15:993–1005. 10.1016/j.cub.2005.04.031 [PubMed: 15936269]

**Fig. 1.**

History of the reading circuitry. **a** Neurologists at the turn of the 20th century debated the location of the brain's visual word form center, which was initially placed in the angular gyrus (Déjerine 1891). This debate played out through the early days of PET until a part of VOTC was discovered as the location of the “visual word form area”. The location was first confirmed by neurologists (Warrington and Shallice 1980). This debate was at the spatial scale of lobes. **b** Within the first decade of fMRI the three main components of the reading circuitry were defined as left inferior frontal, inferior parietal and occipitotemporal cortex (Pugh et al. 1996; Shaywitz et al. 2002). This early model continues to be influential and outlines the circuitry at the spatial scale of ~4cm (general locations within a lobe). **c** Over the next decade a more nuanced understanding of these regions emerged. Regions were precisely defined relative to sulcal landmarks at the millimeter spatial scale. Language-related regions are in red, while the other colors are used to illustrate retinotopic maps in visual cortex (Wandell et al. 2012). **d, e, f** A similar historical progression of anatomical precision for white matter pathways. **d** Ludwig and Klingler's brain model provides a representation of the white matter tracts. Using Klingler's method of dissecting the human brain after freezing it, these models revealed that axonal connections form large bundles, or fascicles (Ludwig and Klingler 1956). **e** Early diffusion tensor imaging based tractography showing a major white matter pathway connecting regions involved in language processing that is also crucial for reading. These in-vivo white matter tract reconstructions corroborated previous post mortem anatomical findings (Catani et al. 2002). **f** Recent advances in diffusion MRI allow for fine-grained representation of white matter pathways in relation to functionally defined regions within individual brains (Yeatman et al. 2012a; Takemura et al. 2015; Weiner et al. 2017b). The current state of the art has led to predictions of functional responses in individual brains, with millimeter precision, based on diffusion MRI measures of an individual's white matter anatomy (Saygin et al. 2012; Grotheer et al.

2021). In each row there is a gradual increase of precision and spatial resolution, which became possible with the accumulation of knowledge from different modalities and the concomitant improvement in imaging technologies. Our current understanding builds upon these previous models and several aspects of the reading circuit are still being investigated in this rapidly evolving field of research. Copyrights: **a** Warrington EK, Shallice T. Word-form dyslexia. *Brain* 1980, Vol 103, 99–112, by permission of Oxford University Press. **b** Reprinted from *Biological Psychiatry*, Vol 52, Shaywitz BA, Shaywitz SE, Pugh KR et al., Disruption of posterior brain systems for reading in children with developmental dyslexia, 101-110, Copyright (2002), with permission from Elsevier. **c** Modified with permission from the *Annual Review of Psychology*, Volume 63 © 2012 by Annual Reviews, <http://www.annualreviews.org> **e** Reprinted from *NeuroImage*, Vol 17, Catani M, Howard RJ, Pajevic S, Jones DK, Virtual in vivo interactive dissection of white matter fasciculi in the human brain, 77–94, Copyright (2002), with permission from Elsevier. **f** Reprinted from *Cortex*, Vol 97, Weiner KS, Yeatman JD, Wandell BA, The posterior arcuate fasciculus and the vertical occipital fasciculus, 274-276, Copyright (2017), with permission from Elsevier.

