Simultaneous cortical responses to multiple written words

Vassiki Chauhan[1], Krystal McCook[1], Mariam Latif[1] & Alex White[1]
[1]Department of Neuroscience & Behavior, Barnard College, Columbia University, New York, New York 10027

ABSTRACT

We have learned much about the brain regions that support reading by measuring neuronal responses to single words, but we know little about how the brain processes multiple words simultaneously. This fMRI study fills that gap by varying the number of English words presented while holding the amount of visual stimulation constant. We adapted the "simultaneous suppression" paradigm, which has demonstrated that the response to multiple stimuli presented simultaneously is typically smaller than the sum of responses to the same stimuli presented sequentially. On each trial, participants viewed rapid sequences of three frames. Each frame contained two character strings, most of which were pseudo-letters with visual features matched to familiar letters. The experimental conditions differed in the number of English words in the sequence: zero words; one word; two words sequentially; or two words simultaneously. Behaviorally, participants performed worse at detecting two words presented simultaneously than sequentially. In the brain, the BOLD response increased linearly with the number of words presented in several reading-related regions of the left hemisphere: the inferior frontal sulcus, the intraparietal sulcus, the superior temporal sulcus, and text-selective occipito-temporal regions. But the mean response magnitudes did not differ between sequential and simultaneous presentation of two words. Nonetheless, the sensitivity of ventral temporal text-selective regions to two words' lexical frequencies was attenuated by simultaneous presentation. This suggests that words were processed less deeply when they competed simultaneously. Our new paradigm, therefore, reveals processing capacity limits in the reading circuitry.

Key words: visual word recognition; visual word form area; VWFA; simultaneous suppression; parallel processing

INTRODUCTION

To extract meaning from written text, a reader must integrate information from many visual elements that are presented all at once. Those elements are arranged hierarchically: the contours within letters, the letters within words, and the words within sentences. For example, the letter R is composed of two line segments and a curved contour, which readers experienced with the Latin alphabet can readily integrate into a single familiar shape. One level up in the hierarchy, each word contains multiple letters, in a particular order, that determine the word's identity. For skilled readers, the time required to process a familiar word increases only modestly with the number of letters in the word (Barton et al., 2014). Many theorists of word recognition therefore posit that the letters within each word are processed in parallel (e.g., Adelman et al., 2010; Grainger et al., 2016).

This efficient processing of words depends, in part, on a specialized region of the left ventral occipito-temporal cortex: the "visual word form area" (**VWFA**). The VWFA is nestled amongst other regions that are involved in visual recognition and that respond preferentially to specific stimulus categories such as faces, bodies, and scenes (Grill-Spector & Weiner, 2014). The VWFA's preference for written words develops as children learn to read (Ben-Shachar et al., 2011; Dehaene et al., 2015; Dehaene-Lambertz et al., 2018; Nordt et al., 2021; Schlaggar & McCandliss, 2007), and its selectivity is weak in dyslexia (Brem et al., 2020; Kubota et al., 2019; Richlan et al., 2009). In healthy sighted readers, the VWFA is generally believed to represent each written word in a format abstracted from its lower-level visual features, and to communicate with language areas in order further process phonological, syntactic and semantic properties (Caffarra et al., 2021; Dehaene & Cohen, 2011; Yeatman & White, 2021).

Localized damage to the VWFA can cause 'pure alexia': a selective reading deficit that often manifests as slow, letter-by-letter reading (Barton et al., 2014; Gaillard et al., 2006; Pflugshaupt et al., 2009). In word recognition tasks, patients with this condition show a marked increase in response times as the number of letters increases, much more so than healthy readers. Thus, the VWFA plays a necessary role in the ability to process the letters within each word holistically.

Processing capacity limits for multiple word recognition

Less is known about the next level up in the hierarchy: how does the brain process multiple words that are presented together? This question is important because readers typically see many

words at once during natural reading. The vast majority of neuroimaging studies of visual word recognition, however, have presented one word at a time. This choice facilitates experimental control and avoids the complications of eye movements that readers naturally make across connected text. However, we know little about how the VWFA responds as a function of the number of words being read, nor whether its processing of each word is affected by neighboring words. That is the gap the present study aims to fill.

In terms of reading behavior, there has been extensive research on a related question: do readers process multiple words in parallel, analogous to how they process letters within words in parallel? Some computational models of reading assume that such parallel processing can occur, with attention distributed over multiple words (Engbert et al., 2005; Reilly & Radach, 2006; Snell et al., 2018). Proponents of those models cite data from several experimental tasks as well as natural reading behavior (reviewed by Snell & Grainger, 2019). There is disagreement about the strength of evidence for parallel processing, however (e.g., Hossain & White, 2023; Milledge et al., 2023; Reichle et al., 2009; White, Boynton, et al., 2019).

Other studies have suggested that two words cannot be recognized simultaneously. One line of experimental research measured the effects dividing attention between two words on recognition accuracy, as compared to when attention was focused on just one word. The results supported a serial processing model: participants could fully process only one word per trial (Campbell et al., 2024; White et al., 2018, 2020). For consistent results with different paradigms, see Brothers (2022) and Johnson et al. (2022). The notion of a “bottleneck” in the reading circuitry has been supported by recordings of VWFA activity when attention is divided between two words (White, Palmer, et al., 2019). That was one of the few fMRI studies to present more than one word at a time (described in detail in the Discussion)

Another behavioral technique to assess processing capacity limits is to compare perceptual performance when multiple stimuli are presented simultaneously as compared to one at a time. The simultaneous presentation of multiple stimuli typically impairs perceptual task performance (Duncan, 1980; Huang & Pashler, 2005; Kleiss & Lane, 1986). This is true for many types of tasks, including for word recognition (Scharff et al., 2011). In a similar vein, the *redundant target* paradigm measures how quickly participants respond “yes” to seeing *any* stimuli that belong to a target category. For word recognition, subjects are slower to respond to two target words than to a single target word. That finding is consistent with a processing capacity limit for semantic recognition, perhaps even requiring words to be processed serially (Mullin & Egeth, 1989; White et al., 2025).

Simultaneous suppression

The neural basis of processing capacity has also been studied by comparing cortical responses to visual stimuli presented sequentially versus simultaneously. *Simultaneous suppression* is the reduced neural response to a stimulus caused by the presence of other stimuli at the same time in the same receptive field. It has been reported both with single-recordings in visual cortex (Miller et al., 1993) and with fMRI (e.g., Kastner et al., 1998; Kastner & Ungerleider, 2001; McMains & Kastner, 2011). Simultaneous suppression provides support for the theory that multiple stimuli compete for perceptual processing resources (Beck & Kastner, 2009; Desimone & Duncan, 1995; Reynolds et al., 1999).

With fMRI, this phenomenon is observed by comparing the BOLD response to multiple stimuli presented simultaneously, at nearby visual field positions, and the response to the same stimuli at the same positions presented sequentially, in rapid succession. The logic is that because of sluggish hemodynamics, the BOLD response integrates the individual responses to the sequential stimuli that are presented within less than a second during a single trial. The typical finding is that the BOLD response to simultaneous stimuli is lower than the response to sequential stimuli. The magnitude of that “suppression” effect varies across retinotopic visual areas, and as a function of parameters such as stimulus size, duration, and eccentricity (Kastner & Ungerleider, 2001; Kupers et al., 2024). Simultaneous suppression has also been found in the fusiform face area, parahippocampal place area, and lateral occipital complex (Kaiser et al., 2014; Shim et al., 2013).

Goals of the present study

We take inspiration from that past literature to investigate how the reading circuitry’s activity varies as a function of the number of words being read. The first goal is to simply assess whether the response increases as the number of words in a display increases from 0 to 1 to 2. The second goal is to assess whether two words can be processed in the same way when presented simultaneously as when presented sequentially.

The present study is the first to investigate whether simultaneous suppression affects the neuronal representation of written words as well as behavioral word recognition accuracy. To do so, we made changes to the classical simultaneous suppression paradigm. Primarily, we distinctly kept visual stimulation constant across all conditions, in terms of the number of onsets and offsets, and the size, number and complexity of shapes. In past studies, the interpretation of simultaneous suppression is complicated because the conditions differ in the total number of stimulus onsets

and offsets. The ‘suppression’ effect could be due not to competitive sensory interactions within spatial receptive fields, but rather to a reduced total response due to the fewer number of onset and offset transients in the simultaneous condition (Kupers et al., 2024).

We kept visual stimulation constant with the following experimental design: each trial was composed of a sequence of 3 images, or “frames.” Each frame contained two 5-character strings, one above and one below the central fixation mark. Each character string was either an English word written in Courier New or a string of characters in the BACS-2 false font. Each BACS-2 character matches a Courier New character in size, number of strokes, symmetry, number of junctions and number of terminations (Vidal et al., 2017). Across conditions, we varied the number of words in each frame, without changing the number of stimulus offsets and offsets, the number of characters or their basic visual features.

Thus, this design aims to isolate the processing of familiar letter strings in cortical regions that are specialized for them. If we find that such a region responds more strongly on 2-word than 1-word trials, we can conclude that it detected two stimuli of its preferred category, and that the effect is not simply the result of increased visual stimulation. If we find a higher response to 2-sequential than 2-simultaneous trials, we would conclude that there are competitive interactions between the processing of two words.

There are three other key features of our experimental design that differ from prior studies of simultaneous suppression. First, most prior studies presented 4 stimuli together, either all in the same visual field quadrant (Kastner, Weerd, et al., 2001) or one in each quadrant (Shim et al., 2013). Here, we present at most two words on each trial. This is because of past findings that the capacity limits for word recognition might be so strong that they prevent even just two words from being processed in parallel (White, Boynton, et al., 2019).

Second, we present the words above and below the point of gaze fixation, at an eccentricity of 1.5 degrees of visual angle. This is because behavioral evidence for serial processing of words is strong with that arrangement (White et al., 2020), and because both words are equally near to fixation and clearly legible. Arranging the words horizontally would better match the conditions of natural reading, but there is a strong right>left hemifield asymmetry (Yeatman & White, 2021) that would complicate the analyses. Crucially, the VWFA is known to respond strongly to words in a region spanning roughly the central 5 degrees of the visual field (R. K. Le et al., 2017; White et al., 2023), including just above and below fixation (Rauschecker

et al., 2012). Thus, the population receptive fields (PRFs) of the voxels we analyze are known to cover both of the stimulus positions.

Third, our stimuli were task-relevant and attended. Many past studies of simultaneous suppression required the participant to perform a task that diverts attention from the primary stimuli (e.g. Kastner et al., 2001). But there is also evidence for simultaneous suppression in visual cortex when the stimuli are attended (Scalf et al., 2011). In this study, the task was to count the number of real words that were presented on each trial, while maintaining central fixation. We chose to make the stimuli task-relevant because much of the VWFA's word-selective processing is dependent on top-down engagement (Chauhan et al., 2024; Mano et al., 2013; White et al., 2023).

METHODS

Participants

We recruited 22 participants (5 self-identified as male) from the Barnard College and Columbia University student body, and general New York City Community (N=4). As described below, one participant had to be excluded from the analysis due to excessive head motion, leaving a final sample of 21 participants. Their ages ranged from 19 to 40 years (mean = 24.4±6.3 years). All but one were right-handed. All participants had learned English before the age of five, and provided written informed consent. They were monetarily compensated for their participation. The study was approved by the Internal Review Board at Barnard College.

