Differential weighting of information during aloud and silent reading: Evidence from representational similarity analysis of fMRI data

Lyam M. Bailey[1], Heath E. Matheson[2], Jonathon M. Fawcett[2], Glen E. Bodner[3], Aaron J. Newman[1]

[1] Dalhousie University, Department of Psychology and Neuroscience, Halifax, NS B3H4R2, Canada

[2] Memorial University of Newfoundland, Department of Psychology, St. John's, NL A1B3X9, Canada

[3] Flinders University, College of Education, Psychology and Social Work, Adelaide, SA5001, Australia

Abstract

Single word reading depends on multiple types of information processing: readers must process low-level visual properties of the stimulus, form orthographic and phonological representations of the word, and retrieve semantic content from memory. Reading aloud introduces an additional type of processing wherein readers must execute an appropriate sequence of articulatory movements necessary to produce the word. To date, cognitive and neural differences between aloud and silent reading have mainly been ascribed to articulatory processes. However, it remains unclear whether articulatory information is used to discriminate unique words, at the neural level, during aloud reading. Moreover, very little work has investigated how other types of information processing might differ between the two tasks. The current work used representational similarity analysis (RSA) to interrogate fMRI data collected while participants read single words aloud or silently. RSA was implemented using a whole-brain searchlight procedure to characterize correspondence between neural data and each of five models representing a discrete type of information. Compared with reading silently, reading aloud elicited greater decodability of visual, phonological, semantic, and articulatory information. This occurred mainly in prefrontal and parietal areas implicated in speech production and cognitive control. By contrast, silent reading elicited greater decodability of orthographic information in right anterior temporal lobe. These results support an adaptive view of reading whereby information is weighted according to its task relevance, in a manner that best suits the reader's goals.

[Abstract word count: 230]

Abbreviations

DLPFC: dorsolateral prefrontal cortex

DMPFC: dorsomedial prefrontal cortex

Pre-SMA: pre-supplementary motor area

RSA: representational similarity analysis

SMA: supplementary motor area

VVP: ventral visual pathway

VWFA: visual word form area

1. Differential weighting of information during aloud and silent reading: Evidence from representational similarity analysis of fMRI data

1.1 Introduction

Single word reading is an automatic and effortless process for most literate individuals. From a cognitive perspective, however, word reading may be considered as a series of computations whereby a printed stimulus is mapped to cognitively relevant representations. In order to characterise the neural underpinnings of word reading, it is useful (and, indeed, commonplace in cognitive neuroscience) to decompose this complex process into discrete types of information processing; while recognizing insights from recent language models which reveal complex interactions between different levels of representation (e.g., Caucheteux et al., 2023; Henningsen-Schomers & Pulvermüller, 2022). In this spirit, we consider five broadly-defined types of information that may be extracted from a printed word: visual, orthographic, phonological, semantic, and (in the case of reading aloud) articulatory. While neural processes concerning each type of information have been studied extensively, it remains unclear how, and to what extent they might differ between reading aloud and reading silently.

This question is particularly relevant in the context of embodied and grounded cognition; that is, the perspective that our internal cognitive states are influenced (if not determined) by the manner in which we receive input from and interact with our environment (Matheson & Barsalou, 2018). From this perspective, viewing cognitive processes (such as reading) as being fundamentally experience-dependent is essential for understanding their neural substrates. In many respects, reading aloud and reading silently entail fundamentally different experiential (and, therefore, cognitive) states. First and foremost, reading aloud entails additional motoric / articulatory processes required for speech production. In turn, the physical act of articulation elicits sensory experiences: the sensation of moving one's mouth and tongue, feeling one's larynx vibrate, and hearing one's own voice. Moreover, reading aloud has been shown to confer a reliable memory advantage (MacLeod et al., 2010), suggesting differences in

cognitive processing beyond low-level motoric and sensorimotor mechanisms. This *production effect* is often attributed to distinctive sensory experiences brought about through articulation, though other mechanisms have been proposed. For example, there is evidence that words read aloud benefit from more elaborate semantic processing (Fawcett et al., 2022), which may facilitate encoding and retrieval. All of this is to say that differences between aloud and silent reading extend beyond the mere physical act of articulation, and entail extensive cognitive differences. Although some work has investigated this possibility from a cognitive-behavioural perspective, it remains unclear how such differences might manifest at the neural level.

This question may be addressed with neuroimaging. In particular, representational similarity analysis (RSA) is a technique that allows us to decode different types of information (defined by formal hypothesis models) in neural patterns. The purpose of the present study was to investigate the presence of information in neural patterns associated with different types of information present in printed words, by applying RSA to functional magnetic resonance imaging (fMRI) data acquired during aloud and silent reading.

Below we provide a brief overview of the functional neuroanatomy of single word reading, emphasising cognitive and neural processes associated with each type of information described above. This is intended to provide some context for RSA literature on single-word reading and, ultimately, the design of the current study.

1.2. Functional Neuroanatomy Of Single Word Reading

When presented with a printed or written word, an individual must process its low-level visual properties—shape and orientation of the constituent letter strokes, size, colour, etc. At the neural level, these perceptual processes are largely governed by primary and associative visual cortices housed in occipital cortex (Cornelissen et al., 2009; Gramfort et al., 2012; Hauk et al., 2012; Tarkiainen et al., 1999). Distinct from this low-level perceptual processing is orthographic processing—the recognition of multi-character strings as *visual word-forms*. Visual word-forms may be described as perceptually invariant mental representations of words,

irrespective of size, color, font, or position in the visual field (Warrington & Shallice, 1980). With respect to the neural correlates of orthographic processing, much emphasis has been placed on the visual word form area (VWFA), located in the temperooccipital portion of the left fusiform gyrus. Neuroimaging work has demonstrated that this area responds preferentially to both real words and orthographically regular pseudowords compared to irregular consonant-string pseudowords (Cohen et al., 2002; Petersen et al., 1990; Polk & Farah, 2002), and also to real words and consonant strings compared to false font or unknown character strings (Baker et al., 2007; Brem et al., 2010; Carreiras et al., 2014; Pleisch et al., 2019). These findings indicate that VWFA is sensitive both to letter strings generally, relative to perceptually similar non-letter visual objects, and also to orthographic regularity. In other words, it appears specialized for detecting visual stimulus properties that adhere to learned orthographic rules (i.e., statistics of written language), thus enabling recognition of familiar visual word forms. Moreover, studies on individuals suffering from pure alexia (selective impairment of the ability to recognise words rapidly and automatically, while identification of individual letters is preserved; McCandliss et al., 2003) often report damage to VWFA (Pflugshaupt et al., 2009; Turkeltaub et al., 2014), suggesting a causal role of this area in orthographic processing.

During reading, orthographic information is associated to an auditory phonological code—that is, a mental representation of the constituent sounds that make up the spoken form(s) of the word they are reading (Leinenger, 2014). The nature of phonological processing in reading has historically been contentious, with opposing connectionist (e.g., Seidenberg & McClelland, 1989) and dual-route (e.g., Coltheart et al., 2001) models describing different mechanisms (see Seidenberg et al., 2022 for discussion). Here we adopt the connectionist principle of weighted spelling-to-sound mapping (Seidenberg, 2005) — which has been integrated into more recent connectionist-dual-process (CDP) models of word reading, and whose purpose is to bridge the gap between the two accounts (e.g., the CPP++; Perry et al., 2010). From this perspective, phonological forms are computed based on weighted connections between orthographic units and sounds; the weightings themselves are determined by the

consistency of spelling-to-sound mappings within the language in question[1]. Words with inconsistent mappings may be considered to impose greater demands on the spelling-to-sound conversion system; hence, structures that are sensitive to spelling-sound consistency likely play a role in this process. A number of fMRI studies have revealed preferential activation of the left inferior frontal gyrus, and neighbouring structures such as anterior insula and anterior cingulate, when participants see words with inconsistent spelling-to-sound mappings relative to those with consistent mappings, implicating these structures in spelling-to-sound mapping (Bolger et al., 2008; Fiez et al., 1999; Fiez & Petersen, 1998; C.-Y. Lee et al., 2004).

Readers also experience semantic processing — that is, rapid and automatic retrieval of the meaning of the word they are reading. Based on a meta-analysis of 180 fMRI studies (Binder et al., 2009), Binder and Desai (2011) proposed a left-lateralized neurobiological model of semantic processing comprising dorsomedial and inferior prefrontal, inferior parietal, and inferior and ventral temporal cortices. Some of these same areas (particularly the inferior frontal gyrus, inferior parietal lobule, and inferior and superior temporal gyri) have also been identified by other meta-analyses of semantic processing (Rodd et al., 2015; Vigneau et al., 2006). In line with these meta-analyses, more recent evidence from RSA of fMRI data (described in more detail in 1.3) has supported the view that representation of semantic knowledge is distributed across bilateral prefrontal, parietal, and ventral temporal areas (e.g. Carota et al., 2021; Devereux et al., 2013; Liuzzi et al., 2020; Nastase et al., 2017).

In the case of reading aloud, an individual must engage additional articulatory (motor) processes, and will experience both proprioceptive feedback from moving one's mouth and tongue, and acoustic stimulation associated with hearing the sound of one's own voice. Indeed, a number of fMRI studies have reported that, compared to silent reading, reading aloud is associated with activation of auditory and sensorimotor cortices (Bailey et al., 2021; Dietz et al., 2005; Qu et al., 2022). Bailey et al. attributed their findings to motoric and sensory experiences involved in articulation, while Dietz et al. emphasised increased phonological processing

[1] For example, the vowel sound in *wave, cave, save,* etc. has high spelling-sound consistency in English, because *-ave* usually produces the eɪ vowel. By contrast, the vowel sound in *have* has low consistency because it does not obey this conventional spelling-sound mapping.

demands in the context of reading aloud. In addition, a meta-analysis by Murphy et al. (2019) identified a single cluster in left STG that responded preferentially to aloud compared to silent reading. Although the authors did not offer an interpretation of this finding, one might take left STG activation to reflect auditory processing while reading aloud, given that STG has been implicated in both basic auditory processing, and speech production specifically (Hickok & Poeppel, 2007; Scott et al., 2000).

1.3. Investigating Aloud And Silent Reading With RSA

Recent years have seen growing popularity of multivariate analysis methods for fMRI, which are all broadly concerned with the distributed patterns of activation elicited by stimuli. One such approach is RSA (Kriegeskorte et al., 2008), which allows us to characterise the informational content of these distributed activation patterns. Rather than making inferences based on contrasts between experimental conditions (as in univariate studies described above), RSA permits direct comparison between neural activation patterns and explicit hypothesis models, which often characterise specific stimulus information. Formally, this method quantifies the representational geometry of stimulus-specific activation patterns—that is, the geometric relationships between stimuli in high-dimensional space—as a *representational dissimilarity matrix* (RDM). This neural RDM may then be compared to one or more researcher-defined RDMs (hypothesis models) derived from quantitative measures of a given stimulus property. Within this framework, we can decode (i.e., detect) specific types of stimulus information (visual, orthographic, semantic, etc.) in a given patch of cortex, by demonstrating statistical dependence between a neural RDM (extracted from patterns in that patch) and a relevant hypothesis model (Kriegeskorte & Kievit, 2013).

Some prior RSA work has investigated decodability of different types of information during single word reading. Much of this work has taken a region-of-interest (ROI)-based approach to examine patterns in ventral temperooccipital cortex, which is largely specialized for visual object processing, and notably includes the VWFA. Studies in which participants see

visually presented words or logographic characters often report often report that phonological and orthographic information can be decoded in the left fusiform gyrus (Fischer-Baum et al., 2017; Qu et al., 2022; Zhao et al., 2017), while one study also decoded low-level visual information in its right-hemisphere homologue (Fischer-Baum et al., 2017). Semantic information may also be decoded in ventral temperoccipital cortices (Fischer-Baum et al., 2018; Wang et al., 2018), consistent with a larger body of work investigating semantic representation of images (as opposed to words) in this area (e.g., Devereux et al., 2013; Liuzzi et al., 2020; Nastase et al., 2017). Beyond ventral temporal cortex, phonological information has been decoded in left prefrontal and inferior parietal cortices (Fischer-Baum et al., 2018; Li et al., 2022, 2022), as well as middle and superior temporal regions (Li et al., 2022). Moreover, Zhang et al. (2020) had participants either read aloud, mouth, or imagine reading aloud simple consonant-vowel syllables, and reported that articulatory and acoustic information was decodable across a range of frontal, temporal, and parietal ROIs previously implicated in speech production and sensation. Taken together, these studies indicate that each type of information that we consider relevant to word reading may be decoded in multiple brain areas during single word reading.