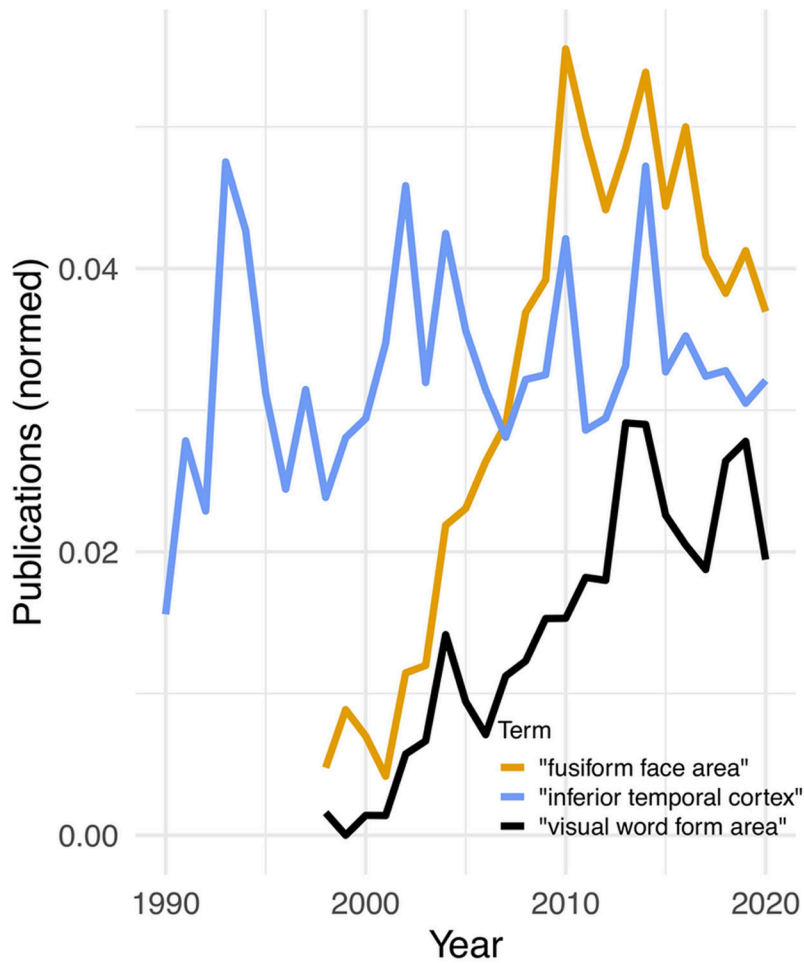


Fig. 2. Use of the term “visual word form area” in publications. Relative number of publications mentioning the terms “visual word form area”, from 1990 to 2020 are shown in solid black. As a comparison, the number of publications including the term “fusiform face area” are shown in yellow, and “inferior temporal cortex” in blue. To account for the overall increase of publications, the number of publications was normalized to the total number of publications mentioning the term “visual cortex”. While topics like inferior temporal cortex show a relatively stable number of publications over the last 30 years, the terms “visual word form area” and “fusiform face area” first appeared in 1998 and the ratio of publications using these terms has increased since the early 2000s. Source: pubmed.gov