The sample size was chosen in advance of data collection. In our two previous fMRI studies of BOLD responses to written words with similar protocols (Chauhan et al., 2024; White et al., 2023), we included 15 and 17 participants, respectively. In this study, we explore a novel effect (simultaneous suppression for words) which we expected might be smaller than the task effects detected in those prior studies. We therefore set out to collect as much data as we had funding for, which allowed full data sets for 22 participants.

The sample size is also justified by a power analysis of the theoretical simultaneous suppression effect in the VWFA, using data from Chauhan et al. (2024). In that study, we estimated the left VWFA's response magnitude to single words and false font strings. We assumed that in this new study, the response to single words would have the same mean (0.32 percent signal change) and standard deviation (0.2) as in the previous study. (The new data matched that prediction quite closely.) We then assumed that the response in the two-word

sequential condition would be equal to the 1-word mean plus the difference between the response to single words and false font strings in Chauhan et al. (2024). This is because both conditions include responses to false font strings. According to a serial processing model, the response to the two-word *simultaneous* condition would be the same as in the one-word condition. That would be the case if a “serial bottleneck” allows only one of the two simultaneous words to be processed, as some of our previous work concluded (White, Palmer, et al., 2019). We then simulated responses in the 2-simultaneous and 2-sequential conditions, using those predicted means and standard deviations. To reject the null hypothesis for the difference between those two conditions (i.e., the simultaneous suppression effect) with 90% power, 18 participants would be sufficient. Rather than stop after that number of participants, we planned in advance to collect as much data as we could, which left us with 22 participants, with 21 subjects in all reported analyses using our exclusion criteria for head motion.

Equipment

We acquired MRI data at the Zuckerman Institute, Columbia University. We used a 3T Siemens Prisma scanner and a 64-channel head coil. At the start of each scanning session, we acquired a T1-weighted structural scan, with 1mm isometric voxels. We collected functional data with T2* echo planar imaging sequences with multiband echo sequencing (SMS3) for whole brain coverage without compromising temporal resolution. The TR was 1.5s, TE was 30 ms, the flip angle was 62° and the voxel size was 2 mm isotropic.

We presented the stimuli using custom code written in MATLAB and the Psychophysics Toolbox (refs). We projected stimuli onto a blank screen that the participants viewed through a mirror, with a viewing distance of 142 cm. The resolution of the display was 1920 by 1080 pixels, and the refresh rate was 60 Hz. We also collected eye-tracking data with an SR Research Eyelink 1000 tracker. To respond to the tasks, the participant used their right hand to press one of three buttons on a handheld MR-safe response pad.

Stimuli

The stimuli and task for the main experiment are depicted in **Figure 1A**. A small black fixation dot was continuously present at the center of the screen against a light background (95% of screen maximum). The dots diameter was 0.11 degrees of visual angle (dva). The stimuli were five-letter English words. We selected 768 such words with no constraints on parts of speech. On the Zipf scale, the lexical frequencies ranged from 2.2 to 6.1, with a median of 4.1. Zipf is a

standardized measure of word frequency calculated through an adjustment of the log frequency per million words in the SUBTLEX-US corpus (Brysbaert & New, 2009; van Heuven et al., 2014).

Real words were presented in the Courier New, a standard monospaced serif font. The other “filler” strings were presented in an unfamiliar, illegible false font called BACS-2 (Vidal et al., 2017). These false font characters are designed to match the size, the symmetry, and the numbers of strokes, junctions and terminations of Courier New. Both these fonts were scaled so that the height of the letter ‘x’ was 0.41 dva. On average, the width of each character string was 2.65 dva, and the height was between 0.41 and 0.79 dva. The distance between centers of neighboring letters was 0.55 dva. The stimuli were rendered in the same black color as the fixation dot.

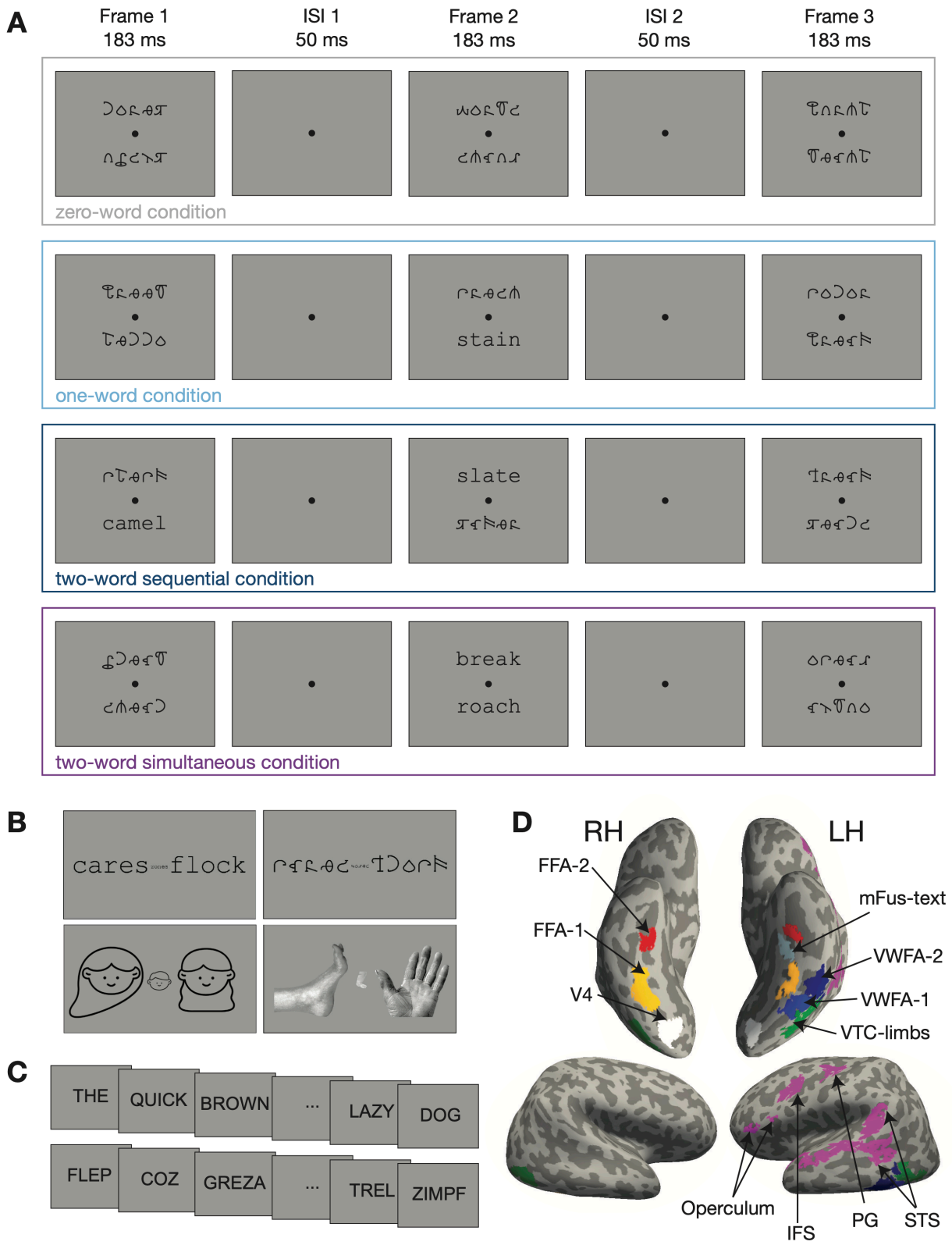

Figure 1: Example stimuli and regions of interest (ROIs). (**A**) Examples of single trials of each presentation condition in the main experiment. The four conditions differed in whether words appeared in the first and/or second frame. Frame 3 always contained two false font strings. The timing and location of

words was counterbalanced within each condition (e.g., on one-word trials the word could be in either frame 1 or 2, on the top or bottom). After frame 3, the participant had 3.65 s to make their response. **(B)** Examples of the four stimulus categories used in the visual category-selective region localizer (faces, limbs, words and false fonts). **(C)** Example stimuli from the language localizer. Each real sentence consisted of 12 real English words, and each jabberwocky sentence consisted of 12 pseudowords. Each (pseudo)word was presented for 450 ms followed by a 50 ms blank. **(D)** Example ROIs in a representative subject's inflated cortical surfaces. The regions on the ventral view (top two images) were defined from the visual category localizer. The lavender regions on the lateral view (bottom images) were defined from the language localizer. Abbreviations: FFA: fusiform face area; VWFA: visual word form area; mFus-Text: mid-fusiform text-selective area; VTC-Limbs: limb area in the ventral temporal surface, extending laterally to include the extrastriate body area (EBA) in some subjects; STS: superior temporal sulcus; PG: Precentral Gyrus; IFS: inferior frontal sulcus. The white V4 region was obtained from a previous study (White et al., 2023).

Trial sequence and task

The stimulus presentation on each trial consisted of three successive frames. Two character strings were presented in each frame, one above and one below the fixation dot. The centers of the character strings were 1.5 dva from the fixation dot. Each frame lasted for 183 ms, and there was a 50 ms inter-stimulus interval between frames (with only the fixation dot presented).

Across the three frames in each trial, there were always 6 character strings, and all of those that were not legible words were in fact English words drawn from the same set, but presented in the illegible false font. The number of words presented in each trial varied across four conditions, which were randomly intermixed. Examples of each condition are illustrated in **Figure 1A**. They are:

1) Zero-word trials: all the character strings were false fonts.
2) One-word trials: one real English word in Courier New was presented in either the first or second frame, either above or below fixation (equally likely).
3) Two-word sequential trials: one word was presented at one location in the first frame, and a different word was presented in the other location in the second frame. The first word was equally likely to be at the top or bottom location.
4) Two-word simultaneous trials: two words appeared at the same time, one above and one below the fixation, in either the first or the second frame (equally likely).

The third frame always contained two false font strings and served as a post-mask. After the third frame, there was a 3651 ms inter-trial interval in which the participant could press a button with their right hand to make their response. Their task was to report the number of words that had appeared in that trial (0, 1 or 2). The distinction between sequential and simultaneous presentation of two words was not relevant to the task. Participants used their ring finger to report zero words, middle finger for one word and index finger for two words. They were instructed to

prioritize accuracy over speed. Participants were also instructed to read the words silently in their head.

After every block of five trials, there was a blank period, the duration of which was randomly and uniformly selected to be 4, 6 or 8 s. There were 64 trials in each run, presented in random order.

Each legible word was presented exactly once to each participant during the experiment. Orthographic neighborhood size and log lexical frequency estimated from the English Lexicon Project were balanced across all experimental conditions in order to minimize differences in task difficulty (Balota et al., 2007). We created 6 sets of word lists that balanced these statistics across the four conditions. One list was randomly chosen for each participant.

The main experiment consisted of 8 five-minute fMRI runs. In total, there were 128 trials of each of the four main conditions. Each run began with written instructions, reminding the participants which fingers to use to make their responses, to respond as accurately as possible, to maintain central gaze fixation, and to read each word silently aloud in their head. There was no feedback at the end of each trial, but the percent correct was displayed once the run ended, along with the percentage of trials in which the participants failed to make a response on time.

Prior to scanning, participants practiced the task in a testing room outside the scanner. During practice, they used the same fingers to report the number of words as during scanning, and their gaze position was tracked. On these practice trials they received auditory feedback: a high-pitched tone for the correct response and a low-pitched tone for the incorrect response. They performed at least 32 practice trials, and repeated the block if their accuracy was below 80%. They also viewed the stimuli for the localizer experiments during the practice session.