Most of the studies described above focussed on aloud or silent reading tasks independently; a critical question, however, is whether the decodability of various types of information might *differ* between these two tasks. This question is particularly relevant to research on neural correlates of the production effect, which is defined as a contrast between aloud and silent reading. Moreover, from an embodied and grounded cognition perspective, this question is important for assessing the degree to which the neural correlates of reading are experience-dependent.

Decoding-based differences between aloud and silent reading would be consistent with prior RSA work showing that semantic decodability in particular is often task- or experience-dependent (e.g., when performing tasks that emphasise different semantic features of presented stimuli; Meersmans et al., 2022; Nastase et al., 2017; Wang et al., 2018). It is possible that decodability of other types of information is similarly variable when comparing

aloud to silent reading. Indeed, there are many reasons to expect such differences, outlined as follows. First and perhaps most obviously, one might expect greater dependence on (and therefore increased decodability of) articulatory information during aloud versus silent reading, because speaking a word necessarily requires participants to execute an appropriate sequence of articulatory movements. In this context, articulatory information should be decodable in primary motor and/or premotor cortices (i.e., precentral gyrus and supplementary motor area). Indeed, previous fMRI work has shown that vocalising different phonemes elicits discriminable responses in the precentral gyrus (Pulvermüller et al., 2006); this area may therefore have the capacity to represent whole words as linear combinations of articulatory features that are required to produce the constituent sounds.

Neural mechanisms distinguishing aloud and silent reading likely extend beyond the motoric components of articulation. Cognitive work surrounding the production effect—whereby words read aloud are more readily remembered compared to words read silently (MacLeod et al., 2010)—has emphasised the role of distinctive sensorimotor experiences during articulation. These include the sensation of moving one's mouth and tongue, feeling one's larynx vibrate, and experiencing auditory feedback from hearing one's own voice. At a higher cognitive level, reading aloud also entails planning and monitoring of speech output, and possibly sensory attenuation of efferent signals (corollary discharge) (Khalilian-Gourtani et al., 2022). At the neural level, we might expect such distinctiveness to be manifest as changes in decodability of articulatory and/or phonological information, perhaps in higher-level (i.e., non-sensorimotor) areas associated with speech planning and monitoring, or episodic encoding. Broadly speaking, the former processes have been linked to medial and lateral prefrontal cortices (Bourguignon, 2014; Hertrich et al., 2021), while episodic encoding is thought to depend largely on frontoparietal and medial temporal areas (e.g., H. Kim, 2011).

There is also evidence that reading aloud enhances semantic processing (Fawcett et al., 2022). We might therefore expect to see higher semantic decodability during aloud versus silent reading, possibly in ventral temporal areas (see 1.2). It has also been suggested that people allocate more attention to words that they read aloud (Fawcett, 2013; Mama et al., 2018;

Ozubko et al., 2012). While attention is not clearly defined in this context (particularly in terms of its relationship to specific types of information), we might broadly consider this term to reflect changes in one's cognitive state whereby words read aloud are assigned greater weighting (i.e., they are prioritised) during perceptual and goal-directed cognitive operations, compared to words read silently. This interpretation is consistent with prior conceptualizations of attention as the allocation of (limited) cognitive resources to a stimulus (P. A. MacDonald & MacLeod, 1998; Mama et al., 2018). Upon seeing a cue to read an upcoming word aloud, participants may experience an overall increased level of arousal as they prepare to map upcoming visual/orthographic information onto an appropriate vocal response. Increased arousal may lead to greater cognitive "investment" in processing multiple types of information (and therefore increased decodability of each), recognising that successful word production depends on accurate mapping of low-level visual information to orthography, phonology, and articulatory features.

The predictions outlined above have received partial support from two studies examining task-dependent changes in decodability during single word reading. Qu et al. (2022) reported that, relative to silent reading, aloud reading elicited greater decodability of orthographic and phonological information in the left anterior fusiform gyrus (Qu et al., 2022). Moreover, Zhang et al. (2020) presented participants with consonant-vowel syllables and reported that reading aloud, mouthing, and imagined reading aloud elicited differential correlations with an articulatory model in a number of areas implicated in speech production and sensation. Notably, articulatory information was decodable in left angular and superior temporal gyri when reading aloud, but not in the other two conditions. These two studies provide proof-of-principle that different reading tasks may modulate decodability of phonological, orthographic, and articulatory information. However, a number of questions still remain. Qu et al.'s (2022) analysis was confined to the fusiform gyrus; therefore, it is unclear how phonological and orthographic information (or, indeed, other types of information) might be differentially represented in other brain areas. Moreover, Zhang et al. (2020) did not explicitly compare aloud reading with silent reading (rather, these authors examined three

production tasks with variable degrees of enactment), and so it remains unclear how articulatory decodability might differ between the former conditions.

1.4. The Current Study

In the above section, we showed that different kinds of information—visual, orthographic, phonological, semantic, and articulatory—may be decoded during single word reading. There is some evidence for differential decodability of phonological and orthographic information between aloud and silent reading, but whether similar effects are present outside occipitotemporal cortex, or for other types of information, remains unclear. As such, the goal of the current work was to establish how reading words aloud versus silently affects the decodability of visual, orthographic, phonological, semantic, and articulatory information throughout the whole brain. We conducted an fMRI experiment in which participants saw visually presented words (each repeated multiple times throughout the experiment) and were instructed to read each word either aloud or silently. We used a searchlight procedure to generate a neural dissimilarity matrix centered on each voxel throughout the brain[2]; we then compared the neural data from each searchlight to hypothesis models representing each of the five types of information discussed above.

Given that articulatory information is essential for speech production, we predicted that, relative to silent reading, reading words aloud would result in increased articulatory decodability in primary and associative motor areas. Articulatory (and possibly phonological) information may also be present in frontoparietal and/or medial temporal areas associated with speech planning/monitoring and episodic encoding, consistent with accounts of the production effect which emphasise sensorimotor experiences during encoding of words spoken aloud. From the view that reading aloud enhances semantic processing, we predicted increased

[2] The searchlight method entails "scanning" a relatively large area (which might be the entire brain or a pre-defined ROI of any size) by parcellating it into a series of smaller searchlight areas, with each centered on a single voxel. An activation pattern for each searchlight area is extracted from its constituent voxels, meaning that one may construct a neural RDM (and, subsequently, compare that RDM to any number of hypothesis models) corresponding to every point in the brain/ROI (Kriegeskorte et al., 2008; Oosterhof et al., 2016).

semantic decodability relative to silent reading. While univariate studies have implicated multiple areas in semantic processing (see 1.2), changes in semantic decodability seems most likely in ventral temperooccipital cortex, as this area has been implicated by multiple RSA studies. Finally, increased attention to aloud words may lead to increases in decodability of multiple types of information. Given that ventral temperooccipital cortex (particularly the left fusiform gyrus) has attracted much attention concerning visual, orthographic, and phonological decodability, it seems plausible that such attentional effects would be present in this area. All of these predictions would be evidenced by higher correlations between the relevant hypothesis model(s) and neural patterns elicited by aloud, compared to silent reading.

2. Methods

2.1. Subjects

Data was collected from 30 participants, aged 18-40 (*M* = 21.43, *SD* = 4.57), 21 female, three left-handed. All participants reported normal or corrected-to-normal vision, proficiency in English, no history of neurological illness or trauma, and no contraindications to MRI scanning. Handedness information was obtained using the Edinburgh Handedness Inventory (Oldfield, 1971). Participants were recruited through on-campus advertising at Dalhousie University, and received $30 CAD reimbursement and a digital image of their brain. All data collection took place at the IWK Health Centre in Halifax, NS. All procedures were approved by the IWK Research Ethics Board. Participants provided informed consent according to the Declaration of Helsinki.

Four participants were excluded from data analysis: one reported difficulty reading visually presented words while in the scanner; one reported that they did not follow task instructions properly (they silently mouthed words in the aloud condition, instead of vocalising the words as instructed); one disclosed that they did not fit inclusion criteria (neurological normality) after having completed the study; one withdrew prior to completing all parts of the experiment. Therefore, data from 26 participants (18 female, two left-handed) were included in our analyses.

2.2. Stimuli And Apparatus

Participants viewed all stimuli while lying supine in the MRI scanner; stimuli were presented using the VisualSystem *HD* stimulus presentation system (Nordic Neuro Lab, Bergen, Norway). Stimuli were presented on an LCD screen positioned behind the scanner and viewed by participants via an angled mirror fixed to the MR head coil. During the experiment, participants made button-press responses using MR-compatible ResponseGrip handles (Nordic Neuro Lab, Bergen, Norway), one in each hand. Stimuli were presented using PsychoPy 2020.2.1

(Peirce et al., 2019). We selected a subset of 30 nouns from Bailey et al. (2021) which, in turn, were sourced from MacDonald and MacLeod (1998). Words were 6 to 10 characters in length, each with a frequency greater than 30 per million (Thorndike & Lorge, 1944). Our full word list is presented in Supplementary Table 1.

All words were presented at the center of the screen in white lowercase Arial font against a dark grey background (RGB: 128, 128, 128). Response instructions for each word (see Procedure) were grayscale icons presented at the start of each trial in the center of the screen—participants were instructed to speak the word aloud if they saw a mouth icon, or to read the word silently if they saw an eye icon. For our active baseline task (see Procedure), individual numbers 1-9 were presented in white Arial font at the center of the screen against the same grey background as the words.

2.1. Hypothesis Models

We constructed a set of hypothesis models used for RSA (see 2.6) reflecting visual, orthographic, phonological, semantic, and articulatory properties of the words presented in our experiment. Unique hypothesis models were generated for each participant and condition separately (because words were randomly allocated to conditions for each participant, see 2.3), in the Python environment using custom scripting and publicly available Python packages. Each hypothesis model comprised a 15 x 15 square dissimilarity matrix (DSM) containing zeros along the diagonal; off-diagonal cells contained pairwise dissimilarity values corresponding to all possible pairs of words within that condition, and matrices were symmetrical about the diagonal (Kriegeskorte et al., 2008). To generate each model, we computed dissimilarity values according to measures that were theoretically relevant to the respective type of information / stimulus property being examined. Each of these measures is briefly outlined below, with more details provided in Supplementary Materials. Hypothesis models from a representative subject and condition are shown in Figure 1.

We computed visual dissimilarity as the correlation distance (1 - Pearson correlation) between vectorized binary silhouette images of each word as it was presented in the

experiment (as per Kriegeskorte et al., 2008). This procedure was implemented using the Pillow package (Umesh, 2012). We computed orthographic dissimilarity as the correlation distance between unconstrained open bigram feature vectors (similar to Fischer-Baum et al., 2017), using the wordkit package (Tulkens et al., 2018). For phonology, words were vectorised according to their syllable-wise spelling-to-sound consistency (Chee et al., 2020), and dissimilarity was computed as the Euclidean distance between vectors (we found that using Euclidean distances produced overall more normally distributed hypothesis models, compared with other metrics such as correlation or cosine distance). While we are not aware of prior RSA work employing this measure, we suggest that spelling-sound consistency provides an appropriate measure of phonological processing as conceptualised in the Introduction. For semantic models, we computed cosine distances (1 - cosine similarity) between word2vec representations of each word's semantic content (consistent with previous work, e.g. Carota et al., 2021; Guo et al., 2022; Liu et al., 2021; Tong et al., 2022; Wang et al., 2018). Word2vec representations were acquired from the publicly available glove-wiki-gigaword-300 model (https://github.com/RaRe-Technologies/gensim-data) using the gensim package (Řehůřek & Sojka, 2010). Finally, we computed articulatory dissimilarity as the feature-weighted phonological edit distance (Fontan et al., 2016) between words, normalized for word length (Beijering et al., 2008; Schepens et al., 2012) using the PhonologicalCorpusTools package (Hall et al., 2019). While this is ostensibly a measure of phonological dissimilarity, it depends on articulatory features necessary to produce each word (e.g., place of articulation, movements of the tongue, teeth, and lips, etc.). As such, we feel that it provides a suitable means for modelling articulatory information. Moreover, our analyses required that our five measures be independent from one another (i.e., minimally correlated); as such, we were motivated to select measures of phonological and articulatory dissimilarity that were as different as possible from one another.