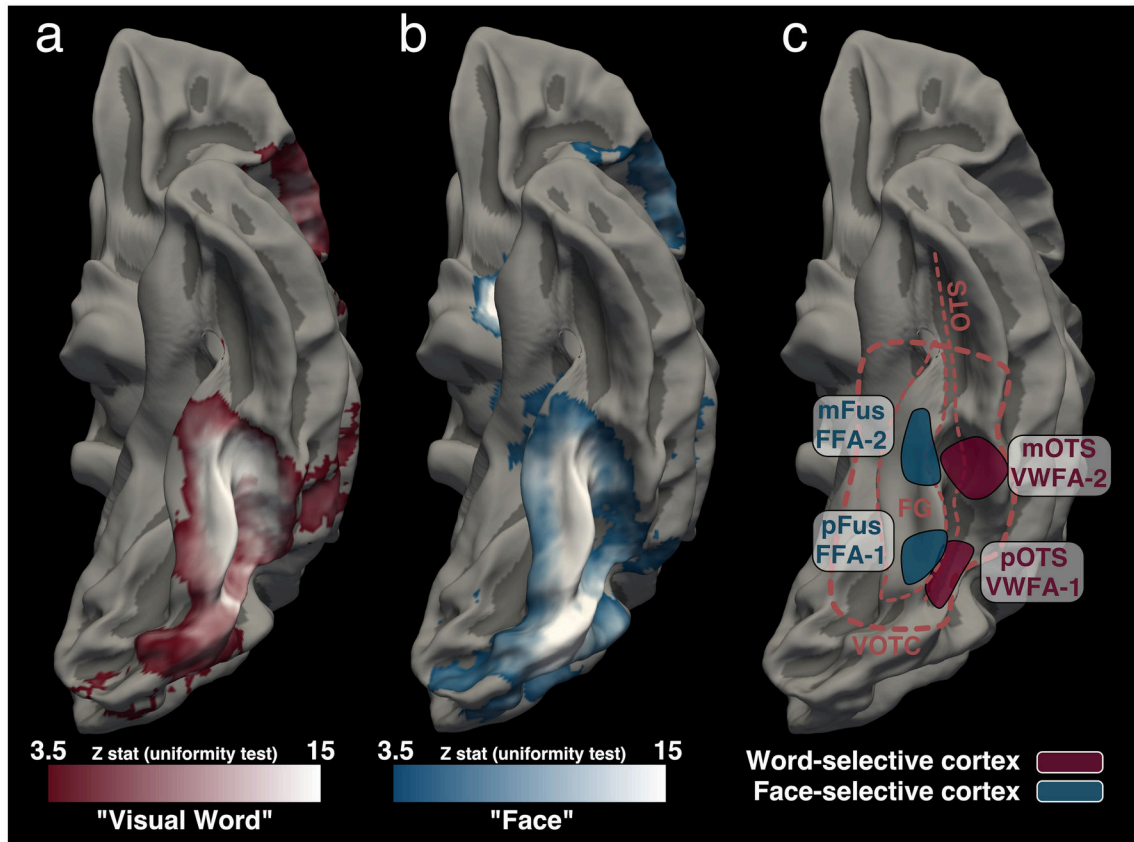


Fig. 3.

Word- and face-selective cortex (a,b) A meta-analysis conducted with Neurosynth of MNI coordinates associated with “Visual Word” (panel a) and “Face” (panel b). Data from (a) 117 and (b) 896 fMRI studies are shown on a group average template. When data are smoothed and analyzed on the MNI template a large portion of VOTC appears to be word-selective. In template space, word-selective and face-selective regions overlap covering much of the fusiform gyrus (FG) and the occipitotemporal sulcus (OTS). This type of group analysis obscures the anatomical organization of word-selective and face-selective cortex. (c) A schematic of the typical layout of word-selective and face-selective regions defined for an individual subject and shown on the fsaverage cortical surface. Word-selective regions can be reliably localized on an individual’s cortical surface and are spatially distinct from face-selective regions. Moreover, a sequence of at least 2 separate word- and face-selective regions can be identified within VOTC (Yeatman and White 2021). Anatomical boundaries of VOTC are shown as a thick dashed line and the FG and OTS are marked as thin dashed lines. Different naming schemes have been used to label these regions. Some groups use an anatomy-function labeling scheme where each region is named based on its anatomical location - posterior fusiform (pFus), mid fusiform (mFus), posterior OTS (pOTS), mid OTS (mOTS) - followed by its functional selectivity (e.g., words; faces) (Grill-Spector and Weiner 2014; Stigliani et al. 2015). Other groups have chosen to keep the widely used terms of “Fusiform Face Area (FFA)” and “Visual Word Form Area (VWFA)” and designated the subregions as FFA-1, FFA-2, VWFA-1, VWFA-2 (White et al. 2019b, a; Yeatman and White 2021). These two naming conventions refer to the same sub-regions and the catch-all