Functional Localizers

Each participant completed two functional localizers: one to define category-selective ventral temporal visual regions, and one to localize core language regions. The visual category localizer is described in White et al. (2023) and Chauhan et al. (2024) (see also https://github.com/Barnard-Vision-Lab/catLoc). Briefly, participants viewed images of four categories of objects: text (including English words and pseudowords), false font strings, faces, and limbs. See examples in **Figure 1B**. These categories were chosen to maximize contrast between activity in the VWFA and surrounding face- and body-selective regions. Each frame included three examples of each category, one at fixation and two larger ones, one to either side. We used two fonts for the text, Courier New and Sloan, and two false fonts with matched visual features, BACS-2 (Vidal et al., 2017) and PseudoSloan (Vildavski et al., 2021). Each 4-second

trial consisted of four frames with images of all the same category. Participants performed a one-back repetition detection task, pushing a button with their index finger whenever they saw the same exact image repeated two times in a row. Repetitions occurred on 33% of trials. Each participant completed four runs, each of which consisted of 54 trials and lasted for about six minutes. Each category was presented 56 times.

The second localizer was the language localizer scan, example trials of which are shown in **Figure 1C**. It was designed to define language processing centers in the frontal lobe and superior temporal lobe. See Fedorenko et al. (2010) for details. Briefly, the stimuli consisted of written words presented one at a time, 450 ms per word, 50 ms interstimulus interval between individual words. In the sentence condition, a sequence of 12 real English words formed a meaningful and grammatical sentence. In the “nonsense” condition, a sequence of 12 pronounceable pseudowords formed a meaningless “Jabberwocky” sentence. After every set of 12 character strings, an icon of a hand appeared. Participants’ task was simply to press a button with their index finger when they saw that hand, to ensure that they were paying attention to the stimuli. Participants performed two runs of this task, each of which consisted of sixteen 6-s trials. Each run also included five blank trials which also lasted 6 s.

Regions of Interest defined using both of these localizers are depicted in **Figure 1D**.

Study Procedure

For 18 out of the 22 subjects, the entire study required three sessions. The first session involved training the participants to perform the tasks and calibrate the eye-tracker. The other two sessions were conducted in the MRI scanner and both included the same sequence of types of scans. First, while we acquired a T1 structural scan, participants performed a few practice trials. Then they performed 4 runs of the main experiment, 2 runs of the visual category localizer and 1 run of the language localizer. The other 4 participants had already completed the localizer scans in a previous study, so they only performed 8 runs of the main experiment during a single MRI session.

Analysis of Behavior

We defined accuracy as the proportion of trials with correct responses made within the response period, which lasted 3.65 s after the offset of the final stimulus frame. On average,

participants failed to respond on time on 1.3% of trials (SEM=0.6%). Those trials were excluded from the analysis of accuracy.

Preprocessing

We pre-processed our MRI data with *fMRIPrep* 21.0.1 (Esteban et al., 2019), which is based on *Nipype* 1.6.1 (Gorgolewski et al., 2011). Depending on the number of sessions, we obtained one or two T1-weighted structural scans for each individual. These structural images were skull-stripped, corrected for intensity non-uniformity, and averaged across sessions for participants who participated in multiple sessions. Cortical surfaces were constructed using the boundaries between gray matter and white matter using Freesurfer.

The functional MRI scans were preprocessed as follows: Magnetic field inhomogeneities in each EPI sequence were corrected using a B_0 nonuniformity map collected during the same session. All functional runs were co-registered to the native anatomical volume (*fsnative*) We then slice-time corrected the multiband sequences and resampled them onto each participant's cortical surface.

We also used fMRIPrep to perform motion correction. Moreover, in order to ensure that our data would not be contaminated by head motion, we used the following exclusion criteria. A run with movement greater than 2 voxels (4 mm) was excluded from analysis. Additionally, if any participant had such excessive head motion in more than half of the runs, that participant was dropped from the analysis. Using these criteria, we excluded one participant from the analysis entirely, leaving a final N of 21. A total of three runs were excluded for head motion from 2 of the included participants.

BOLD Response Estimation

We estimated BOLD responses (beta weights) to each single trial with GLMSingle in Python (Prince et al., 2022). The result of this analysis is a beta weight for each trial, for each cortical surface node in each participant. We first resampled the BOLD time-series from 1.5s to 1s. GLMSingle optimizes the hemodynamic response function for each node in the surface space and performs denoising using fractional ridge regression to remove correlated noise. For the main experiment, the design matrix consisted of a unique column for each stimulus condition (zero, one, two-sequential or two-simultaneous words), uniquely coding trials in which the top word appeared in the first or second frame, and uniquely coding the frame in which the words in the 1-word or 2-simultaneous conditions appeared. For the visual category localizer, each stimulus

category received its own column. For the language localizer, there were two columns: sentence trials and nonsense trials.

ROI Definition

We sought to define nine regions of interest (ROIs) in each hemisphere in each participant. See **Figure 1D** for these ROIs on an example participant's inflated cortical surfaces. Details about the number of participants in which we were able to localize each ROI are provided in Table 1. Six of the ROIs in the ventral temporal cortex were obtained from the visual category localizer scans. For each category (words, faces and limbs), we computed t-statistics for the difference in BOLD responses to that category with all other categories (including false fonts). The t-values were projected onto each participant's inflated cortical surfaces, and thresholded at $t \geq 3$.

We defined up to three text-selective areas in each hemisphere using the following anatomical and functional constraints. VWFA-1 and VWFA-2 were in the occipitotemporal sulcus. VWFA-1 was defined as being anterior to hV4. VWFA-2 was defined as a separate peak in the contrast anterior to VWFA-1. In some subjects, there was a continuous patch that we segmented into VWFA-1 and -2. The third text-selective region was mFus-text, in the middle of the fusiform gyrus or mid-fusiform sulcus, medial to and often anterior of VWFA-2. The face-selective areas, FFA-1 and FFA-2, were always medial to the VWFAs and located on the fusiform gyrus. The limb-selective area was lateral to the VWFAs, and in some subjects extended into posterior inferior temporal sulcus/middle temporal gyrus (the location of the putative extrastriate body area", Downing et al., 2001)

We also used the language localizer to define four core language regions with the contrast of sentences vs nonsense sequences. These language ROIs were in the precentral gyrus (PG), the inferior frontal sulcus (IFS), the frontal operculum, and the superior temporal sulcus (STS).

Lastly, we obtained one bilateral retinotopic visual region, hV4, from a previous study that used retinotopic mapping (White et al., 2023). This average-subject ROI was projected from the *fsaverage* template into each participant's native surface space.

Statistical Analysis of Behavior and ROI Responses

We used linear mixed effects modelling to analyze both behavioral and brain responses on single trials. These models included fixed effects for experimental conditions and random slopes and intercepts for subjects. For behavior, we built a generalized linear model with a binomial linking function to predict single-trial accuracy.

To analyze responses within each ROI, we used two different linear mixed effects models to predict the beta weight on each trial. The first model tested whether the BOLD response

increased from 0 to 1 to 2-word trials. We excluded the two-word simultaneous condition because a goal of this first analysis was to verify that the BOLD magnitude represented an integration of the response to two sequential words. The second model tested the simultaneous suppression effect, only for the two-word trials. The fixed effect was experimental condition, simultaneous or sequential presentation.

Our primary ROI is the left VWFA, given extensive past research that it plays a necessary and unique role in word recognition. We applied the same analyses to all the other visual and language ROIs described above, and for these we corrected the p-values of each test for a false discovery rate of 0.05 (Benjamini & Hochberg, 1995). The statistics in the text and the tables report these corrected p-values.

Whole-cortex Analyses

We performed two whole-cortex analyses within each participant. The first was the "number of words effect": a linear slope fit to beta weights as a function of the number of words presented in the 0, 1-word, and 2-words sequential conditions. The second contrast was of the "simultaneous suppression" effect: the difference in beta weights between two-word sequential trials and two-word simultaneous trials. Therefore, for the number of words effect, each node was assigned a slope, and for the simultaneous suppression contrast, each node was assigned the difference in response to sequential and simultaneous trials. After computing each statistic in each subject's native surface space, we spatially smoothed the data with a 5 mm full-width at half-maximum kernel, and projected the smoothed data onto *fsaverage* space using the *mri_surf2surf* function from Freesurfer. We then concatenated all participants' contrasts in *fsaverage* space and computed the mean at each node. Next, we performed a one-sample t-test to determine the nodes in which the value was consistently different from zero. Finally, we corrected the associated p-values for multiple comparisons using FDR correction, implemented as *fdrcorrection* in the Python library *statsmodels.* We then saved the t-values for nodes with correct $p<0.05$.

RESULTS

The experiment included four main conditions: zero-word trials, one-word trials, two-word sequential trials and two-word simultaneous trials (**Figure 1A**). The design allowed us to test two hypotheses: 1) That the response magnitude in reading-related regions increases with the number of items of their preferred category (i.e., words), while the amount of visual stimulation is held constant. 2) That the response to two simultaneously presented words is lower than the total

response to two sequentially presented words, due to processing capacity limits. This second hypothesis is based on earlier studies of simultaneous suppression in visual cortex (e.g., Kastner, Weerd, et al., 2001; Kupers et al., 2024).

Behavioral accuracy shows a cost of simultaneous presentation of two words

First, we calculated response accuracy in each experimental condition (**Figure 2A**). We focus on accuracy because the experimental design was not optimized to detect effects in response times (RTs). Participants were instructed to emphasize accuracy over speed, and comparing RTs across the sequential and simultaneous presentation conditions is difficult because the critical stimuli are presented over different amounts of time.

According to a linear mixed-effect model of single trials, accuracy differed across presentation conditions ($F(1, 10635)=8.72$, $p=0.003$). A paired t-test revealed that participants were less accurate on two-word simultaneous trials than on two-word sequential trials: mean ± standard error proportion correct = 0.87±0.03 vs. 0.95±0.01, respectively ($t(20)=2.98$, $p=0.007$, BF=6.53). Of all the two-word simultaneous trials in which participants made an error, they mostly reported seeing one word (94% of total incorrect trials) as compared to zero words (6% of incorrect trials). Thus, it was more difficult to detect two words when they were presented simultaneously than sequentially, which suggests a processing capacity limit (Scharff et al., 2011b).

The VWFA response increases linearly with the number of words but does not show simultaneous suppression

Our primary region of interest is the left visual word form area (VWFA), a region in the occipito-temporal sulcus that has been shown to play a critical role in reading (Dehaene et al., 2015; Dehaene & Cohen, 2011). We expected it to be one of the few cortical regions that distinguishes words from false font strings, and therefore sensitive to the number of words embedded in the sequences of false fonts. Moreover, prior research has also implicated the left VWFA as a "bottleneck" that prevents simultaneous processing of two words, so it is the prime candidate for demonstrating simultaneous suppression in this task (White, Palmer, et al., 2019). Therefore, we began by analyzing single-trial BOLD responses in the left VWFA, averaging over the sub-regions VWFA-1 and VWFA-2, which had indistinguishable response patterns. The mean responses are shown in **Figure 2B.**

We found a significant effect of experimental condition on VWFA responses (linear mixed effect model, $F(1, 10234)=57.249$, $p<0.001$). The VWFA's response cleared increased with the

number of words on each trial, even as the amount of visual stimulation was held constant. The estimate of the slope of that effect, obtained from the linear model, was 0.11 psc/word and significantly greater than 0 ($t(19)=8.69$, $p<0.001$). In pairwise comparisons between conditions, we observed an increase in the VWFA response from the zero-word to one-word condition ($t(19)=6.85$, $p<0.001$, BF=12182.83), as well as from the one-word to two-word sequential condition ($t(19)=5.43$, $p<0.001$, BF=804.59) (**Figure 2B, right subplot**). The sizes of the stepwise increases from 0 to 1 word and from 1 word to 2-sequential were not significantly different from each other (mean difference of differences = 0.02±0.02 p.s.c., CI=[-0.04, 0.03] $t(19)=0.7$, $p=0.42$, BF=0.29). That suggests that there is a steady linear increase in the response magnitude across trials with 0, 1 and 2 sequential words.