To ensure that our five measures were independent from one another, we computed pairwise correlations between (vectorised) exemplar models. Exemplar models contained all words in our stimulus set, and were computed for each measure. We reasoned that, because

stimulus allocation to each condition was randomised for each participant, comparing exemplar models (containing all possible stimulus combinations) was the best way to approximate potential correlations between models used in our analysis. Correlations between our five exemplar models are displayed in Table 1. Accompanying parenthetical values in Table 1 show Bayes factors BF_{10} (outlined in 2.6.2) for each correlation, computed using the bayesFactor package (Krekelberg, 2022) in the MATLAB environment, using default JZS priors (Rouder et al., 2009). These comparisons revealed moderate evidence for a true correlation between the phonological and semantic models (r = 0.106, BF_{10} = 3.38); all other comparisons revealed equivocal evidence ($BF_{10} < 2.0$).

Table 1. Correlation matrix for the five dissimilarity measures used in this study. Values are correlation coefficients for each pair of measures; values in parentheses are Bayes factors indicating strength of evidence of a true (anti-)correlation.

	Visual	**Orthographic**	**Phonological**	**Semantic**	**Articulatory**
Visual		-0.037 (0.051)	-0.043 (0.057)	0.03 (0.046)	> 0.001 (0.038)
Orthographic			-0.028 (0.045)	0.106 (0.449)	0.134 (1.934)
Phonological				-0.143 (3.382)	0.118 (0.792)
Semantic					0.01 (0.039)
Articulatory					

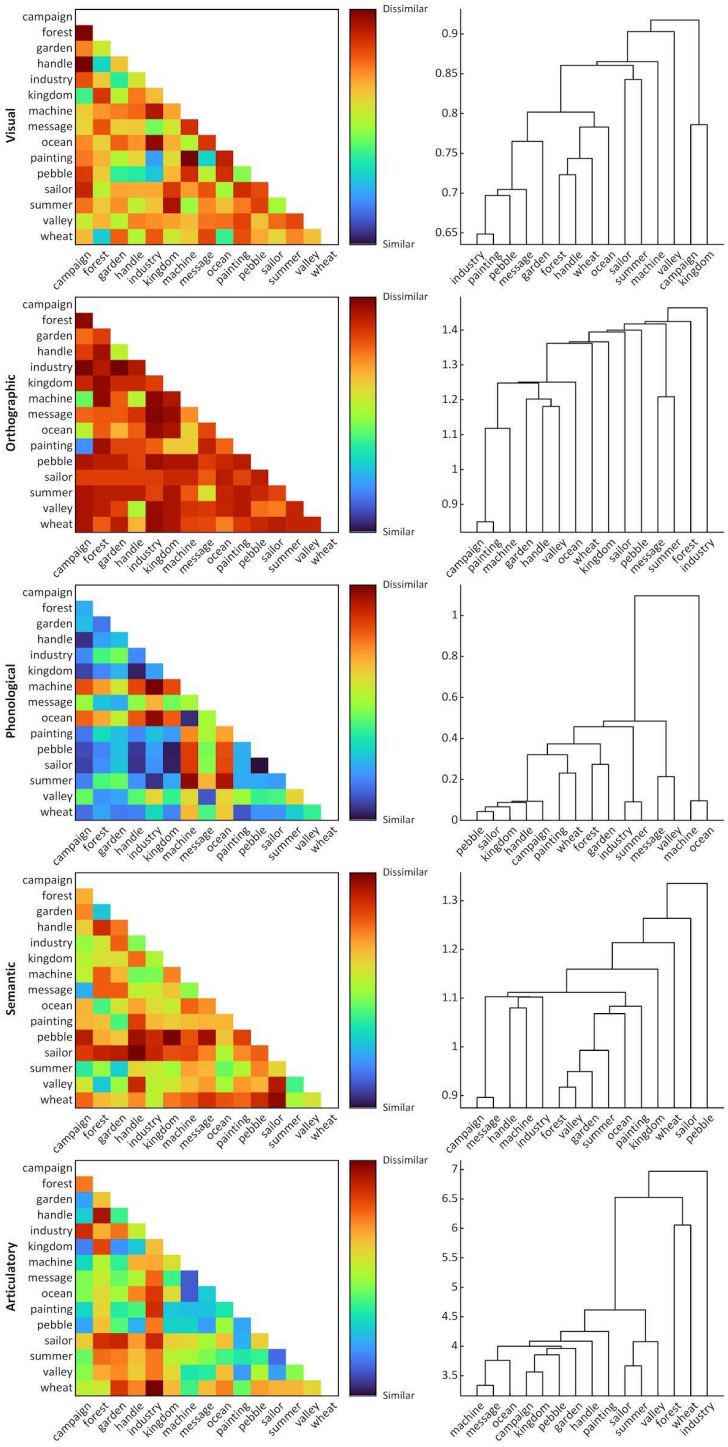


Figure 1. Hypothesis models (i.e., dissimilarity matrices) from a representative subject and condition

2.3. Procedure

After providing informed consent and completing MRI safety screening and the Edinburgh Handedness Inventory, participants completed a shortened version of the experiment described below (using a different stimulus list) on a laptop computer to familiarize them with the task. Participants were also asked to confirm (verbally) that they understood the task requirements before entering the scanner. Once participants were positioned in the MRI scanner, a brief scout scan determined head position. Following this, participants completed three functional scans as part of another study, followed by a structural scan. After the structural scan, participants completed four three-minute functional scans that comprised the data for the present study.

During each three-minute scan, participants performed a word reading task, in which the 30 words (S Table 1) were randomly allocated to one of two conditions (aloud or silent; 15 words in each condition) for each participant. Word-condition mappings were randomised between participants. Each trial of the word reading task began with a fixation cross ("+") presented for 500 ms to alert the participant that the trial was starting, followed by a cue presented for 1000 ms instructing participants to read the upcoming word either aloud (if they saw a mouth icon) or silently (if they saw an eye icon). Following the cue, a word from the corresponding condition was presented for 2500 ms. Each word appeared once in each functional run, and each of the four runs used a different random order of presentation.

On every trial, following word presentation, participants completed a short active baseline task: a randomly generated number between 1 and 9 was presented for 2000 ms, and participants were instructed to decide whether the number was odd or even, and to make an appropriate button-press response. During this baseline period, response mapping cues were presented as a reminder in the top corners of the screen (left-hand button press for odd, right-hand for even). Response mappings were the same for all participants. This active baseline served two purposes: (i) to lengthen SOA of each trial to 6 seconds, and therefore ensure reliable item-level estimation (Zeithamova et al., 2017), while (ii) ensuring that participants did not mind-wander between trials. We selected a simple odd-even judgement task because prior

work has shown this task to be an effective alternative to the traditional resting baseline in fMRI (Stark & Squire, 2001). A schematic of our trial structure is shown in Figure 2. Each functional run included 30 trials.

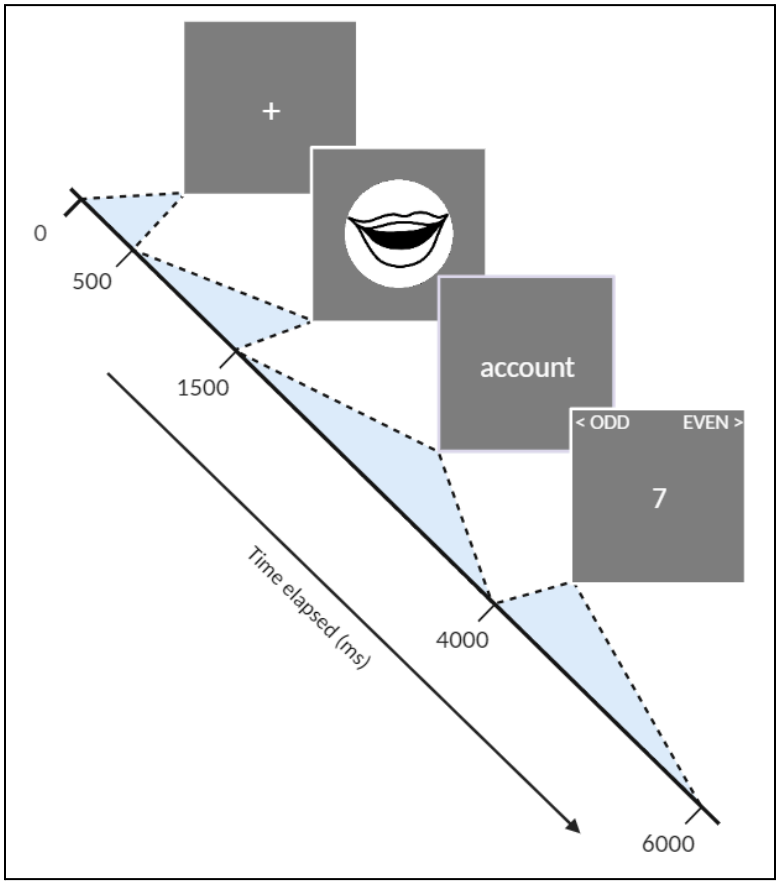


Figure 2. Schematic of an example trial

2.4. MRI Data Acquisition

MRI data was acquired on a 1.5 Tesla GE MRI system (GE Medical Systems, Waukesha, WI) equipped with a 19 channel head coil. Each participant completed a structural scan followed by four functional scans. As noted above, participants also completed three additional functional scans at the start of their session, but those functional scans are beyond the scope of the current work. For the structural scan, a T1-weighted anatomical image was obtained using a

magnetization-prepared rapid acquisition gradient echo (MPRAGE) sequence, TI = 1134 ms, flip angle = 8°, NEX = 2, FOV = 224 mm, matrix = 224 x 224, resulting in an in-plane voxel resolution of 1 x 1 mm. Functional scans used a gradient echo-planar pulse sequence, TR = 1800 ms, TE = 23 ms, FOV = 240 mm, flip angle = 90°. Images were obtained in 34 axial slices[3] (no gap, sequential acquisition) of thickness = 3.75 mm, matrix = 64 x 64, resulting in an in-plane voxel resolution of 3.75 x 3.75 mm. The FOV included full cortical coverage and partial cerebellar coverage. For each run we collected 100 functional volumes. Five additional volumes were acquired at the start of each run, but were discarded following acquisition.

2.5. fMRI Data Processing

All fMRI data processing was implemented with custom bash scripting, unless otherwise stated. To improve efficiency of our analysis pipeline, we used GNU Parallel (Tange, 2011) to perform many of the steps described below in parallel for multiple subjects/runs. The fMRI data were processed using functions from FEAT (fMRI Expert Analysis Tool) Version 6.00, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). Preprocessing steps included non-brain removal with BET (Smith, 2002), grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor, high pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 50.0 s), and spatial smoothing (using a Gaussian kernel of FWHM 6 mm). We also performed motion correction using MCFLIRT (Jenkinson et al., 2002) to mitigate the impact of any potential movement during functional runs. We visually inspected motion correction output; our a-priori threshold for excessive motion between contiguous time points was 2 mm, however no runs exceeded this threshold, therefore no runs were removed due to excessive motion.