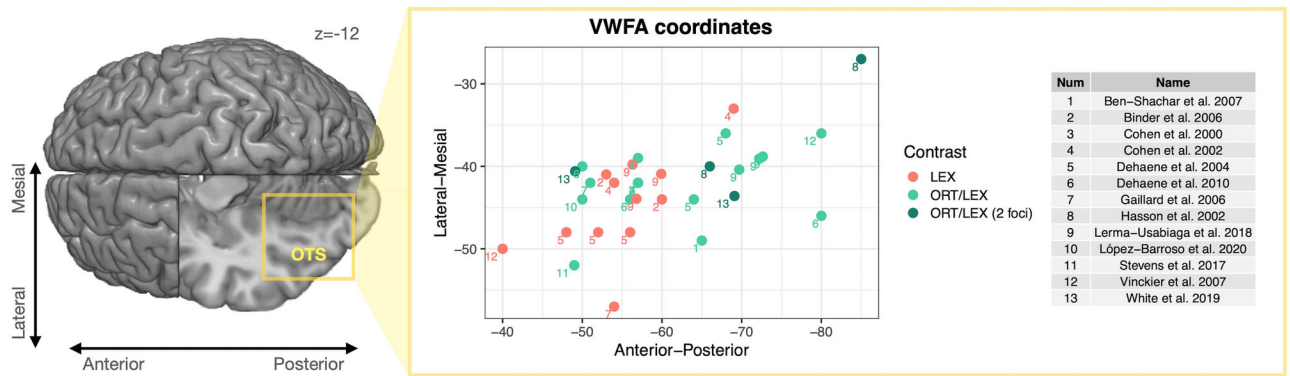
terms “word-selective cortex” and “face-selective cortex” are used to refer to the collection of cortical regions that are selective for a given category.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Fig. 4.**

Left: Top view of MNI152 brain template with a cutout over the left occipital temporal sulcus (OTS) at the level of $z = -12$. Right: Spatial representations of VWFA coordinates reported in a sample of fMRI papers examining different types of stimulus contrasts in adults (Ben-Shachar et al. 2007; Stevens et al. 2017; López-Barroso et al. 2020). Contrasts were categorized as “lexical” when different types of text stimuli were compared (LEX e.g., words vs letter strings, words vs false fonts, words vs pseudowords; as in Lerma-Usabiaga et al. 2018). Contrasts were considered “orthographic/lexical” (i.e., text-selective) when text stimuli were compared with non-text controls (ORT/LEX e.g., letters vs rest, words vs symbol strings, words vs chequerboards, words vs. scrambled word images). While the VWFA coordinates of lexical contrasts cluster in more anterior portions of VOTC, the coordinates of orthographic/lexical contrasts cover a wider spatial range and extend further posterior. This spatial variability might be associated with the processing of both lexical and perceptual features that are relevant when comparing text with non-linguistic material. Interestingly, two of the abovementioned studies that used an “orthographic/lexical” contrast (marked in dark green) specifically reported two distinct VWFA foci that could be distinguished on the posterior-anterior axis (Hasson et al. 2002; White et al. 2019b). Additional VWFA coordinates can be added to the table available at https://github.com/SendyCaffarra/VWFA_coordinates/tree/main and the figure can be reproduced and updated using the available code.