However, we did not find an effect of simultaneous suppression in the VWFA. Mean response magnitudes in the two-word simultaneous condition were only slightly lower than in the two-word sequential condition (0.40 vs 0.41 p.s.c), and that difference was not significant ($t(19)=0.49$, $p=0.627$, BF=0.259). Thus, the VWFA responds more strongly to two words than to one, regardless of whether the two words are presented in rapid succession or flashed simultaneously

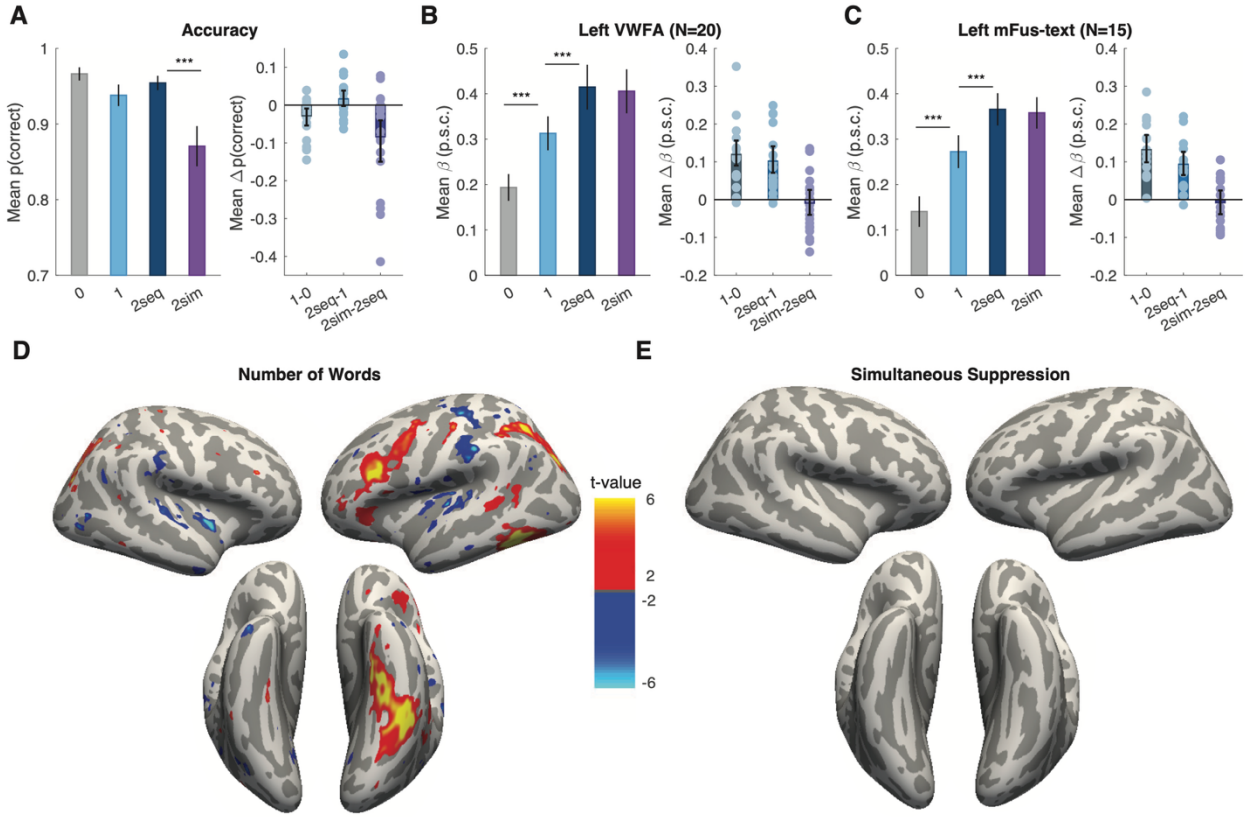

Figure 2: Effects of presentation condition on behavior and brain responses. **(A)** The left panel shows mean accuracy for reporting the number of words as a function of the presentation condition. "2seq" means two words presented sequentially, "2sim" means two words presented simultaneously. Error bars are ± 1 SEM. The right plot in panel A shows the *differences* in accuracy across pairs of conditions. The dots are individual participants. The bars are the means, and the error bars are 95% bootstrapped confidence intervals. **(B)** Mean BOLD responses in the left VWFA, averaged over VWFA-1 and -2. The units of the responses, or β weights, are percent signal change (p.s.c). As in Panel A, the left plot shows the means in each condition and the right shows the differences between pairs of conditions. The VWFA response increases with the number of words presented, but does not reliably differ between two-word sequential and two-word simultaneous conditions. **(C)** Mean BOLD responses in left mFus-text. Similar to left VWFA responses, left mFus-text responses increased with the number of words, but showed no simultaneous suppression effect. **(D)** Whole-cortex maps of the effect of the number of words presented per trial. Colors represented t-values on the mean slope across participants, thresholded with false discovery rate correction across all surface notes. We found an effect of the number of words in IFS, IPS and text-selective regions on the ventral surface. **(E)** Whole-cortex analysis of the simultaneous suppression effect. After correction for multiple comparisons, no surface nodes showed a significant difference between two-word simultaneous and sequential trials.

In 15 of our 21 participants, we also identified a text-selective area in the middle of the left fusiform gyrus or mid-fusiform sulcus, which we call mFus-text (see example in **Figure 1D**). This might overlap with a region identified with direct intracranial recordings that plays a causal role in word reading (Woolnough et al., 2021) and picture naming (Forseth et al., 2018; Woolnough & Tandon, 2024). As shown in **Figure 2C,** the pattern of results in this region was consistent with the left VWFA. The response significantly increased with the number of words (mean slope = 0.11 p.s.c/word, $t(14)=9.92$, $p<0.001$), linearly from zero-word to one-word trials and one-word to two-word-sequential trials (mean difference of differences=0.04±0.03 p.s.c., CI = [-0.09, 0.01], $t(14)=1.4$, $p=0.18$, BF=0.59). There was no simultaneous suppression effect in mFus-text (mean difference between sequential and simultaneous trials = 0.007±0.02, CI = [-0.04, 0.02] p.s.c., $t(14)=0.47$, $p=0.64$, BF=0.28).

Whole-cortex analyses reveal a language network that responds strongly to multiple words

A whole-cortex analysis revealed three large clusters in which the response increased significantly with the number of words presented. **Figure 3C** shows the map of t-values for the positive linear slope of the BOLD response with the number of words per trial, thresholded for FDR-corrected $p<0.05$. The first cluster was in the left ventral temporal cortex, covering parts of the occipito-temporal sulcus and the fusiform gyrus. This corresponds with the locations of the VWFA and mFus-text in most subjects. The second cluster sensitive to the number of words was in the left frontal cortex, including the inferior frontal sulcus (IFS) and extending posteriorly to the premotor cortex. This includes regions that are often referred to as "Broca's area", and it overlaps with some of the language ROIs that are analyzed separately below. The third cluster was in the

intraparietal sulcus (IPS), bilaterally. These patches overlap with regions known to be involved in attentional control (Lauritzen et al., 2009; Silver et al., 2005). In the left hemisphere, this IPS region extended more anteriorly than in the right, and may be more specifically involved in word recognition and language (Forseth et al., 2018; Ossmy et al., 2014; Rapp et al., 2016; Woolnough et al., 2022). All of these word-sensitive regions are known to show correlated activity during word recognition tasks (Chauhan et al., 2024; Stevens et al., 2017; Vogel et al., 2012; White et al., 2023).

We also performed a whole-cortex analysis of the simultaneous suppression effect: the difference in responses between two-word simultaneous and two-word sequential trials (**Figure 2D**). After correction for multiple comparisons, we found no surface nodes with a significant effect.

However, just because the mean beta response for two words presented simultaneously is equivalent to when they are presented sequentially, that does not mean that the words were processed to the same extent. To investigate whether more information can be extracted from words when their presentation is separated in time, we performed the following analysis.

Evidence of a capacity limit from lexical frequency effects

To investigate whether the two words presented simultaneously were processed equally well as two words presented sequentially, we performed a lexical frequency analysis. The VWFA and mFus-text are known to respond more strongly to *less* common words (e.g., Chauhan et al., 2024; Kronbichler et al., 2004; Woolnough et al., 2021). This effect of lexical frequency is thought to reflect the process of lexical access, when a letter string is matched to a known word in the orthographic lexicon. Here we investigated whether text-selective regions are sensitive to the lexical frequencies of two simultaneous words (inspired by White et al., 2019).

We separately analyzed two sets of trials: those in which both words were of low frequency (below the median of 4.1) and those in which both words were of high frequency (above the median). The results for left VWFA and mFus-text are plotted in **Figure 3.** To determine whether the lexical frequency effect on BOLD responses differed across simultaneous and sequential presentation conditions, we fit a linear mixed effects model with two interacting fixed effects: lexical frequency bin and presentation condition. We also conducted pairwise t-tests of the lexical frequency effect within each presentation condition. In the left VWFA, we found a significant main effect of lexical frequency ($F(1, 2715)=5.7$, $p=0.017$) but no significant effects of presentation condition ($F(1, 2715)=3.2$, $p=0.073$) or interaction ($F(1, 2715)=2.25$, $p=0.133$). In the sequential

condition, the response to low-frequency words was on average 0.07±0.025 p.s.c greater than to high-frequency words (95% CI = [0.03, 0.12]; t(20)=2.82, p=0.011; BF=4.85). In the simultaneous condition, that mean difference was smaller: 0.017±0.021 p.s.c (95% CI = [-0.03, 0.05]; t(20)=0.77, p=0.44; BF=0.30).

This pattern of results was stronger in mFus-text (**Figure 3B)**. The main effects of lexical frequency bin and presentation condition were not significant (F(1,1961)=1.86, p=0.172; F(1,1961)=3.40, p=0.07, respectively). However, there was a significant interaction between lexical frequency and presentation condition (F(1, 1961)=4.41, p=0.0356). The lexical frequency effect was robust in the sequential condition (mean difference = 0.117±0.031 p.s.c, CI: [0.061, 0.184], (t(14)=3.65, p=0.003, BF=17.1) but not in the simultaneous condition (mean difference = 0.029±0.05 p.s.c., CI: [-0.043, 0.159], (t(20)=0.54, p=0.60, BF=0.30).

These results demonstrate that although on average across the whole stimulus set there was no simultaneous suppression (**Figure 2**), the most anterior text-selective region (mFus text) processed two words less deeply when they were presented simultaneously than sequentially. The lexical frequency effect is a signature of familiarity with the individual words, beyond mere detection of words amidst false font strings, and it was strong only when words were presented sequentially.