Following preprocessing, we used FEAT to estimate activation patterns for each trial using the least-squares-all (LSA) method. For each functional run, we constructed a single GLM that included one regressor for each trial in the run. Each regressor comprised a timeseries

[3] We added 4 axial slices (total = 38) to the protocol for one participant, in order to accommodate their entire cerebral cortex.

modelling the period during which the word for that trial was presented (duration: 2500 ms), convolved with a gamma function (lag: 6 sec, sigma: 3 s) as a model of the haemodynamic response function (HRF). Additionally, we included parameters derived from motion correction as regressors of no interest. We defined 30 contrasts of interest, with each contrast comprising one regressor. This procedure generated four contrasts of parameter estimates (i.e., COPE images) for every unique item; one per trial. Because functional data from different runs are rarely perfectly aligned (due to small head movements between runs), we next aligned all the COPEs from each subject to a common native functional space. We used Advanced Normalization Tools (ANTs; http://stnava.github.io/ANTs/) to rigidly align COPEs from the second, third and fourth runs with the example functional volume (example_func in FSL) from the first fMRI run. We additionally masked the aligned data with a COPE image from the first run, thus removing non-overlapping voxels across runs.

Item-level pattern estimation (described as follows) was performed in the MATLAB environment using functions from the CoSMoMVPA toolbox (Oosterhof et al., 2016). We estimated item-level activity patterns for each unique item by averaging across COPEs from that item's four respective trials. We then subtracted the mean pattern (i.e., the mean value at each voxel across items) from item-level patterns in each condition separately, in order to remove the activation pattern common to all items in each condition (Walther et al., 2016)[4].

[4] While mean pattern subtraction is a contentious topic (Diedrichsen et al., 2011; Garrido et al., 2013), we argue that it is necessary in the case of our study. Shared mean activation has been shown to artificially inflate pairwise correlations between patterns (Walther et al., 2016); this in turn will add noise to correlation-based neural dissimilarity matrices. This has major implications for our study, the purpose of which was to compare decodability (that is, correspondence between neural DSMs and hypothesis models) between aloud and silent reading. Aloud reading reliably elicits more univariate activation than silent reading (Bailey et al., 2021; Dietz et al., 2005; Qu et al., 2022); in turn, item-level patterns for words read aloud likely share more common activation than do patterns for words read silently. As a result, we ought to see relatively greater inflation of pairwise correlations between patterns for words read aloud. Stated differently, failing to control for shared activation (within each condition) would likely result in systematically different degrees of noise contributing to decodability in aloud and silent reading, which would obfuscate any comparisons of decodability between these conditions.

2.6. Representational Similarity Analysis

2.6.1. Searchlight Analyses

We performed RSA using a whole-brain searchlight approach, separately for each subject and condition (aloud or silent reading), in each subject's native functional space. Searchlights were implemented in the MATLAB environment using functions from the CoSMoMVPA toolbox (Oosterhof et al., 2016). Within each searchlight area (spherical searchlight, radius = 3 voxels), we first generated a neural DSM comprising pairwise Pearson correlation distances (i.e., 1 minus the Pearson correlation) between item-level patterns for all items within that condition. Each neural DSM was vectorized and submitted to a GLM with five regressors, with each regressor comprising one of the five (vectorized) hypothesis models described above. Both the neural data and models were standardized using the *z* transform prior to estimating regression coefficients (Oosterhof et al., 2016). Each searchlight analysis generated five whole-brain statistical maps with a beta (β) value at each voxel; one map was generated for each hypothesis model. β values reflected correspondence between the neural DSM at that voxel (computed from patterns in the corresponding searchlight sphere) and the respective hypothesis model. A separate searchlight analysis was performed for each condition, such that each subject was associated with ten β maps in native functional space.

We next transformed each subject's searchlight output (i.e., whole-brain β maps) from native functional space to template MNI152 space using ANTs (http://stnava.github.io/ANTs/). We first computed a structural-to-EPI transformation matrix by rigidly aligning participants' high-resolution structural T1 image to their example functional volume from the first fMRI run. We next computed an MNI152-to-structural transformation matrix using linear affine, and then nonlinear methods (the latter implemented using the "SyN" algorithm: Avants et al., 2008)[5]. We then applied the inverse of the structural-to-EPI and MNI152-to-structural transformation

[5]To be clear: the structural-to-EPI and MNI152-to-structural transformation matrices were computed using the EPI and structural T1 as reference images respectively. While it is more common to compute these transformations in the opposite direction (i.e., EPI-to-structural and structural-to-MNI152, using structural T1 and MNI152 as reference images), we found that the inverse procedure yielded qualitatively better transformation of searchlight maps—that is, better alignment to the MNI152 template, based on visual inspection of registration output.

matrices to each subject's respective searchlight output, re-sliced to 2 mm isotropic resolution. Each subject was therefore associated with ten β maps in MNI152 space. Following this, we masked out the cerebellum (as defined by the Harvard-Oxford cortical atlas supplied by FSL) from all of the spatially normalized β maps; the motivation for this was that our fMRI scans included only partial cerebellar coverage.

2.6.2. Group-Level Analyses

For group-level analyses, we computed voxel-wise Bayes factors BF_{10} to evaluate our hypotheses. This is in contrast to conventional null-hypothesis statistical testing (e.g., permutation cluster statistics), which relies on *p* values to determine whether or not a null hypothesis may be rejected at each voxel. Unlike *p* values, Bayes factors quantify the strength of evidence in favour of a null or alternative hypothesis. BF_{10} is a ratio that expresses the likelihood of the data under the alternative hypothesis (H_1) relative to the null hypothesis (H_0) (Dienes, 2014), and may range between 0-∞. Within this framework, incrementally larger values > 1 indicate greater support for the data under H_1; incrementally lower values < 1 indicate greater support for the data under H_0. Recently, a growing number of studies have adopted Bayes factors for group-level neuroimaging analyses, particularly in the context of MVPA (e.g., Grootswagers, Robinson, & Carlson, 2019; Grootswagers, Robinson, Shatek, et al., 2019; Kaiser et al., 2018; Matheson et al., 2023; Moerel et al., 2022; Proklova et al., 2019; Teichmann et al., 2021). Teichmann et al. (2021) and Matheson et al. (2023) have argued in favour of Bayes factors over *p* values in MVPA research, primarily because Bayes factors are actually informative as to strength of evidence (as opposed to dichotomous acceptance or rejection of H_1), and are not susceptible to multiple comparison problems (Dienes, 2016; Teichmann et al., 2022).

We computed Bayes factors using voxel-wise Bayes *t*-tests implemented with functions from the bayesFactor package (Krekelberg, 2022) in the MATLAB environment, using default JZS priors (Rouder et al., 2009). We first tested whether each type of information was decodable in either condition. Here, we performed a right-tailed one-sample Bayes *t*-test at every voxel against the hypothesis that β was greater than zero; this procedure was repeated independently for each condition and type of information. This produced ten statistical maps of BF_{10} values (2

conditions x 5 information types); heron in, we will refer to these as *within-condition* Bayes maps. The results of these analyses are not reported here; rather, they were used to constrain results from contrasts between aloud and silent reading, described as follows.

We next tested the hypothesis that there was a difference in decodability, for each type of information, between conditions. Here, we performed a two-tailed paired-samples Bayes *t*-test at every voxel against the hypothesis that β values differed between aloud and silent reading. This procedure was repeated independently for each type of information, generating five statistical maps of BF_{10} values; hereon in, *between-condition* Bayes maps.

As a convention, Bayes factors are often assigned qualitative labels to express the strength of evidence for H_1, whereby BF_{10} values < 3.0 are considered to provide "weak" or "anecdotal" evidence, values < 10.0 provide "moderate" evidence, and values >= 10.0 are considered to provide "strong" evidence (Dienes, 2014; Lee & Wagenmakers, 2014). We chose to only consider voxels where there was at least moderate evidence for any hypothesis (that is, decodability greater than 0, or differences between conditions); we feel that reporting weak evidence would not be particularly valuable, and moreover would distract from data points where there was relatively stronger support for our hypotheses. As such, we thresholded all Bayes maps (both within- and between-conditions) at BF_{10} >= 3.0. To establish directionality of our results concerning differences between conditions, we first computed (for each type of information) two average contrast maps: an aloud > silent contrast representing the average searchlight results (across participants) from the aloud condition minus those of the silent condition; vice-versa for the silent > aloud contrast. These average contrast maps were used to mask the thresholded between-condition Bayes maps, thus generating one Bayes map representing aloud > silent, and one representing silent > aloud, for each type of information. Each contrast map was then masked by the within-condition map for its respective minuend condition; this ensured that the contrast maps only included voxels where information was actually decodable in the minuend condition (i.e., it served to exclude any between-condition differences driven entirely by negative values in the subtrahend condition).

We generated interpretable tables of results using FSL's cluster function to identify clusters of contiguous voxels, separately for each model and between-conditions contrast. We only considered clusters with a minimum spatial extent of 20 voxels, to avoid our results being skewed by noisy voxels. Tables report the spatial location (in MNI coordinates) of the center of gravity (COG) for each cluster, as well as each cluster's mean (averaged over constituent voxels) and maximum BF_{10} value. Maximum BF_{10} values represent the strongest available evidence for a given hypothesis (e.g., for a difference between conditions) within a cluster, while we consider mean values to reflect evidential strength across the entire cluster. Tables also report anatomical labels for each cluster; these were identified from the Harvard-Oxford cortical atlas using FSL's atlasquery function.

3. Results

We performed RSA on fMRI data to investigate decodability of information across five types of information—visual, orthographic, phonological, semantic, and articulatory—during aloud and silent word reading. We used a whole-brain searchlight to characterise, at every voxel, correspondence between the neural data and formal hypothesis models representing each type of information, for each condition separately. We then explored differences in decodability between the two conditions. Below, all references to "moderate" or "strong" evidence describe quantitative benchmarking of mean BF_{10} values that were computed for each cluster, based on conventional norms (Dienes, 2014; Lee & Wagenmakers, 2014).

3.1. Aloud Reading > Silent Reading

Visual, phonological, semantic, and articulatory—but not orthographic—information were all more decodable in the aloud reading condition compared to the silent reading condition; these results are displayed in Table 2 and Figure 3. For convenience, all references to decodability in this section describe differences between aloud and silent reading.

3.1.1. Visual Information

We identified a single cluster exhibiting moderate evidence for visual decodability. This cluster was situated in the occipital portion of the left fusiform gyrus (more precisely, on the fusiform bank of the posterior collateral sulcus; Lehman et al., 2016) and extended into lateral occipital cortex.

3.1.2. Phonological Information

Decodability of phonological information was notably more extensive than visual, semantic, or articulatory information, and was detected in bilateral frontal and parietal areas. In the frontal lobes, clusters were situated bilaterally on the ventrolateral / orbital surfaces of the frontal poles; notably, the cluster in the left frontal pole exhibited the strongest evidence for this contrast (mean BF_{10} = 172.96). In addition, a single cluster with strong evidence spanned

prefrontal cortices in both hemispheres, encompassing dorsolateral (middle and superior frontal gyri, frontal pole) and dorsomedial (medial surfaces of the superior frontal gyrus and frontal pole) areas bilaterally. One additional cluster was situated on the lateral surface of the right precentral gyrus, though this cluster exhibited only moderate evidence.