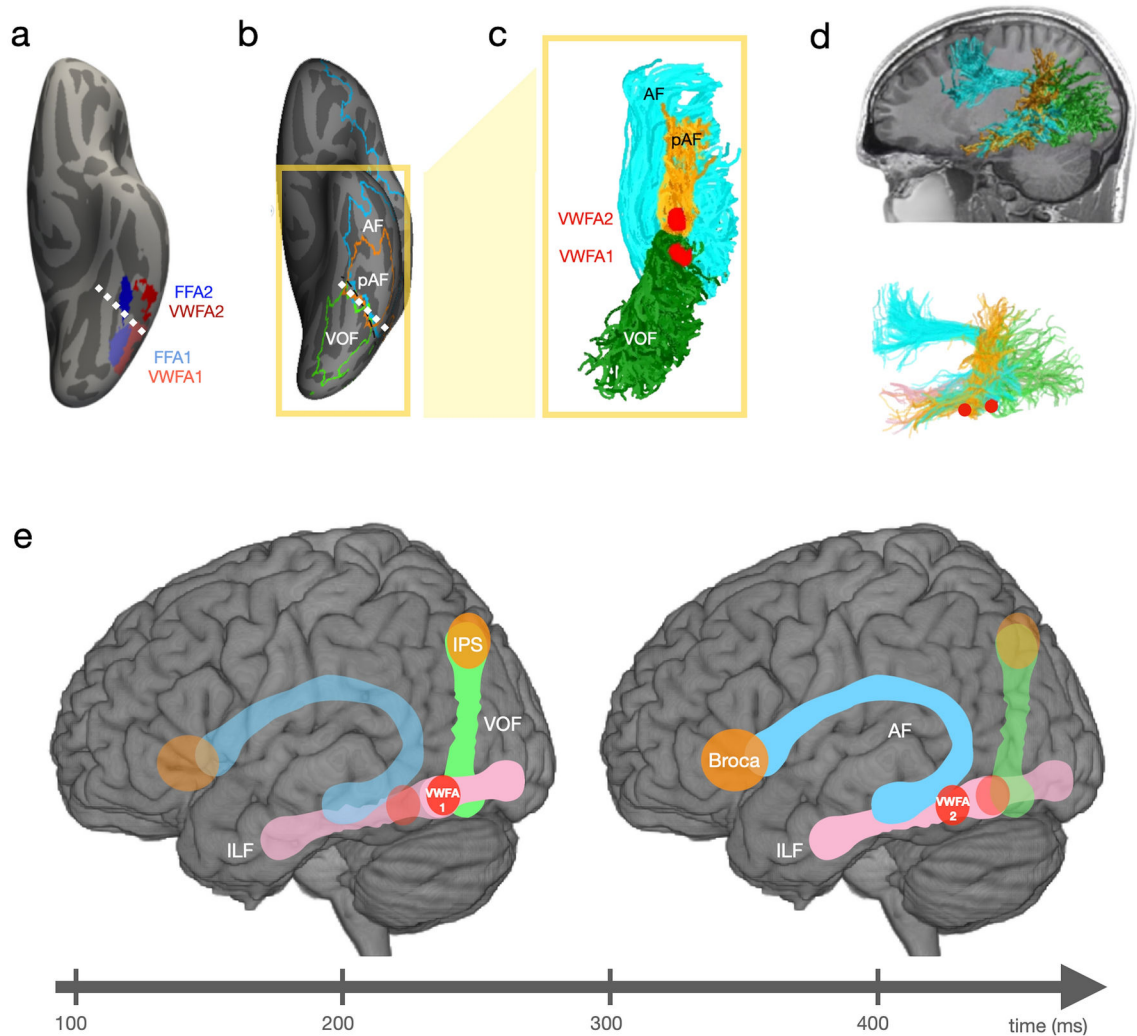


Fig. 5. Structure of word-selective VOTC. **a.** Localization of VWFA-1, VWFA-2, FFA-1 and FFA-2 on the cortical surface. Unpublished data from a sample of adult typical readers performing a category recognition experiment. Colored patches denote the location of maximal overlap of individually identified regions transformed to the fsaverage cortical surface with surface-based alignment (Fischl 2012). **b.** Cortical endpoints of the left arcuate fasciculus (AF), the posterior arcuate fasciculus (pAF) and the vertical occipital fasciculus (VOF, from Weiner et al. 2017b). **c.** and **d.** Ventral and sagittal view of VWFA-1, VWFA-2, AF, pAF, and VOF. Data are from an example adult participant. In d, the upper plot shows AF, pAF and VOF superimposed on the participant's T1 weighted image. The lower plot in panel d shows the AF, pAF, VOF, the inferior longitudinal fasciculus (ILF), VWFA-1 and VWFA-2. **e.** Schematic representation of VWFA subdivisions and anatomical connections subserving early (150-250ms) and late (300-500ms) word recognition processes.