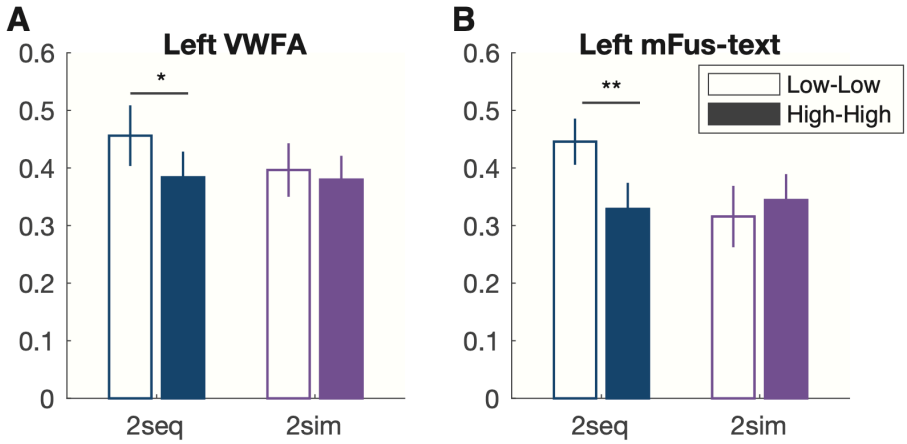

Figure 3: Lexical frequency effects on responses to two words in text-selective regions. **(A)** Average responses in the left VWFA. “2-seq” means two words presented sequentially, “2-sim” means two words presented simultaneously. Filled bars are for trials with two high-frequency words, empty bars are for trials with two low-frequency words. Error bars are ± 1 SEM. **(B)** Average responses in left mFus-text.

No simultaneous suppression in other visual or language areas

We also quantified the effects of the number of words and of simultaneous suppression in seven other individually-defined regions, which we report in two sets: bilateral occipito-temporal visual regions, and then left frontal and temporal language regions. The first set includes: V4; the fusiform face area (FFA), and a limb-selective area (VTC-limbs). These results are plotted in **Figure 4**, and statistics are reported in **Table 1.**

We focus first on the effect of the number of words in each trial. V4 and VTC-limbs were not sensitive to the number of words. However, there was a significant effect in the fusiform face area, both in the left hemisphere ($t(19)=5.68$, $p<0.001$) and the right hemisphere ($t(20)=2.72$, $p=0.01$). Unlike the left VWFA, however, neither of these regions showed a significant increase from the 1-word to 2-word conditions. The response magnitudes were larger in the left FFA than the right FFA, and the slope was twice as large in the left. This partial sensitivity to words in the left FFA, which we have observed before (Chauhan et al., 2024; White et al., 2023), could be due to partial overlap between primarily word-selective and primarily face-selective neural populations. Our ROI boundaries are somewhat arbitrary, it is unsurprising that the FFA shows a pattern of sensitivity to words that is similar to the neighboring VWFA, but weaker.

Importantly, none of these visual regions showed evidence of simultaneous suppression: responses were statistically indistinguishable between the 2-sequential and 2-simultaneous conditions, as in the VWFA.

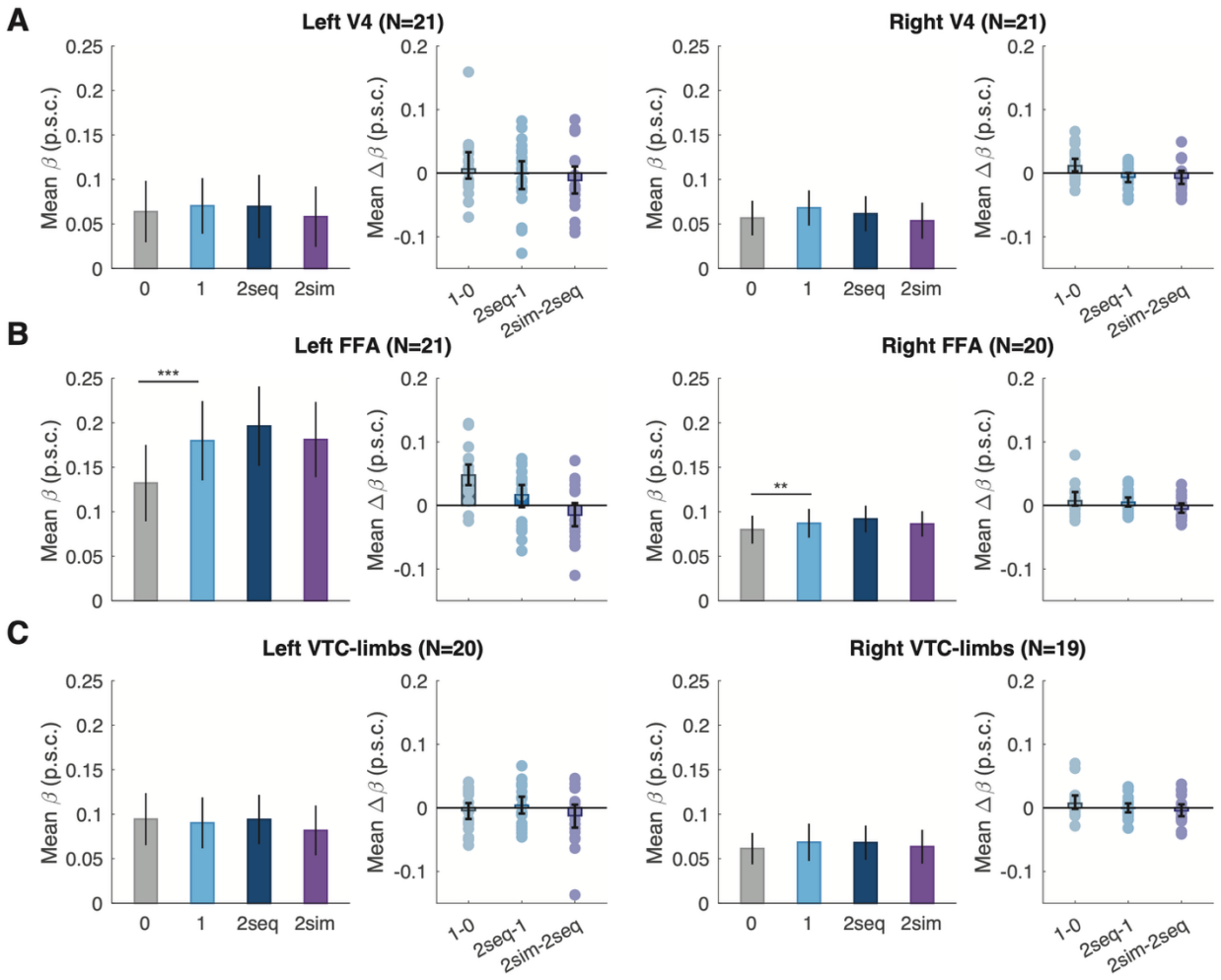

Figure 4: Effect of stimulus condition on BOLD responses in control visual ROIs. Colors represent experimental conditions, left panels indicate mean BOLD responses (error bars represent SEM), right panels depict differences in beta responses between pairwise conditions (error bars represent bootstrapped 95% confidence intervals, dots correspond to differences within individual subjects).

Results for the language regions are visualized in **Figure 5**. These regions were defined in individual participants from a contrast of response to meaningful sentences versus meaningless pseudoword sequences. We begin with the three regions in the left frontal lobe. In none of these frontal regions was there evidence of simultaneous suppression, although they differed in how sensitive they were to the number of words presented on each trial (see **Table 1)**. First, activity in the most dorsal region, the precentral gyrus (PG), did not vary significantly with the number of words ($t(12)=1.67$, $p=0.14$). Second, a region in the left inferior frontal sulcus (IFS) demonstrated a pattern of responses similar to the VWFA (**Figure 2C)**. The linear slopes fit to the response magnitude was significantly greater than zero ($t(17)=5.44$, $p<0.001$), with increases from 0 to 1 and from 1 to 2 words (**Figure 5B**). Third, in the operculum, there was a significantly positive

slope as a function of the number of words ($t(13)=3.16$, $p=0.004$). However, for PG and the operculum, only the increase from 0 to 1-word conditions was statistically significant on its own.

Interestingly, on zero-word trials (in which only false font strings were presented), the average response in the operculum dipped below the baseline BOLD level of 0 (by 0.05±0.02 p.s.c. on average; CI = [-0.009, 0.093], $t(13)=2.21$, $p=0.045$; BF=1.69). See the gray bar in **Figure 5C**. This is reminiscent of our previous finding that unfamiliar characters elicit suppression of the reading circuitry (but only when the participant is looking for familiar words; Chauhan et al., 2024).

Lastly, we identified a large language-sensitive region in the left superior temporal sulcus (STS; **Figure 5D**). Its response steadily increased with the number of words presented ($t(19)=3.56$, $p<0.001$). Interestingly, it *did* respond more weakly in the 2-word simultaneous condition compared to the sequential condition. The 95% bootstrapped confidence interval for that difference excludes 0, as visualized by the purple bar in the right panel of **Figure 5D**. The Bayes factor for that comparison was 2.0, indicating a weak effect. After adjusting p-values for multiple comparisons across regions, that simultaneous suppression effect was not statistically significant. This also matches the absence effect anywhere on the cortex (**Figure 2D**), so the small simultaneous suppression effect in the STS may be a false positive. Nonetheless, it presents an interesting opportunity for further research. Perhaps a capacity limit for processing two words at once arises at the level of combining orthography with phonology.

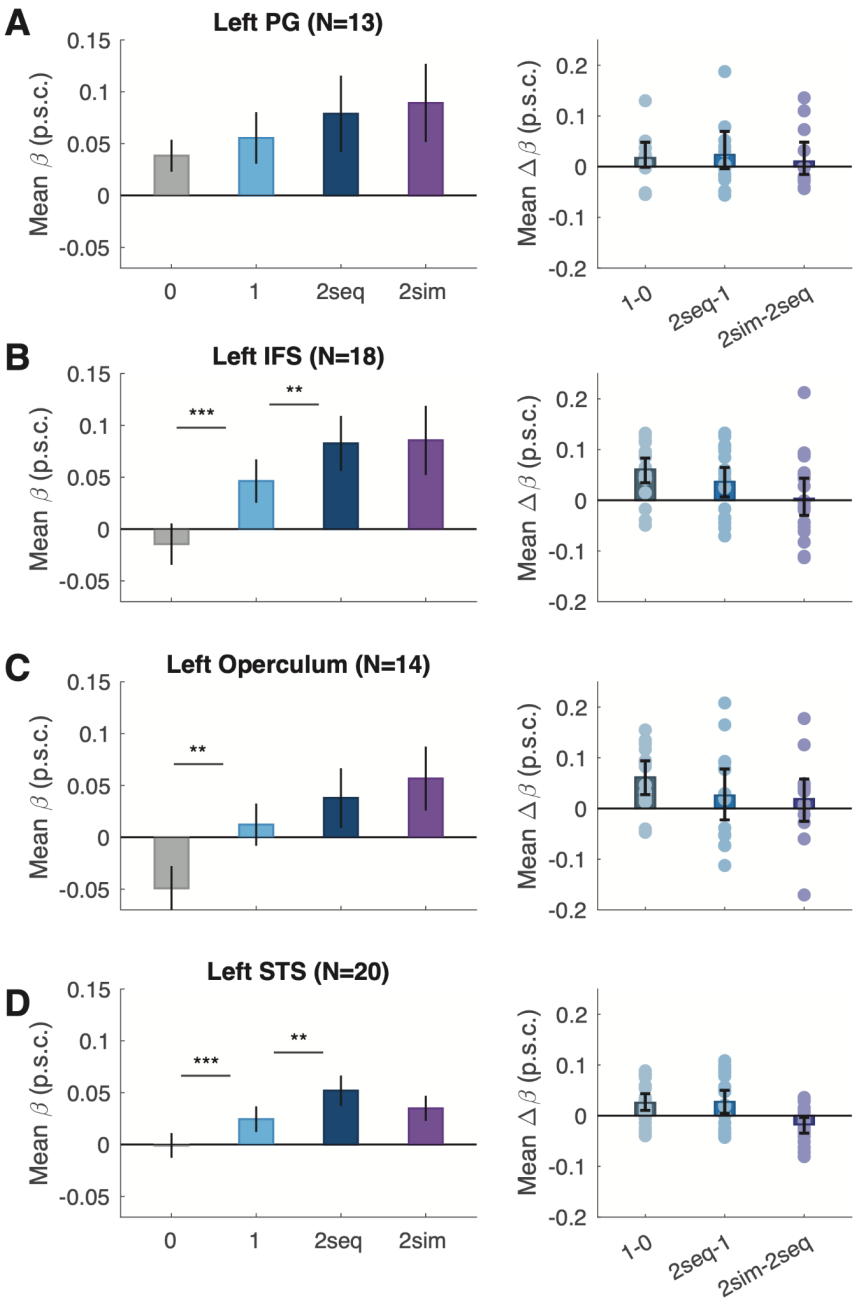

Figure 5: Effect of stimulus condition on BOLD responses in four language ROIs. Format as in Figure 4.