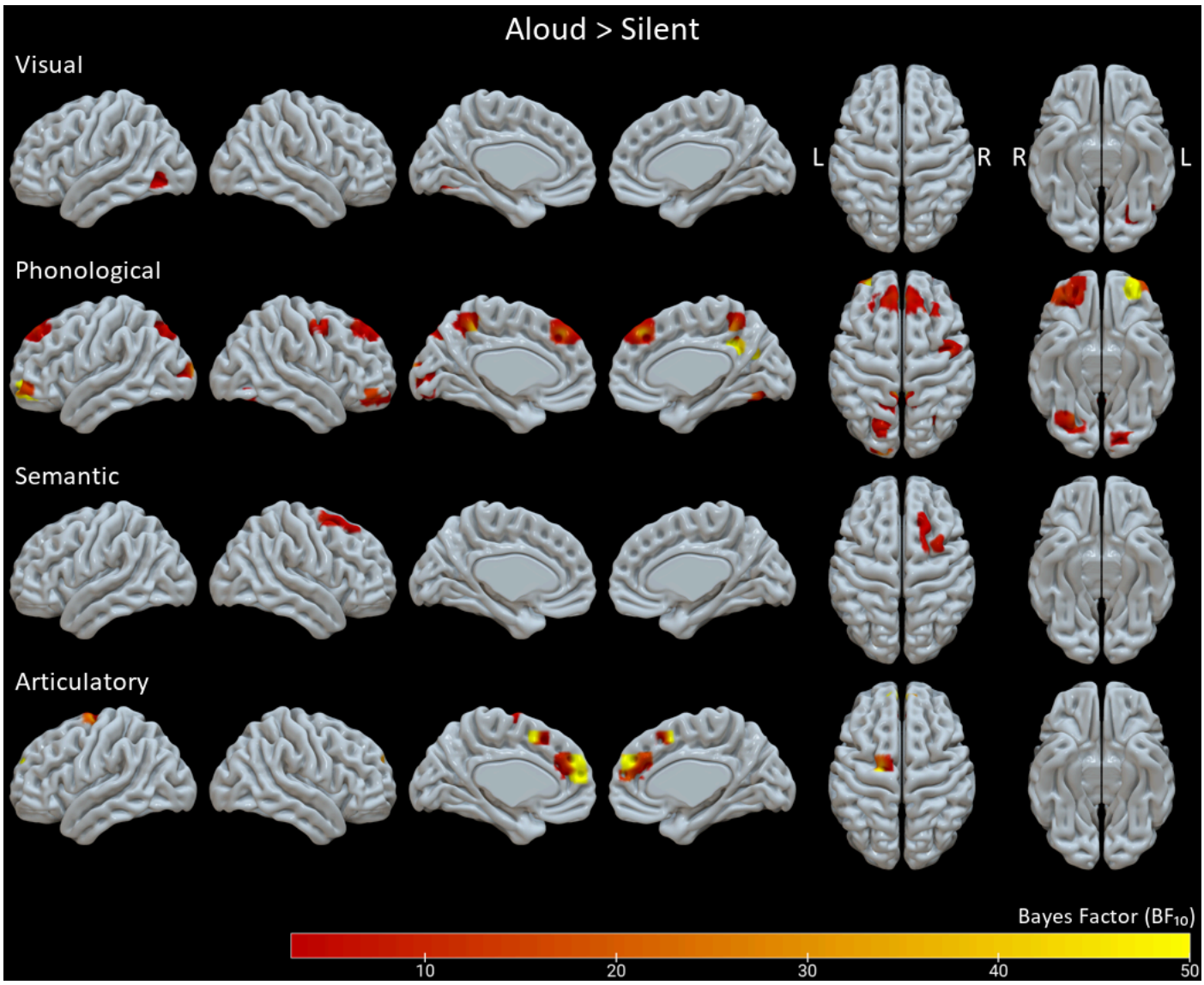


Figure 3. BF_{10} maps for the Aloud > Silent contrasts. Highlighted areas show evidence of greater decodability in the aloud reading condition relative to the silent reading condition. Surfaces from left to right show left lateral, right lateral, left medial, right medial, dorsal bilateral, and ventral bilateral views.

In the parietal lobes, clusters were situated in the precuneus bilaterally, though the right cluster exhibited markedly stronger evidence (mean BF_{10} = 68.24 [right], 15.71 [left]) and was situated more posteriorly and ventrally, at the parieto-occipital boundary. Clusters were also present in the occipital lobe. On the lateral surface of the left hemisphere, one strong cluster was situated at the parieto-occipital boundary (mainly encompassing lateral occipital cortex) and, more posteriorly, a moderate cluster extended from the lateral surface of the occipital pole to posterior ventral occipital cortices. Finally, one cluster with moderate evidence was present in the occipital portion of the right fusiform gyrus.

3.2.3. Semantic Information

We identified two adjacent clusters in right prefrontal cortex (middle and superior frontal gyri), both of which exhibited moderate evidence for semantic decodability. The position of these clusters approximately corresponds to dorsal premotor cortex (Genon et al., 2017).

3.2.4. Articulatory Information

Articulatory information was decodable in medial frontal areas bilaterally, as well as in left premotor cortex. On the medial surface, two large clusters encompassed areas in both hemispheres. The anteriormost cluster encompassed anterior dorsomedial cortex (superior frontal and paracingulate gyri); the more posterior cluster was situated in the pre-supplementary motor area (pre-SMA). On the lateral surface, one cluster was situated in the posteriormost portion of the superior frontal gyrus; corresponding to the supplementary motor area proper (SMA). All of these clusters exhibited strong evidence.

Table 2. RSA results table detailing clusters identified by the Aloud > Silent contrasts. Clusters show evidence of greater phonological, semantic, and articulatory decodability in the aloud reading condition relative to the silent reading condition.

Type of information	Cluster #	Cluster extent (mm^3)	Mean BF	Max BF	x	y	z	Anatomical label(s)
Visual	1	58	9.41	29.9	-33	-69	-5	**L Occipital Fusiform Gyrus (36)**
								L Lateral Occipital Cortex inferior (22)
Phonological	1	118	172.96	882	-27	53	-8	**L Frontal Pole (118)**
	2	212	68.24	542	20	-55	29	**R Precuneus Cortex (141)**
								R Cingulate Gyrus posterior (29)
								R Cuneal Cortex (2)
								R Lateral Occipital Cortex superior (22)
								R Superior Parietal Lobule (2)
	3	116	15.71	57.3	-3	-48	54	**L Precuneus Cortex (83)**
								L Postcentral Gyrus (2)
								L Superior Parietal Lobule (1)
								R Precuneus Cortex (30)
	4	164	12.19	37.8	-18	-65	45	**L Lateral Occipital Cortex superior (136)**
								L Precuneus Cortex (24)
								L Superior Parietal Lobule (4)
	5	863	11.69	93.2	-1	39	46	**L Superior Frontal Gyrus (297)**
								L Frontal Pole (201)
								L Paracingulate Gyrus (2)
								R Frontal Pole (124)
								R Middle Frontal Gyrus (12)
								R Paracingulate Gyrus (5)
								R Superior Frontal Gyrus (220)
	6	140	10.41	23.2	29	48	-14	**R Frontal Pole (140)**
	7	190	9.02	35.1	-19	-90	5	**L Occipital Pole (90)**
								L Intracalcarine Cortex (23)

								L Lateral Occipital Cortex inferior (15)
								L Lateral Occipital Cortex superior (13)
								L Lingual Gyrus (9)
								L Occipital Fusiform Gyrus (40)
	8	22	8.29	20.3	28	-72	-14	**R Occipital Fusiform Gyrus (22)**
	9	61	6.51	13.3	44	-2	50	**R Precentral Gyrus (60)**
								R Middle Frontal Gyrus (1)
Semantic	1	102	8.58	24.6	30	9	65	**R Superior Frontal Gyrus (61)**
								R Middle Frontal Gyrus (41)
	2	33	3.97	5.21	24	28	59	**R Superior Frontal Gyrus (32)**
								R Middle Frontal Gyrus (1)
Articulatory	1	368	36.71	294	0	49	25	**L Paracingulate Gyrus (193)**
								L Frontal Pole (5)
								L Superior Frontal Gyrus (27)
								R Paracingulate Gyrus (115)
								R Superior Frontal Gyrus (28)
	2	33	25.32	81.4	-19	-7	72	**L Superior Frontal Gyrus (33)**
	3	27	23.6	96	-3	15	52	**L Paracingulate Gyrus (12)**
								L Superior Frontal Gyrus (13)
								R Paracingulate Gyrus (1)
								R Superior Frontal Gyrus (1)

3.2. Silent Reading > Aloud Reading

Only orthographic information showed greater decodability in the silent reading condition compared to the aloud condition; these results are displayed in Table 3 and Figure 4. This contrast revealed two clusters with strong-to-moderate evidence (mean BF_{10} = 11.32, 8.93 respectively) in the right anterior temporal lobe. The stronger cluster spanned a relatively large area of anterior ventral temporal cortex, with the peak value (max BF_{10} = 132) situated at the tip

of the temporal pole, extending posteriorly into (anterior portions of) the fusiform and parahippocampal gyri. The second cluster was on the lateral surface, at the border between the temporal pole and superior temporal gyrus.

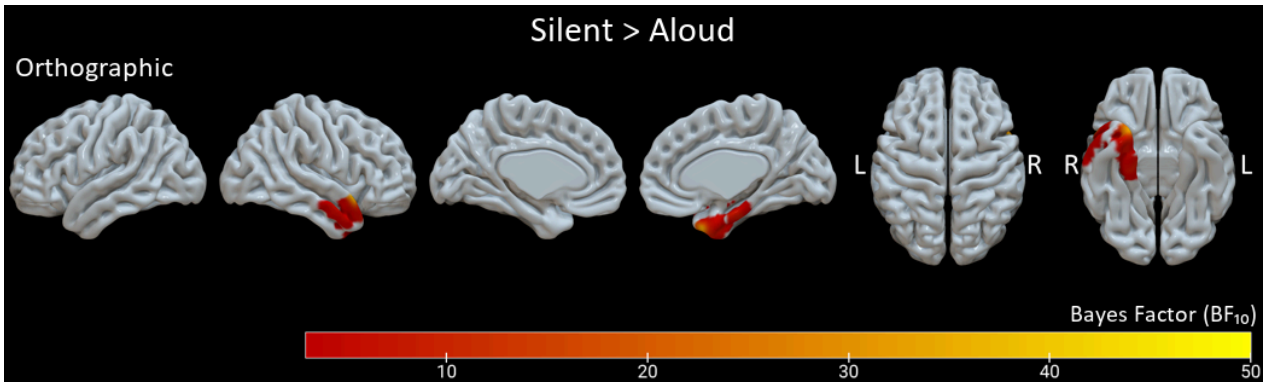


Figure 4. BF_{10} map for the Silent > Aloud contrast. Highlighted areas show evidence of greater decodability in the silent reading condition relative to the aloud reading condition. Surfaces from left to right show left lateral, right lateral, left medial, right medial, dorsal bilateral, and ventral bilateral views.

Table 3. RSA results table detailing clusters identified by the Silent > Aloud contrast. Clusters show evidence of greater orthographic decodability in the silent reading condition relative to the aloud reading condition.

Type of information	Cluster #	Cluster extent (mm^3)	Mean BF	Max BF	x	y	z	Anatomical label(s)
Orthographic	1	357	11.32	85	27	7	-38	**R Temporal Pole (186)**
								R Inferior Temporal Gyrus anterior (4)
								R Parahippocampal Gyrus anterior (94)
								R Parahippocampal Gyrus posterior (3)
								R Temporal Fusiform Cortex anterior (70)
								Right Hippocampus (16)
	2	252	8.93	132	56	12	-21	**R Temporal Pole (179)**
								R Middle Temporal Gyrus anterior (58)
								R Middle Temporal Gyrus posterior (15)

4. Discussion

The current work aimed to characterise differences in the decodability of five types of stimulus-relevant information between aloud and silent single-word reading. We used RSA to compare fMRI data (acquired during each reading task) to formal hypothesis models representing visual, orthographic, phonological, semantic, and articulatory information. Our results revealed differential decodability for all types of information between the two tasks. Interestingly, and contrary to our initial predictions (see Introduction), we did not find evidence that decodability is unilaterally enhanced by aloud reading. Instead, visual, phonological, semantic, and articulatory information were all more decodable during aloud reading, while silent reading entailed greater decodability of orthographic information. We interpret these results to reflect differential weighting of information, depending on cognitive demands imposed by either task.

4.1. Decoding Information In Aloud Reading

We found moderate evidence that visual information was more decodable during aloud versus silent reading, in a cluster spanning the posterior portion of the left collateral sulcus and inferior lateral occipital cortex. In the context of reading, these posterior occipital areas have been associated with low-level orthographic analysis (Jobard et al., 2003; Levy et al., 2009); for example, they tend to respond preferentially to consonants versus false fonts (Thesen et al., 2012; Vinckier et al., 2007). We therefore suggest that our findings concerning visual information reflect greater attention (that is, allocation of more cognitive resources) to low-level perceptual properties of the visually presented words, consistent with the notion that participants allocate more attention to words spoken aloud (Fawcett, 2013; Mama et al., 2018; Ozubko et al., 2012).

Phonological and articulatory information were both decodable in dorsomedial prefrontal cortex (DMPFC). In addition, articulatory information was present in (pre-)supplementary motor areas (pre-SMA, SMA), while phonological information was present

in dorsolateral prefrontal cortex (DLPFC) and the frontal poles. DMPFC, DLPFC, and (pre-)SMA are functionally connected to each other and to major language processing centers in perisylvian cortex, and are all associated with planning and cognitive control of speech processes, amongst other functions (Bourguignon, 2014; Hertrich et al., 2016, 2021). Left DLPFC in particular is commonly associated with speech planning (see Hertrich et al., 2021 for review), while DMPFC has been linked to domain-general online response monitoring. For example, in the context of a Stroop task, DMPFC may monitor (and resolve) conflicting information arising on a trial-to-trial basis (Freund et al., 2021; C. Kim et al., 2013; A. W. MacDonald et al., 2000). With respect to its role in language production, a recent study linked DMPFC to learning arbitrary associations between visually presented stimuli and orofacial and vocal responses, as well as physically performing those responses (Loh et al., 2020). These authors also note that pre-SMA was recruited for learning speech-based vocal responses, but not non-speech vocalizations or orofacial movements, suggesting some degree of specialization for articulatory processes. Pre-SMA has otherwise been associated with selecting and encoding complex motor sequences (particularly in the context of speech production tasks; Alario et al., 2006; Tremblay & Gracco, 2009, 2010), while SMA proper is thought to play a role in their implementation (Hertrich et al., 2016). As for the frontal pole, while its specific role in language production is less clear, it has been suggested that this region plays a domain-general role in goal-dependent feedback monitoring (Tsujimoto et al., 2011) and goal selection and maintenance (Fine & Hayden, 2021; Koechlin, 2011).

Given the above information, we suggest that the presence of phonological and articulatory information in the aforementioned areas reflects top-down planning and maintenance of vocal responses. To speak each word correctly, participants had to plan and coordinate an appropriate articulatory response comprising a unique sequence of movements of the mouth, jaw, and tongue. The presence of articulatory information in pre-SMA and SMA likely reflects preparation and implementation (respectively) of this complex motor sequence, while DMPFC may be involved in on-line monitoring to ensure that the produced responses match computed phonological information (see next paragraph). Thus while it is unsurprising

that a vocal response should involve motor planning and monitoring, our results demonstrate that *stimulus-specific* articulatory information can be decoded in these areas. As such, it is accurate to say that responses in SMA, pre-SMA, and DMPFC reflect unique articulatory properties of specific words.

With respect to phonological information (which in our study concerned spelling-to-sound consistency), executive control processes may be recruited to ensure that visually presented words are converted to appropriate phonological codes for the purpose of reading aloud. These control processes likely include speech planning (DLPFC), online monitoring during speech production (DMPFC) and goal-related cognition (frontal poles). With respect to the latter point, phonological information might be relevant to goal maintenance (e.g., updating one's current goal—to read the presented visual word aloud—to incorporate the desired sounds) or goal-related feedback (e.g., registering that the perceived visual word was produced as planned). As an aside, our interpretations here may seem at odds with the view that whole phonological forms of familiar/high frequency words (such as our stimuli) are retrieved automatically, without the need for explicit spelling-to-sound mapping (Coltheart et al., 2001). Direct phonological retrieval may have occurred in our study, but been supplemented by additional spelling-to-sound conversion, perhaps as an internal "check" to ensure that the appropriate sounds were retrieved. This suggestion is consistent with the view that the degree to which readers engage in explicit spelling-to-sound conversion is dependent upon their goals and/or task demands (Harm & Seidenberg, 2004). In our case, checks on phonological output may be driven by participants' goal to read the words aloud.

Phonological information was also present in the precuneus. Notably, the area identified by our analysis appears to correspond to part of the paracingulate network (Dadario & Sughrue, 2023). This network has been proposed to support integration of external input and prior knowledge to guide goal-directed behaviour, via its functional connections to other areas. These other areas notably include the insula and anterior cingulate (both implicated in spelling-to-sound conversion; see Introduction), medial frontal areas (possibly involved in monitoring articulatory responses, see above), and sensorimotor cortices (necessary for

implementing articulatory responses; see Introduction) (Dadario & Sughrue, 2023). From this perspective, our findings may reflect goal-directed access to phonological information, perhaps to facilitate appropriate speech production.

We also detected phonological information in left occipital areas, including superior lateral occipital cortex, the occipital pole, and ventral occipital cortex. The latter (posteriormost) areas are sensitive to low-level orthographic properties (see first paragraph of this section); it is likely that such information is relevant to spelling-to-sound conversion. The role of superior lateral occipital cortex is not clear, though it is worth noting that immediately adjacent parietal cortices have been implicated in spelling-to-sound mapping (Levy et al., 2009).

Semantic information was decodable in right dorsal premotor cortex. This area has been implicated in a range of higher-order cognitive functions that may support integration of semantic information during articulation and / or episodic encoding. In particular, the cluster identified in our analysis approximately corresponds to the boundary of the dorsal and rostral (anterior) segments of dorsal premotor cortex, identified in an anatomical parcellation of this area by Genon et al. (2017). These authors suggest that the rostral segment in particular may be involved in higher-order processes, citing its functional connections to DLPFC and inferior parietal lobules. The role of DLPFC in language processing is discussed above, while the inferior parietal lobules have been linked to multisensory integration, particularly in the context of episodic encoding (Mesulam, 1998; Pasalar et al., 2010; Seghier, 2013; Zeller et al., 2015). As such, we posit that the role of dorsal premotor cortex in our study may reflect integration of each word's semantic features during articulatory planning or monitoring, via connections to DLPFC; recognizing that accessing semantic features of a word can facilitate access to that word's phonological representations (e.g., Bailey et al., 2023). Another possibility is that this area supports integration of semantic features into multisensory episodic representations (consistent with Fawcett et al., 2022), via connections to inferior parietal cortex.

4.2. Decoding Information In Silent Reading

We found that, relative to aloud reading, silent reading was associated with greater decodability of orthographic information in the right anterior temporal lobe. Notably, evidence for orthographic decodability increased in the posterior-to-anterior direction along the (anterior portion of the) ventral visual pathway (VVP)[6]. This finding is consistent with Zhao et al. (2017), who identified increasing decodability of orthographic information along the posterior-to-anterior axis of the VVP bilaterally. We interpret this finding in our study to reflect a shift towards direct print-to-meaning (i.e., orthographic-to-semantic) mapping, as opposed to phonologically-mediated access to meaning, during silent reading. While at face value this may seem like a contradictory statement—why would such an effect not be revealed by the semantic model?—we reason that that extraction of meaning directly from the visual word form should entail sensitivity to orthographic features in the area(s) responsible for that process (elaborated below). Importantly, our interpretation does not imply deeper semantic processing in the silent condition (as we did not observe silent > aloud semantic decodability); rather, it concerns the *routes* by which meaning is extracted.

Print-to-meaning mapping has been linked to connectivity along the VVP, particularly between the fusiform gyrus and anterior temporal regions (Taylor et al., 2017). The anterior temporal lobes in turn have long been associated with semantic access; in particular, the hub-and-spoke model of semantics (Patterson et al., 2007) considers these areas as a semantic 'hub' where inputs from multiple modalities (including visual word forms) converge. While recent versions of this model consider print-to-meaning operations to be left-lateralized (Ralph et al., 2017; Rice et al., 2015), some empirical work has supported right-sided contributions (Pobric et al., 2010; Taylor et al., 2017).

Why should aloud and silent reading entail different degrees of print-to-meaning mapping? Connectionist models of reading consider access to meaning as the sum of parallel inputs from both direct and phonologically-mediated pathways (Harm & Seidenberg, 2004).

[6] The VVP runs from early visual cortex to anterior ventral temporal cortex, with subsequent projections to prefrontal areas (Kravitz et al., 2013).

From this perspective, the 'division of labour' between pathways depends on ease of mapping in each—for example, sound-meaning mapping is ambiguous in the case of homophones (e.g., *ewes-use*), and may be resolved by assigning greater weighting to the direct pathway (Harm & Seidenberg, 2004). We reason that such division of labour may also be affected by task-related availability of phonological information. That is, if spelling-to-sound computations are already being promoted to meet task demands, as appears to be the case when reading aloud, then phonologically mediated access to semantics may be assigned greater weighting simply because the relevant computations are already being executed. By extension, reduced demands for phonological information when reading silently may result in *relatively* greater weighting on direct print-to-meaning mapping.

4.3. General Discussion

Our study provides evidence that aloud and silent reading entail differential decodability of multiple types of stimulus information. Broadly, this finding is consistent with the view of embodied and grounded cognition, whereby different experiences (e.g., reading aloud versus silently) give rise to fundamental changes in how stimuli are perceived and evaluated (Matheson & Barsalou, 2018). More specifically, we interpret our results to reflect flexible cognitive states during reading, whereby certain types of information are weighted according to the speakers' goals. Such goal-dependent weighting, we argue, is primarily driven by the demands of reading aloud.

How is such weighting determined? Consider that some types of information are more relevant or useful for certain tasks; consider also that any reading task will recruit cognitive resources that must be allocated in a manner which best serves the reader's goals. We suggest that, when reading aloud, phonological and articulatory information are weighted more heavily because those features are useful for planning, execution, and monitoring of speech output. Similarly, visual information is necessary for recognising orthographic units; increased weighting here might therefore facilitate spelling-to-sound conversion. Semantic information may also

facilitate access to phonological representations; alternatively, increased weighting of semantic information may reflect an intrinsic component of the encoding process.

When speech production is not required, the aforementioned types of information are less relevant to one's goal (reading the word), and so receive less weighting. We have argued that decreased weighting on phonological information in particular gives rise to greater weighting on orthographic information for the purpose of semantic access. More precisely, de-weighting of spelling-to-sound computations may necessitate relatively greater emphasis on direct print-to-meaning mapping when reading silently. Overall, our findings provide an adaptive view of information processing, whereby each type of information is weighted according to its task relevance.

A subtly different interpretation, and one which aligns with attentional accounts of the production effect (Fawcett, 2013; Fawcett et al., 2023; Mama et al., 2018; Ozubko et al., 2012), is overall reduced cognitive investment when reading words silently. Silent words might be processed relatively superficially—that is, readers can perceive and extract meaning from visually presented words based on orthographic information alone, without needing to invest further cognitive resources in phonological and articulatory computations. Such superficial processing may apply to all stimulus properties, rather than specific types of information that are directly relevant to production. Indeed, this provides a satisfying explanation for (relatively) reduced decodability of visual information in the silent condition: this information is arguably no more relevant to aloud reading than it is to silent reading, therefore decreased decodability may reflect a global decrease in cognitive investment.