Lastly, none of these visual or language ROIs (other than the text-selective ones plotted in **Figure 3**) were sensitive to the lexical frequencies of the words, and none of them yielded a significant interaction between lexical frequency bin and presentation condition.

ROI	Experimental Condition	Hemisphere	N	Slope	SEM	95% CI	t value	p value
VWFA	Number of Words	Left	20	0.11	0.012	0.08, 0.136	8.69	0
		Right	13	0.007	0.004	-0.0008, 0.015	1.772	0.123
	Simultaneous Suppression	Left	20	-0.005	0.008	-0.022, 0.012	-0.594	0.552
		Right	13	0.002	0.003	-0.004, 0.010	0.779	0.674
mFus-text	Number of Words	Left	15	0.113	0.011	0.091, 0.135	9.923	0
	Simultaneous Suppression	Left	15	-0.004	0.01	-0.025, 0.016	-0.42	0.674
FFA	Number of Words	Left	21	0.032	0.005	0.021, 0.0436	5.685	0
		Right	20	0.006	0.002	0.002, 0.010	2.73	0.013
	Simultaneous Suppression	Left	21	-0.008	0.004	-0.017, 0.0008	-1.773	0.434
		Right	20	0.003	0.002	-0.0006, 0.0007	1.63	0.433
VTC-limbs	Number of Words	Left	21	-0.0004	0.005	-0.009, 0.008	-0.105	0.483
		Right	19	0.003	0.002	-0.001, 0.008	1.398	0.2
	Simultaneous Suppression	Left	20	0.006	0.005	-0.003, 0.015	1.279	0.468
		Right	19	0.003	0.002	-0.001, 0.007	1.144	0.504
V4	Number of Words	Left	21	0.003	0.004	-0.006, 0.0116	0.585	0.617
		Right	21	0.002	0.002	-0.002, 0.008	1.031	0.353

	Simultaneous Suppression	Left	21	-0.006	0.005	-0.016, 0.004	-1.116	0.504
		Right	21	-0.004	0.003	-0.009, 0.0008	-1.629	0.434
STS	Number of Words	Left	20	0.026	0.007	0.012, 0.041	3.564	0.001
		Right	17	0.006	0.003	0.0003, 0.011	2.07	0.07
	Simultaneous Suppression	Left	20	-0.008	0.004	-0.016, -0.0006	-2.103	0.372
		Right	17	0.002	0.002	-0.003, 0.006	0.643	0.674
PG	Number of Words	Left	13	0.02	0.012	-0.003, 0.044	1.67	0.138
	Simultaneous Suppression	Left	13	-0.005	0.008	-0.02, 0.01	-0.635	0.674
IFS	Number of Words	Left	18	0.048	0.009	0.03, 0.066	5.544	0
	Simultaneous Suppression	Left	18	0.001	0.009	-0.017, 0.019	0.131	0.895
Operculum	Number of Words	Left	14	0.043	0.014	0.016, 0.07	3.159	0.004
	Simultaneous Suppression	Left	14	-0.01	0.01	-0.03, 0.011	-0.934	0.613

Table 1: Statistical results obtained from the two linear mixed effects models: Number of Words and Simultaneous Suppression. The first modelled beta activity as a function of the number of words presented. The second only used two-word sequential and two-word simultaneous trials and modelled beta activity as a function of experimental condition. Here, we report the estimates of the slope, the standard error, 95% bootstrapped confidence intervals, t-values and associated p-values corrected for multiple tests using FDR correction. N corresponds to the number of subjects in which a given ROI is present. Significant results are presented in bold text.

DISCUSSION

Summary

The primary goal of this study was to investigate how reading-related regions in the cortex respond to multiple written words. Words are most often seen with other words, but nearly all prior fMRI studies of visual word recognition presented just one word at a time. We first established that in text-selective ventral temporal regions (the left VWFA and mFus-Text) and in frontal and temporal language-processing regions, the BOLD response magnitude steadily increases as the number of words presented in a single trial increases from 0 to 1 to 2. Importantly, these trial types were equated in visual stimulation: they contained an equal number of characters and an equal number of stimulus offsets and onsets. What differed between them was the number of meaningful words that could be read.

Given the response increase from 1 word to 2 words in these regions, we could then ask whether they show a simultaneous suppression effect. In the behavioral data, we did observe a deficit in reporting the number of words when two words were presented simultaneously as compared to sequentially. This suggests a processing capacity limit, consistent with past research (e.g., Scharf et al. 2011; White et al. 2020). The neural correlate of that capacity limit could have manifest as a weaker BOLD response to two words presented simultaneously as compared to sequentially, as has been observed for other types of stimuli in visual cortex (Kaiser et al., 2014; Kastner, Weerd, et al., 2001; Miller et al., 1993; Shim et al., 2013). However, we did not observe statistically significant simultaneous suppression in any of our regions of interest nor in a whole-cortex analysis. In the left VWFA, the evidence favored the null hypothesis for equivalent response magnitudes between the simultaneous and sequential conditions (BF=0.27). Similar patterns obtained in all other text- and language-selective regions, and in the two VWFA sub-regions analyzed separately. Thus, these data demonstrate that the reading circuitry can respond to two written words simultaneously. The comparison of the BOLD response to the sequential presentation condition does not explain the deficit observed behaviorally.

The one possible exception to the absence of the simultaneous suppression effect was the left superior temporal sulcus (STS), where a 95% bootstrapped confidence interval for the simultaneous suppression effect excluded 0 (with a modest BF of 2.0). However, this effect was not statistically significant after correcting for multiple comparisons across ROIs. Future research may investigate whether the processing function carried out by the STS is the source of the processing capacity limits observed behaviorally.

However, the absence of a simultaneous suppression effect is not equivalent to an absence of capacity limits in text-selective regions. This difference is established through our analysis of lexical frequency. Specifically, in the mid-fusiform text-selective area, the response was higher for two low-frequency words than two high-frequency words, but only when the words were presented sequentially. Our interpretation is that in the simultaneous presentation condition, a processing capacity limit hindered lexical access, resulting in less sensitivity to the words' frequencies. The fact that this result was clearest in mFus-text is consistent with demonstrations of this region's special role in lexical access. In intracranial recordings, sensitivity to lexical frequency arises in the mid-fusiform area before it arises in the most posterior VWFA (Woolnough et al., 2021). When the words were presented simultaneously, they may have been processed less deeply or faithfully.

The analysis of mean BOLD response, across all words in the stimulus set, simply relies upon the brain's *detection* of letter shapes in a stream of false font shapes. Thus, the absence of simultaneous suppression in the main analysis could mean that the reading circuitry can detect two strings of legible letters at once. However, an effect of lexical frequency is dependent on the familiarity of whole words (as specific combinations of letters or as units of meaning), which comes with expertise. That level of processing seems to have limited capacity when two words are processed simultaneously.

There is another way to interpret the lexical frequency results: Had all the words in the stimulus set been of low frequency, we may have observed a simultaneous suppression effect in the main analysis (note the difference in blue and purple empty bars in **Figure 3**). Relatedly, had our task required semantic judgements of the words, rather than simply detection, we may have observed a simultaneous suppression effect overall. Therefore, an open question for further research is whether the difficulty of the task influences how processing capacity limits manifest as simultaneous suppression.

Differences from previous studies of simultaneous suppression

Considering past research on simultaneous suppression in visual cortex, it is remarkable that in the main analysis (**Figure 2**), we found that equally strong responses to two words presented simultaneously as sequentially, both being significantly above the response to a single word. But in our experimental design we made several modifications to the typical simultaneous suppression paradigm that could explain why our results differ.

First, we attempted to equate the amount of visual stimulation across conditions. Not only did we equalize the number of onsets and offsets in each condition (zero-word, one-word, two-word sequential and two-word simultaneous), we also embedded the words in sequences of false font strings that are matched to the words in several visual features (Vidal et al., 2017). This approach was meant to isolate the response in the reading circuitry that is selective for familiar letter strings. Had we not done so, it is possible that we would have found a simultaneous suppression effect in VWFA and other language regions. However, in such a scenario, we might not be certain if the effect arose from a limit on processing two words *as words* simultaneously, or a non-linearity in the temporal integration of responses to stimulus transients, as Kupers et al. (2024) proposed. Moreover, if we had simply presented words without the equalizing “filler” false font strings, it would have been trivial to observe that the response steadily increases with the number of words. In this study, the linear increase in the VWFA's response magnitude from 0 to 1 to 2 words is therefore important. It shows that, despite the matched visual stimulation, the trial-evoked BOLD response does linearly add together the responses to two words.

Second, the stimulus duration may be an important factor. Each word was presented for 183 ms, which we chose to minimize the chance of goal-directed eye movements from one location to another within a frame. The time between frame onsets was 233 ms, roughly the average fixation duration during natural reading (Rayner et al., 2016). This stimulus onset asynchrony is also not far from the 250 ms used in classic simultaneous suppression studies (e.g., Kastner et al., 1998; Kastner & Ungerleider, 2001; McMains & Kastner, 2011). However, there is some evidence that the effect increases as stimulus durations increase above 200 ms (Kupers et al., 2024). We reasoned that if we made the stimulus duration too long, attention could move from one word to another within a single frame, allowing for *serial* processing of two words presented simultaneously. In other words, even if there is a serial bottleneck, there would be no evidence of a capacity limit in the data if there is enough time for the two simultaneously presented words to be processed sequentially. Therefore, we might have found more evidence of simultaneous suppression if the presentation durations were even *shorter*. That possibility remains to be explored with future research.

Third, the stimulus location may have played a role in explaining our results. We presented our stimuli within the large population receptive fields of voxels in the VWFA, which cover roughly the central 5 degrees of the visual field (R. Le et al., 2017). Given past demonstrations that simultaneous suppression increases up the visual hierarchy and with increasing population receptive field size, we would expect a strong effect in ventral temporal cortex (Kastner, De

Weerd, et al., 2001; Kupers et al., 2024). To ensure that both words were as close to fixation as possible, we chose to present the words above and below the fixation, instead of to the left and right, as words are typically read. If anything, we would expect *less* of a capacity limit in more natural presentation formats, and stronger simultaneous suppression for words stacked vertically that are perhaps more difficult to process together.