Our findings also have implications for research on the production effect (and, more generally, for research on how words are encoded in memory). The production effect is reliably elicited when memory for aloud versus silent reading is tested, including contexts in which participants are not informed in advance that they will be tested on the studied material (e.g., P. A. MacDonald & MacLeod, 1998). Thus while we did not explicitly assess participants' memory for the words in the present experiment (largely because our stimulus list was relatively short, with each word repeated 4 times), the procedures employed in our experiment (other than

stimulus repetition) accurately reflect the encoding conditions that reliably give rise to the production effect. Moreover, our findings align with theoretical accounts of the production effect. One account states that words read aloud are encoded as highly distinctive memory traces comprising sensorimotor information elicited during articulation (MacLeod et al., 2010). Our finding that reading aloud increased decodability of articulatory information appears to support this position. At the very least, we have shown that articulatory information is *available* for encoding at the neural level, which is a major assumption of distinctiveness accounts. Moreover, the presence of phonological information in parietal areas (widely considered hubs for multisensory integration) also supports the notion of distinctive, integrated memory representations. As described in the Introduction, we do not feel that increased semantic decodability is incompatible with a distinctiveness-based account. It may be that the canonical memory advantage for words spoken aloud is facilitated both by distinctive sensorimotor experiences *and* enhanced semantic processing (a position previously expressed by Fawcett et al., 2022).

One possibility we must consider is that the results we observed are epiphenomenal to the production effect. Future research should therefore aim to link decodability (as measured in this study) directly to subjects' behavioural memory performance; this would provide a more complete understanding of which type(s) of information contribute to encoding. Moreover, future research might consider incorporating formal models of cognition. For example, the MINERVA-2 model of memory has proven able to reproduce the production effect by simulating sensory distinctiveness (Jamieson et al., 2016). In a different vein, some models consider cognitive representations as hierarchical, interconnected structures encompassing all types of stimulus information, as opposed to information being represented in terms of discrete processes (Matheson & Barsalou, 2018; Meyer & Damasio, 2009). Applying such models in the context of RSA may provide a valuable means of testing formal cognitive theories of cognitive and neural representations.

Data and Code Availability

Code for all analyses reported in this manuscript is available on GitHub: https://github.com/lbailey25/Production_Effect_MVPA; additional materials that are necessary for analyses are stored in an Open Science Framework (OSF) repository: https://osf.io/czb26/?view_only=86a66caf1d71484d8ef0293cfa2371df. Twelve participants consented to their anonymized data being made publicly available. Data from those participants are available on the OSF repository.

Acknowledgements

We wish to thank the following individuals for their assistance with data collection: Matt Rogers, Laura McMillan, Cindy Hamon-Hill. We also wish to thank Philip Cook for his assistance with spatial normalization of MVPA output. This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) to GEB, JMF and AJN, and from the Nova Scotia Health Research Foundation to AJN. LMB was supported by a Killam Predoctoral scholarship.

CRediT authorship statement

Lyam M. Bailey: Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Data curation, Writing - original draft, Visualization, Project administration, Funding acquisition. **Heath E. Matheson:** Conceptualization, Methodology, Software, Writing - review & editing. **Jonathan M. Fawcett:** Conceptualization, Writing - review & editing. **Glen E. Bodner:** Conceptualization, Funding acquisition, **Aaron J. Newman:** Conceptualization, Methodology, Supervision, Project administration, Funding acquisition.

References

Alario, F.-X., Chainay, H., Lehericy, S., & Cohen, L. (2006). The role of the supplementary motor area (SMA) in word production. *Brain Research*, *1076*(1), 129–143. https://doi.org/10.1016/j.brainres.2005.11.104

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*(1), 26–41. https://doi.org/10.1016/j.media.2007.06.004

Bailey, L. M., Bodner, G. E., Matheson, H. E., Stewart, B. M., Roddick, K., O'Neil, K., Simmons, M., Lambert, A. M., Krigolson, O. E., Newman, A. J., & Fawcett, J. M. (2021). Neural correlates of the production effect: An fMRI study. *Brain and Cognition*, *152*, 105757. https://doi.org/10.1016/j.bandc.2021.105757

Bailey, L. M., Lockary, K., & Higby, E. (2023). Cross-linguistic influence in the bilingual lexicon: Evidence for ubiquitous facilitation and context-dependent interference effects on lexical processing. *Bilingualism: Language and Cognition*, 1–20. https://doi.org/10.1017/S1366728923000597

Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, *104*(21), 9087–9092. https://doi.org/10.1073/pnas.0703300104

Beijering, K., Gooskens, C., & Heeringa, W. (2008). Predicting intelligibility and perceived linguistic distance by means of the Levenshtein algorithm. *Linguistics in the Netherlands*,

25(1), 13–24.

Binder, J. R., & Desai, R. H. (2011). The Neurobiology of Semantic Memory. *Trends in Cognitive Sciences*, *15*(11), 527–536. https://doi.org/10.1016/j.tics.2011.10.001

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex*, *19*(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Bolger, D. J., Hornickel, J., Cone, N. E., Burman, D. D., & Booth, J. R. (2008). Neural correlates of orthographic and phonological consistency effects in children. *Human Brain Mapping*, *29*(12), 1416–1429. https://doi.org/10.1002/hbm.20476

Bourguignon, N. J. (2014). A rostro-caudal axis for language in the frontal lobe: The role of executive control in speech production. *Neuroscience & Biobehavioral Reviews*, *47*, 431–444. https://doi.org/10.1016/j.neubiorev.2014.09.008

Brem, S., Bach, S., Kucian, K., Kujala, J. V., Guttorm, T. K., Martin, E., Lyytinen, H., Brandeis, D., & Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter–speech sound correspondences. *Proceedings of the National Academy of Sciences*, *107*(17), 7939–7944. https://doi.org/10.1073/pnas.0904402107

Carota, F., Nili, H., Pulvermüller, F., & Kriegeskorte, N. (2021). Distinct fronto-temporal substrates of distributional and taxonomic similarity among words: Evidence from RSA of BOLD signals. *NeuroImage*, *224*, 117408. https://doi.org/10.1016/j.neuroimage.2020.117408

Carreiras, M., Armstrong, B. C., Perea, M., & Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences*, *18*(2), 90–98. https://doi.org/10.1016/j.tics.2013.11.005

Caucheteux, C., Gramfort, A., & King, J.-R. (2023). Evidence of a predictive coding hierarchy in the human brain listening to speech. *Nature Human Behaviour*, *7*(3), Article 3. https://doi.org/10.1038/s41562-022-01516-2

Chee, Q. W., Chow, K. J., Yap, M. J., & Goh, W. D. (2020). Consistency norms for 37,677 english words. *Behavior Research Methods*, *52*(6), 2535–2555. https://doi.org/10.3758/s13428-020-01391-7

Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, *125*(5), 1054–1069. https://doi.org/10.1093/brain/awf094

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A Dual Route Cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Cornelissen, P. L., Kringelbach, M. L., Ellis, A. W., Whitney, C., Holliday, I. E., & Hansen, P. C. (2009). Activation of the Left Inferior Frontal Gyrus in the First 200 ms of Reading: Evidence from Magnetoencephalography (MEG). *PLoS ONE*, *4*(4), e5359. https://doi.org/10.1371/journal.pone.0005359

Dadario, N. B., & Sughrue, M. E. (2023). The functional role of the precuneus. *Brain*, *146*(9), 3598–3607. https://doi.org/10.1093/brain/awad181

Devereux, B. J., Clarke, A., Marouchos, A., & Tyler, L. K. (2013). Representational Similarity Analysis Reveals Commonalities and Differences in the Semantic Processing of Words and Objects. *Journal of Neuroscience*, *33*(48), 18906–18916. https://doi.org/10.1523/JNEUROSCI.3809-13.2013

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*. https://www.frontiersin.org/articles/10.3389/fpsyg.2014.00781

Dienes, Z. (2016). How Bayes factors change scientific practice. *Journal of Mathematical Psychology*, *72*, 78–89.

Dietz, N. A. E., Jones, K. M., Gareau, L., Zeffiro, T. A., & Eden, G. F. (2005). Phonological decoding involves left posterior fusiform gyrus. *Human Brain Mapping*, *26*(2), 81–93. https://doi.org/10.1002/hbm.20122

Fawcett, J. M. (2013). The production effect benefits performance in between-subject designs: A meta-analysis. *Acta Psychologica*, *142*(1), 1–5. https://doi.org/10.1016/j.actpsy.2012.10.001

Fawcett, J. M., Baldwin, M. M., Whitridge, J. W., Swab, M., Malayang, K., Hiscock, B., Drakes, D. H., & Willoughby, H. V. (2023). Production improves recognition and reduces intrusions in between-subject designs: An updated meta-analysis. *Canadian Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale*, *77*(1), 35–44. https://doi.org/10.1037/cep0000302

Fawcett, J. M., Bodner, G. E., Paulewicz, B., Rose, J., & Wakeham-Lewis, R. (2022). Production can enhance semantic encoding: Evidence from forced-choice recognition with homophone versus synonym lures. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-022-02140-x

Fiez, J. A., Balota, D. A., Raichle, M. E., & Petersen, S. E. (1999). Effects of Lexicality, Frequency, and Spelling-to-Sound Consistency on the Functional Anatomy of Reading. *Neuron*, *24*(1), 205–218. https://doi.org/10.1016/S0896-6273(00)80833-8

Fiez, J. A., & Petersen, S. E. (1998). Neuroimaging studies of word reading. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(3), 914–921. https://doi.org/10.1073/pnas.95.3.914

Fine, J. M., & Hayden, B. Y. (2021). The whole prefrontal cortex is premotor cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1844), 20200524. https://doi.org/10.1098/rstb.2020.0524

Fischer-Baum, S., Bruggemann, D., Gallego, I. F., Li, D. S. P., & Tamez, E. R. (2017). Decoding levels of representation in reading: A representational similarity approach. *Cortex*, *90*, 88–102. https://doi.org/10.1016/j.cortex.2017.02.017

Fischer-Baum, S., Kook, J. H., Lee, Y., Ramos-Nuñez, A., & Vannucci, M. (2018). Individual Differences in the Neural and Cognitive Mechanisms of Single Word Reading. *Frontiers in Human Neuroscience*, *12*. https://www.frontiersin.org/articles/10.3389/fnhum.2018.00271

Fontan, L., Ferrané, I., Farinas, J., Pinquier, J., & Aumont, X. (2016). Using phonologically weighted Levenshtein distances for the prediction of microscopic intelligibility. *Annual Conference Interspeech (INTERSPEECH 2016)*, 650–654.

Freund, M. C., Bugg, J. M., & Braver, T. S. (2021). A Representational Similarity Analysis of Cognitive Control during Color-Word Stroop. *Journal of Neuroscience*, *41*(35), 7388–7402. https://doi.org/10.1523/JNEUROSCI.2956-20.2021

Genon, S., Li, H., Fan, L., Müller, V. I., Cieslik, E. C., Hoffstaedter, F., Reid, A. T., Langner, R., Grefkes, C., Fox, P. T., Moebus, S., Caspers, S., Amunts, K., Jiang, T., & Eickhoff, S. B. (2017). The Right Dorsal Premotor Mosaic: Organization, Functions, and Connectivity.