Lastly, the role of attention and top-down influences must be considered. While the earliest studies of simultaneous suppression in visual regions demonstrated a stronger effect when stimuli are ignored rather than attended (Kastner et al., 1998; S. McMains & Kastner, 2011), the effect is still present even when the stimuli are attended (Scalf et al., 2011). As our primary question is whether two words can be *read* simultaneously, we made our stimuli task-relevant. It is known that the VWFA's activity is strongly modulated by voluntary effort to read words, likely via top-down signals from frontal and parietal regions (Chauhan et al., 2024; Kay & Yeatman, 2017; Mano et al., 2013; White et al., 2023). Thus, it is possible that top-down signals counteracted a simultaneous suppression effect that we might have observed if the stimuli were ignored. At the same time, a simultaneous suppression effect might have emerged with a task that required deeper lexical or semantic processing (as already discussed in the interpretation of the lexical frequency effects). That question deserves to be explored further. However, the concern is mitigated by three findings: that there was a behavioral deficit in the simultaneous condition; that higher-level language regions were engaged and sensitive to the number of words presented; and that mFus-text was less sensitive to the lexical frequencies of two simultaneous words. This last result indicates a capacity limit.

Comparison to other reports of a bottleneck in the reading circuitry

Only a few fMRI studies have measured VWFA responses to more than one word presented at a time, and none have directly compared how these responses change as the number of words is increased. White, Palmer, et al. (2019) measured BOLD responses to pairs of words while participants performed a semantic categorization task and either focused attention on one word or divided attention between both. They concluded that a posterior sub-region of the VWFA (VWFA-1) was capable of processing two words at once, with separate "spatial channels" for one word on the left of fixation and another on the right. The more anterior sub-region (VWFA-2) however, seemed to lack those separable spatial channels and was sensitive to the lexical frequency of only one attended word per trial. Thus, the activity of VWFA-2 was proposed to account for the behavioral deficits in processing two words at once with divided attention, consistent with a "serial bottleneck" in the reading circuitry.

In that 2019 study as well as the present study, the neuronal sensitivity to lexical frequencies was key to revealing a region with a processing capacity limit. We found that mFus-text (a region just anterior/medial to VWFA-2) was sensitive to the lexical frequencies of two words only when they were presented sequentially. That mirrors the previous finding of how spatial attention gates lexical frequency effects for one word at a time in VWFA-2 (White et al. 2019). Nonetheless, those previous findings led us to predict an overall effect of simultaneous suppression in VWFA-2, which we did not find. This apparent conflict between studies could be accounted for by the difference in tasks and top-down control, or the more rapid presentation durations (50 ms) in White et al. (2019). It is also possible that the comparison of simultaneous vs sequential word presentations revealed a capacity for parallel processing (perhaps at a sub-lexical level) that could not be detected in the previous study.

Future directions

Overall, we found that activity in several reading-related regions increased as the number of words increased, up to two simultaneously presented words. An important question for future research is how far this goes: as the number of words increases, at what point does the response asymptote? And is that threshold the same for all regions?

Future research investigating simultaneous suppression in word recognition would involve systematic behavioral testing of the effects of stimulus location, duration and task demands. Our findings based on lexical frequency effects highlight the importance of measures that go beyond mean BOLD responses, by taking the linguistic properties of words into consideration. Going even further, our results present interesting avenues for exploring how responses to multiple words depend on the linguistic relationships between them: such as pairs of words that form a compound word, semantically related words, or words that form syntactic phrases.

REFERENCES

Adelman, J. S., Marquis, S. J., & Sabatos-DeVito, M. G. (2010). Letters in Words Are Read Simultaneously, Not in Left-to-Right Sequence. *Psychological Science*, *21*(12), 1799–1801. https://doi.org/10.1177/0956797610387442

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project.

Behavior Research Methods, *39*(3), 445–459. https://doi.org/10.3758/BF03193014

Barton, J. J. S., Hanif, H. M., Eklinder Björnström, L., & Hills, C. (2014). The word-length effect in reading: A review. *Cognitive Neuropsychology*, *31*(5–6), 378–412. https://doi.org/10.1080/02643294.2014.895314

Beck, D. M., & Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Research*, *49*(10), 1154–1165. https://doi.org/10.1016/j.visres.2008.07.012

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Ben-Shachar, M., Dougherty, R. F., Deutsch, G. K., & Wandell, B. A. (2011). The development of cortical sensitivity to visual word forms. *Journal of Cognitive Neuroscience*, *23*(9), 2387–2399. https://doi.org/10.1162/jocn.2011.21615

Brem, S., Maurer, U., Kronbichler, M., Schurz, M., Richlan, F., Blau, V., Reithler, J., van der Mark, S., Schulz, E., Bucher, K., Moll, K., Landerl, K., Martin, E., Goebel, R., Schulte-Körne, G., Blomert, L., Wimmer, H., & Brandeis, D. (2020). Visual word form processing deficits driven by severity of reading impairments in children with developmental dyslexia. *Scientific Reports*, *10*(1), 1–14. https://doi.org/10.1038/s41598-020-75111-8

Brothers, T. (2022). Capacity limits in sentence comprehension: Evidence from dual-task judgements and event-related potentials. *Cognition*, *225*(May), 105153. https://doi.org/10.1016/j.cognition.2022.105153

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, Instruments & Computers*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Caffarra, S., Karipidis, I. I., Yablonski, M., & Yeatman, J. D. (2021). Anatomy and physiology of

word-selective visual cortex: From visual features to lexical processing. *Brain Structure and Function*. https://doi.org/10.1007/s00429-021-02384-8

Campbell, M., Oppenheimer, N., & White, A. L. (2024). Severe processing capacity limits for sub-lexical features of letter strings. *Attention, Perception, & Psychophysics*, *86*(2), 643–652. https://doi.org/10.3758/s13414-023-02830-1

Chauhan, V. S., McCook, K. C., & White, A. L. (2024a). Reading Reshapes Stimulus Selectivity in the Visual Word Form Area. *eNeuro*, *11*(7), ENEURO.0228-24.2024. https://doi.org/10.1523/ENEURO.0228-24.2024

Chauhan, V. S., McCook, K. C., & White, A. L. (2024b). Reading Reshapes Stimulus Selectivity in the Visual Word Form Area. *eNeuro*, *11*(7). https://doi.org/10.1523/ENEURO.0228-24.2024

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), 254–262. https://doi.org/10.1016/j.tics.2011.04.003

Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015a). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews Neuroscience*, *16*(4), 234–244. https://doi.org/10.1038/nrn3924

Dehaene, S., Cohen, L., Morais, J., & Kolinsky, R. (2015b). Illiterate to literate: Behavioural and cerebral changes induced by reading acquisition. *Nature Reviews Neuroscience*, *16*(4), 234–244. https://doi.org/10.1038/nrn3924

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 Biology*, *16*(3), e2004103. https://doi.org/10.1371/journal.pbio.2004103

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A Cortical Area Selective for Visual Processing of the Human Body. *Science*, *293*(5539), 2470–2473. https://doi.org/10.1126/science.1063414

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, *87*(3), 272–300. https://doi.org/10.1037/0033-295X.87.3.272

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: a dynamical model of saccade generation during reading. *Psychological Review*, *112*(4), 777–813. https://doi.org/10.1037/0033-295X.112.4.777

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., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*, *16*(1), 111–116. https://doi.org/10.1038/s41592-018-0235-4

Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New Method for fMRI Investigations of Language: Defining ROIs Functionally in Individual Subjects. *Journal of Neurophysiology*, *104*(2), 1177–1194. https://doi.org/10.1152/jn.00032.2010

Forseth, K. J., Kadipasaoglu, C. M., Conner, C. R., Hickok, G., Knight, R. T., & Tandon, N. (2018). A lexical semantic hub for heteromodal naming in middle fusiform gyrus. *Brain: A Journal of Neurology*, *141*(7), 2112–2126. https://doi.org/10.1093/brain/awy120

Gaillard, R., Naccache, L., Pinel, P., Clémenceau, S., Volle, E., Hasboun, D., Dupont, S., Baulac, M., Dehaene, S., Adam, C., & Cohen, L. (2006). Direct Intracranial, fMRI, and Lesion Evidence for the Causal Role of Left Inferotemporal Cortex in Reading. *Neuron*, *50*(2), 191–204. https://doi.org/10.1016/j.neuron.2006.03.031

Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data

Processing Framework in Python. *Frontiers in Neuroinformatics*, *5*. https://doi.org/10.3389/fninf.2011.00013

Grainger, J., Dufau, S., & Ziegler, J. C. (2016). A Vision of Reading. *Trends in Cognitive Sciences*, *20*(3), 171–179. https://doi.org/10.1016/j.tics.2015.12.008

Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience*, *15*(8), 536–548. https://doi.org/10.1038/nrn3747

Hossain, J., & White, A. L. (2023). The transposed word effect is consistent with serial word recognition and varies with reading speed. *Cognition*, *238*, 105512. https://doi.org/10.1016/j.cognition.2023.105512

Huang, L., & Pashler, H. (2005). Attention capacity and task difficulty in visual search. *Cognition*, *94*(3), B101-111. https://doi.org/10.1016/j.cognition.2004.06.006

Johnson, M. L., Palmer, J., Moore, C. M., & Boynton, G. M. (2022). Evidence from partially valid cueing that words are processed serially. *Psychonomic Bulletin & Review*, *0123456789*. https://doi.org/10.3758/s13423-022-02230-w

Kaiser, D., Stein, T., & Peelen, M. V. (2014). Object grouping based on real-world regularities facilitates perception by reducing competitive interactions in visual cortex. *Proceedings of the National Academy of Sciences*, *111*(30), 11217–11222. https://doi.org/10.1073/pnas.1400559111

Kastner, S., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1998). Mechanisms of Directed Attention in the Human Extrastriate Cortex as Revealed by Functional MRI. *Science*, *283*, 108–111. https://doi.org/10.1016/s1053-8119(18)30877-2

Kastner, S., De Weerd, P., Pinsk, M. A., Elizondo, M. I., Desimone, R., & Ungerleider, L. G. (2001). Modulation of Sensory Suppression: Implications for Receptive Field Sizes in the Human Visual Cortex. *Journal of Neurophysiology*, *86*(3), 1398–1411. https://doi.org/10.1152/jn.2001.86.3.1398

Kastner, S., & Ungerleider, L. G. (2001). The neural basis of biased competition in human visual cortex. *Neuropsychologia*, *39*, 1263–1276.

Kastner, S., Weerd, P. D., Pinsk, M. A., Elizondo, M. I., Desimone, R., & Ungerleider, L. G. (2001). Modulation of Sensory Suppression: Implications for Receptive Field Sizes in the Human Visual Cortex. *Journal of Neurophysiology*, *86*, 1398–1411. https://doi.org/0022-3077/01

Kay, K. N., & Yeatman, J. D. (2017). Bottom-up and top-down computations in word- and face-selective cortex. *eLife*, *6*, e22341. https://doi.org/10.7554/eLife.22341

Kleiss, J. A., & Lane, D. M. (1986). Locus and persistence of capacity limitations in visual information processing. *Journal of Experimental Psychology. Human Perception and Performance*, *12*(2), 200–210. https://doi.org/10.1037//0096-1523.12.2.200

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: Evidence from a parametric fMRI study. *NeuroImage*, *21*(3), 946–953. https://doi.org/10.1016/j.neuroimage.2003.10.021

Kubota, E. C., Joo, S. J., Huber, E., & Yeatman, J. D. (2019). Word selectivity in high-level visual cortex and reading skill. *Developmental Cognitive Neuroscience*, *36*(June 2018), 100593. https://doi.org/10.1016/j.dcn.2018.09.003

Kupers, E. R., Kim, I., & Grill-Spector, K. (2024a). Rethinking simultaneous suppression in visual cortex via compressive spatiotemporal population receptive fields. *Nature Communications*, *15*(1), 6885. https://doi.org/10.1038/s41467-024-51243-7

Kupers, E. R., Kim, I., & Grill-Spector, K. (2024b). Rethinking simultaneous suppression in visual cortex via compressive spatiotemporal population receptive fields. *Nature Communications*, *15*(1), 6885. https://doi.org/10.1038/s41467-024-51243-7

Lauritzen, T. Z., D'Esposito, M., Heeger, D. J., & Silver, M. A. (2009). Top–down flow of visual spatial attention signals from parietal to occipital cortex. *Journal of Vision*, *9*(13), 18.

https://doi.org/10.1167/9.13.18

Le, R. K., Witthoft, N., Ben-Shachar, M., & Wandell, B. A. (2017). The field of view available to the ventral occipito-temporal reading circuitry. *Journal of Vision*, *17*(4), 1–19. https://doi.org/10.1167/17.4.6

Le, R., Witthoft, N., Ben-Shachar, M., & Wandell, B. (2017a). The field of view available to the ventral occipito-temporal reading circuitry. *Journal of Vision*, *17*(4), 6. https://doi.org/10.1167/17.4.6

Le, R., Witthoft, N., Ben-Shachar, M., & Wandell, B. (2017b). The field of view available to the ventral occipito-temporal reading circuitry. *Journal of Vision*, *17*(4), 6. https://doi.org/10.1167/17.4.6

Mano, Q. R., Humphries, C., Desai, R. H., Seidenberg, M. S., Osmon, D. C., Stengel, B. C., & Binder, J. R. (2013). The role of left occipitotemporal cortex in reading: Reconciling stimulus, task, and lexicality effects. *Cerebral Cortex*, *23*(4), 988–1001. https://doi.org/10.1093/cercor/bhs093

McMains, S. A., & Kastner, S. (2011). Interactions of top-down and bottom-up mechanisms in human visual cortex. *Journal of Neuroscience*, *31*(2), 587–597. https://doi.org/10.1523/JNEUROSCI.3766-10.2011

McMains, S., & Kastner, S. (2011). Interactions of top-down and bottom-up mechanisms in human visual cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(2), 587–597. https://doi.org/10.1523/JNEUROSCI.3766-10.2011

Milledge, S. V., Bhatia, N., Mensah-Mcleod, L., Raghvani, P., A. McGowan, V., Elsherif, M. M., Cutter, M. G., Wang, J., Liu, Z., & Paterson, K. B. (2023). The transposed-word effect provides no unequivocal evidence for parallel processing. *Attention, Perception, & Psychophysics*, *85*(8), 2538–2546. https://doi.org/10.3758/s13414-023-02721-5

Miller, E. K., Gochin, P. M., & Gross, C. G. (1993). Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque by addition of a second stimulus. *Brain*

Research, *616*(1), 25–29. https://doi.org/10.1016/0006-8993(93)90187-R

Mullin, P. A., & Egeth, H. E. (1989). Capacity limitations in visual word processing. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(1), 111–123. https://doi.org/10.1037/0096-1523.15.1.111

Nordt, M., Gomez, J., Natu, V. S., Rezai, A. A., Finzi, D., Kular, H., & Grill-Spector, K. (2021). Cortical recycling in high-level visual cortex during childhood development. *Nature Human Behaviour*, *5*, 1686–1697. https://doi.org/10.1038/s41562-021-01141-5

Ossmy, O., Ben-Shachar, M., & Mukamel, R. (2014). Decoding letter position in word reading. *Cortex*, *59*, 74–83. https://doi.org/10.1016/j.cortex.2014.07.002

Pflugshaupt, T., Gutbrod, K., Wurtz, P., Von Wartburg, R., Nyffeler, T., De Haan, B., Karnath, H. O., & Mueri, R. M. (2009). About the role of visual field defects in pure alexia. *Brain*, *132*(7), 1907–1917. https://doi.org/10.1093/brain/awp141

Prince, J. S., Charest, I., Kurzawski, J. W., Pyles, J. A., Tarr, M. J., & Kay, K. N. (2022). Improving the accuracy of single-trial fMRI response estimates using GLMsingle. *eLife*, *11*, e77599. https://doi.org/10.7554/eLife.77599

Rapp, B., Purcell, J., Hillis, A. E., Capasso, R., & Miceli, G. (2016). Neural bases of orthographic long-term memory and working memory in dysgraphia. *Brain: A Journal of Neurology*, *139*(Pt 2), 588–604. https://doi.org/10.1093/brain/awv348

Rauschecker, A. M., Bowen, R. F., Parvizi, J., & Wandell, B. A. (2012). Position sensitivity in the visual word form area. *Proceedings of the National Academy of Sciences*, *109*(24), E1568–E1577. https://doi.org/10.1073/pnas.1121304109

Rayner, K., Schotter, E. R., Masson, M. E. J., Potter, M. C., & Treiman, R. (2016). So much to read, so little time: How do we read, and can speed reading help? In *Psychological Science in the Public Interest* (Vol. 17). https://doi.org/10.1177/1529100615623267

Reichle, E. D., Liversedge, S. P., Pollatsek, A., & Rayner, K. (2009). Encoding multiple words simultaneously in reading is implausible. *Trends in Cognitive Sciences*, *13*(3), 115–119.

https://doi.org/10.1016/j.tics.2008.12.002

Reilly, R. G., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research*, *7*(1), 34–55. https://doi.org/10.1016/j.cogsys.2005.07.006

Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive Mechanisms Subserve Attention in Macaque Areas V2 and V4. *The Journal of Neuroscience*, *19*(5), 1736–1753. https://doi.org/10.1523/JNEUROSCI.19-05-01736.1999

Richlan, F., Kronbichler, M., & Wimmer, H. (2009). Functional abnormalities in the dyslexic brain: A quantitative meta-analysis of neuroimaging studies. *Human Brain Mapping*, *30*(10), 3299–3308. https://doi.org/10.1002/hbm.20752

Scalf, P. E., Basak, C., & Beck, D. M. (2011). Attention does more than modulate suppressive interactions: Attending to multiple items. *Experimental Brain Research*, *212*(2), 293–304. https://doi.org/10.1007/s00221-011-2730-z

Scharff, A., Palmer, J., & Moore, C. M. (2011a). Extending the Simultaneous-Sequential Paradigm to Measure Perceptual Capacity for Features and Words. *Journal of Experimental Psychology. Human Perception and Performance*, *37*(3), 813–833. https://doi.org/10.1037/a0021440

Scharff, A., Palmer, J., & Moore, C. M. (2011b). Extending the simultaneous-sequential paradigm to measure perceptual capacity for features and words. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(3), 813–833. https://doi.org/10.1037/a0021440

Schlaggar, B. L., & McCandliss, B. D. (2007). Development of Neural Systems for Reading. *Annual Review of Neuroscience*, *30*(1), 475–503. https://doi.org/10.1146/annurev.neuro.28.061604.135645

Shim, W. M., Jiang, Y. V., & Kanwisher, N. (2013). Redundancy gains in retinotopic cortex. *Journal of Neurophysiology*, *110*(9), 2227–2235. https://doi.org/10.1152/jn.00175.2013

Silver, M. A., Ress, D., & Heeger, D. J. (2005). Topographic maps of visual spatial attention in human parietal cortex. *Journal of Neurophysiology*, *94*(2), 1358–1371. https://doi.org/10.1152/jn.01316.2004

Snell, J., & Grainger, J. (2019). Readers Are Parallel Processors. *Trends in Cognitive Sciences*, *23*(7), 537–546. https://doi.org/10.1016/j.tics.2019.04.006

Snell, J., Van Leipsig, S., Grainger, J., & Meeter, M. (2018). OB1-reader: A model of word recognition and eye movements in text reading. *Psychological Review*, *125*(6), 969–984. https://doi.org/10.1037/rev0000119

Stevens, W. D., Kravitz, D. J., Peng, C. S., Tessler, M. H., & Martin, A. (2017). Privileged Functional Connectivity between the Visual Word Form Area and the Language System. *Journal of Neuroscience*, *37*(21), 5288–5297. https://doi.org/10.1523/JNEUROSCI.0138-17.2017

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A New and Improved Word Frequency Database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Vidal, C., Content, A., & Chetail, F. (2017). BACS: The Brussels Artificial Character Sets for studies in cognitive psychology and neuroscience. *Behavior Research Methods*, *49*(6), 2093–2112. https://doi.org/10.3758/s13428-016-0844-8

Vildavski, V., Lo Verde, L., Blumberg, G., Parsey, J., & Norcia, A. (2021). PseudoSloan: A perimetric-complexity and area-controlled font for vision and reading research. *Journal of Vision*, *21*(9), 2857. https://doi.org/10.1167/jov.21.9.2857

Vogel, A. C., Miezin, F. M., Petersen, S. E., & Schlaggar, B. L. (2012). The putative visual word form area is functionally connected to the dorsal attention network. *Cerebral Cortex (New York, N.Y.: 1991)*, *22*(3), 537–549. https://doi.org/10.1093/cercor/bhr100

White, A. L., Boynton, G. M., & Yeatman, J. D. (2019). You Can’t Recognize Two Words

Simultaneously. *Trends in Cognitive Sciences*, *23*(10), 812–814. https://doi.org/10.1016/j.tics.2019.07.001

White, A. L., Kay, K. N., Tang, K. A., & Yeatman, J. D. (2023). Engaging in word recognition elicits highly specific modulations in visual cortex. *Current Biology*, *33*(7), 1308-1320.e5. https://doi.org/10.1016/j.cub.2023.02.042

White, A. L., Palmer, J., & Boynton, G. M. (2018). Evidence of Serial Processing in Visual Word Recognition. *Psychological Science*, *28*(7), 1062–1071. https://doi.org/doi.org/10.1177/09567976177518

White, A. L., Palmer, J., & Boynton, G. M. (2020). Visual word recognition: Evidence for a serial bottleneck in lexical access. *Attention, Perception, & Psychophysics*, *82*(4), 2000–2017. https://doi.org/10.3758/s13414-019-01916-z

White, A. L., Palmer, J., Boynton, G. M., & Yeatman, J. D. (2019). Parallel spatial channels converge at a bottleneck in anterior word-selective cortex. *Proceedings of the National Academy of Sciences*, *116*(20), 10087–10096. https://doi.org/10.1073/pnas.1822137116

White, A. L., Palmer, J., Sanders, G., Hossain, J., & Zabinsky, Z. B. (2025). Negative effects of redundant targets. *PsyArxiv*. https://doi.org/10.31234/osf.io/6mfce

Woolnough, O., Donos, C., Curtis, A., Rollo, P. S., Roccaforte, Z. J., Dehaene, S., Fischer-Baum, S., & Tandon, N. (2022). A Spatiotemporal Map of Reading Aloud. *Journal of Neuroscience*, *42*(27), 5438–5450. https://doi.org/10.1523/JNEUROSCI.2324-21.2022

Woolnough, O., Donos, C., Rollo, P. S., Forseth, K. J., Lakretz, Y., Crone, N. E., Fischer-Baum, S., Dehaene, S., & Tandon, N. (2021). Spatiotemporal dynamics of orthographic and lexical processing in the ventral visual pathway. *Nature Human Behaviour*, *5*(3), 389–398. https://doi.org/10.1038/s41562-020-00982-w

Woolnough, O., & Tandon, N. (2024). Dissociation of reading and naming in ventral occipitotemporal cortex. *Brain*, *147*(7), 2522–2529. https://doi.org/10.1093/brain/awae027

Yeatman, J. D., & White, A. L. (2021). Reading: The Confluence of Vision and Language. *Annual Review of Vision Science*, *7*(1), 487–517. https://doi.org/10.1146/annurev-vision-093019-113509