Cerebral Cortex (New York, NY), *27*(3), 2095–2110. https://doi.org/10.1093/cercor/bhw065

Gramfort, A., Pallier, C., Varoquaux, G., & Thirion, B. (2012). Decoding Visual Percepts Induced by Word Reading with fMRI. *2012 Second International Workshop on Pattern Recognition in NeuroImaging*, 13–16. https://doi.org/10.1109/PRNI.2012.20

Grootswagers, T., Robinson, A. K., & Carlson, T. A. (2019). The representational dynamics of visual objects in rapid serial visual processing streams. *NeuroImage*, *188*, 668–679. https://doi.org/10.1016/j.neuroimage.2018.12.046

Grootswagers, T., Robinson, A. K., Shatek, S. M., & Carlson, T. A. (2019). Untangling featural and conceptual object representations. *NeuroImage*, *202*, 116083. https://doi.org/10.1016/j.neuroimage.2019.116083

Guo, W., Geng, S., Cao, M., & Feng, J. (2022). Functional Gradient of the Fusiform Cortex for Chinese Character Recognition. *eNeuro*, *9*(3). https://doi.org/10.1523/ENEURO.0495-21.2022

Hall, K. C., Mackie, J. S., & Lo, R. Y.-H. (2019). Phonological CorpusTools: Software for doing phonological analysis on transcribed corpora. *International Journal of Corpus Linguistics*, *24*(4), 522–535. https://doi.org/10.1075/ijcl.18009.hal

Harm, M. W., & Seidenberg, M. S. (2004). Computing the Meanings of Words in Reading: Cooperative Division of Labor Between Visual and Phonological Processes. *Psychological Review*, *111*, 662–720. https://doi.org/10.1037/0033-295X.111.3.662

Hauk, O., Coutout, C., Holden, A., & Chen, Y. (2012). The time-course of single-word reading: Evidence from fast behavioral and brain responses. *NeuroImage*, *60*(2), 1462–1477.

https://doi.org/10.1016/j.neuroimage.2012.01.061

Henningsen-Schomers, M. R., & Pulvermüller, F. (2022). Modelling concrete and abstract concepts using brain-constrained deep neural networks. *Psychological Research*, *86*(8), 2533–2559. https://doi.org/10.1007/s00426-021-01591-6

Hertrich, I., Dietrich, S., & Ackermann, H. (2016). The role of the supplementary motor area for speech and language processing. *Neuroscience & Biobehavioral Reviews*, *68*, 602–610. https://doi.org/10.1016/j.neubiorev.2016.06.030

Hertrich, I., Dietrich, S., Blum, C., & Ackermann, H. (2021). The Role of the Dorsolateral Prefrontal Cortex for Speech and Language Processing. *Frontiers in Human Neuroscience*, *15*. https://www.frontiersin.org/articles/10.3389/fnhum.2021.645209

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), Article 5. https://doi.org/10.1038/nrn2113

Jamieson, R. K., Mewhort, D. J. K., & Hockley, W. E. (2016). A computational account of the production effect: Still playing twenty questions with nature. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *70*(2), 154.

Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: A metanalysis of 35 neuroimaging studies. *Neuroimage*, *20*(2), 693–712.

Kaiser, D., Moeskops, M. M., & Cichy, R. M. (2018). Typical retinotopic locations impact the time course of object coding. *NeuroImage*, *176*, 372–379. https://doi.org/10.1016/j.neuroimage.2018.05.006

Khalilian-Gourtani, A., Wang, R., Chen, X., Yu, L., Dugan, P., Friedman, D., Doyle, W., Devinsky, O., Wang, Y., & Flinker, A. (2022). A Corollary Discharge Circuit in Human Speech. *bioRxiv*,

2022.09.12.507590. https://doi.org/10.1101/2022.09.12.507590

Kim, C., Chung, C., & Kim, J. (2013). Task-dependent response conflict monitoring and cognitive control in anterior cingulate and dorsolateral prefrontal cortices. *Brain Research*, *1537*, 216–223. https://doi.org/10.1016/j.brainres.2013.08.055

Kim, H. (2011). Neural activity that predicts subsequent memory and forgetting: A meta-analysis of 74 fMRI studies. *NeuroImage*, *54*(3), 2446–2461. https://doi.org/10.1016/j.neuroimage.2010.09.045

Koechlin, E. (2011). Frontal pole function: What is specifically human? *Trends in Cognitive Sciences*, *15*(6), 241.

Krekelberg, B. (2022). *BayesFactor* (v2.2.0) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.7006300

Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: Integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, *17*(8), 401–412. https://doi.org/10.1016/j.tics.2013.06.007

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*. https://www.frontiersin.org/articles/10.3389/neuro.06.004.2008

Lee, C.-Y., Tsai, J.-L., Kuo, W.-J., Yeh, T.-C., Wu, Y.-T., Ho, L.-T., Hung, D. L., Tzeng, O. J., & Hsieh, J.-C. (2004). Neuronal correlates of consistency and frequency effects on Chinese character naming: An event-related fMRI study. *Neuroimage*, *23*(4), 1235–1245.

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge university press.

Lehman, V. T., Black, D. F., Bernstein, M. A., & Welker, K. M. (2016). Temporal lobe anatomy: Eight imaging signs to facilitate interpretation of MRI. *Surgical and Radiologic Anatomy*, *38*(4), 433–443. https://doi.org/10.1007/s00276-015-1582-9

Leinenger, M. (2014). Phonological coding during reading. *Psychological Bulletin*, *140*(6), 1534–1555. https://doi.org/10.1037/a0037830

Levy, J., Pernet, C., Treserras, S., Boulanouar, K., Aubry, F., Démonet, J.-F., & Celsis, P. (2009). Testing for the Dual-Route Cascade Reading Model in the Brain: An fMRI Effective Connectivity Account of an Efficient Reading Style. *PLOS ONE*, *4*(8), e6675. https://doi.org/10.1371/journal.pone.0006675

Li, A., Yang, R., Qu, J., Dong, J., Gu, L., & Mei, L. (2022). Neural representation of phonological information during Chinese character reading. *Human Brain Mapping*, *43*(13), 4013–4029. https://doi.org/10.1002/hbm.25900

Liu, J., Zhang, H., Yu, T., Ren, L., Ni, D., Yang, Q., Lu, B., Zhang, L., Axmacher, N., & Xue, G. (2021). Transformative neural representations support long-term episodic memory. *Science Advances*, *7*(41), eabg9715. https://doi.org/10.1126/sciadv.abg9715

Liuzzi, A. G., Aglinskas, A., & Fairhall, S. L. (2020). General and feature-based semantic representations in the semantic network. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-65906-0

Loh, K. K., Procyk, E., Neveu, R., Lamberton, F., Hopkins, W. D., Petrides, M., & Amiez, C. (2020). Cognitive control of orofacial motor and vocal responses in the ventrolateral and dorsomedial human frontal cortex. *Proceedings of the National Academy of Sciences*, *117*(9), 4994–5005. https://doi.org/10.1073/pnas.1916459117

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, *288*(5472), 1835–1838.

MacDonald, P. A., & MacLeod, C. M. (1998). The influence of attention at encoding on direct and indirect remembering. *Acta Psychologica*, *98*(2–3), 291–310. https://doi.org/10.1016/s0001-6918(97)00047-4

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 671–685. https://doi.org/10.1037/a0018785

Mama, Y., Fostick, L., & Icht, M. (2018). The impact of different background noises on the Production Effect. *Acta Psychologica*, *185*, 235–242. https://doi.org/10.1016/j.actpsy.2018.03.002

Matheson, H. E., & Barsalou, L. W. (2018). Embodiment and Grounding in Cognitive Neuroscience. In *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (pp. 1–27). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119170174.epcn310

Matheson, H. E., Kenett, Y. N., Gerver, C., & Beaty, R. E. (2023). Representing creative thought: A representational similarity analysis of creative idea generation and evaluation. *Neuropsychologia*, *187*, 108587. https://doi.org/10.1016/j.neuropsychologia.2023.108587

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*(7), 293–299.

https://doi.org/10.1016/S1364-6613(03)00134-7

Meersmans, K., Storms, G., De Deyne, S., Bruffaerts, R., Dupont, P., & Vandenberghe, R. (2022). Orienting to different dimensions of word meaning alters the representation of word meaning in early processing regions. *Cerebral Cortex*, *32*(15), 3302–3317. https://doi.org/10.1093/cercor/bhab416

Meyer, K., & Damasio, A. (2009). Convergence and divergence in a neural architecture for recognition and memory. *Trends in Neurosciences*, *32*(7), 376–382. https://doi.org/10.1016/j.tins.2009.04.002

Moerel, D., Grootswagers, T., Robinson, A. K., Shatek, S. M., Woolgar, A., Carlson, T. A., & Rich, A. N. (2022). The time-course of feature-based attention effects dissociated from temporal expectation and target-related processes. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-10687-x

Murphy, K. A., Jogia, J., & Talcott, J. B. (2019). On the neural basis of word reading: A meta-analysis of fMRI evidence using activation likelihood estimation. *Journal of Neurolinguistics*, *49*, 71–83. https://doi.org/10.1016/j.jneuroling.2018.08.005

Nastase, S. A., Connolly, A. C., Oosterhof, N. N., Halchenko, Y. O., Guntupalli, J. S., Visconti di Oleggio Castello, M., Gors, J., Gobbini, M. I., & Haxby, J. V. (2017). Attention Selectively Reshapes the Geometry of Distributed Semantic Representation. *Cerebral Cortex*, *27*(8), 4277–4291. https://doi.org/10.1093/cercor/bhx138

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: multi-modal multivariate

pattern analysis of neuroimaging data in Matlab/GNU Octave. *Frontiers in Neuroinformatics*, *10*, 27.

Ozubko, J. D., Gopie, N., & MacLeod, C. M. (2012). Production benefits both recollection and familiarity. *Memory & Cognition*, *40*(3), 326–338. https://doi.org/10.3758/s13421-011-0165-1

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*(12), Article 12. https://doi.org/10.1038/nrn2277

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the Connectionist Dual Process (CDP++) model. *Cognitive Psychology*, *61*(2), 106–151. https://doi.org/10.1016/j.cogpsych.2010.04.001

Petersen, S. E., Fox, P. T., Snyder, A. Z., & Raichle, M. E. (1990). Activation of Extrastriate and Frontal Cortical Areas by Visual Words and Word-Like Stimuli. *Science*, *249*(4972), 1041–1044. https://doi.org/10.1126/science.2396097

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: A Journal of Neurology*, *132*(Pt 7), 1907–1917. https://doi.org/10.1093/brain/awp141

Pleisch, G., Karipidis, I. I., Brem, A., Röthlisberger, M., Roth, A., Brandeis, D., Walitza, S., & Brem, S. (2019). Simultaneous EEG and fMRI reveals stronger sensitivity to orthographic strings

in the left occipito-temporal cortex of typical versus poor beginning readers. *Developmental Cognitive Neuroscience*, *40*, 100717. https://doi.org/10.1016/j.dcn.2019.100717

Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2010). Amodal semantic representations depend on both anterior temporal lobes: Evidence from repetitive transcranial magnetic stimulation. *Neuropsychologia*, *48*(5), 1336–1342. https://doi.org/10.1016/j.neuropsychologia.2009.12.036

Polk, T. A., & Farah, M. J. (2002). Functional MRI evidence for an abstract, not perceptual, word-form area. *Journal of Experimental Psychology: General*, *131*(1), 65–72. https://doi.org/10.1037/0096-3445.131.1.65

Proklova, D., Kaiser, D., & Peelen, M. V. (2019). MEG sensor patterns reflect perceptual but not categorical similarity of animate and inanimate objects. *NeuroImage*, *193*, 167–177. https://doi.org/10.1016/j.neuroimage.2019.03.028

Pulvermüller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(20), 7865–7870. https://doi.org/10.1073/pnas.0509989103

Qu, J., Pang, Y., Liu, X., Cao, Y., Huang, C., & Mei, L. (2022). Task modulates the orthographic and phonological representations in the bilateral ventral Occipitotemporal cortex. *Brain Imaging and Behavior*. https://doi.org/10.1007/s11682-022-00641-w

Ralph, M. A. L., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*(1), Article 1.

https://doi.org/10.1038/nrn.2016.150

Řehůřek, R., & Sojka, P. (2010). Software Framework for Topic Modelling with Large Corpora. *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, 45–50.

Rice, G. E., Lambon Ralph, M. A., & Hoffman, P. (2015). The Roles of Left Versus Right Anterior Temporal Lobes in Conceptual Knowledge: An ALE Meta-analysis of 97 Functional Neuroimaging Studies. *Cerebral Cortex*, *25*(11), 4374–4391. https://doi.org/10.1093/cercor/bhv024

Rodd, J. M., Vitello, S., Woollams, A. M., & Adank, P. (2015). Localising semantic and syntactic processing in spoken and written language comprehension: An Activation Likelihood Estimation meta-analysis. *Brain and Language*, *141*, 89–102. https://doi.org/10.1016/j.bandl.2014.11.012

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. https://doi.org/10.3758/PBR.16.2.225

Schepens, J., Dijkstra, T., & Grootjen, F. (2012). Distributions of cognates in Europe as based on Levenshtein distance*. *Bilingualism: Language and Cognition*, *15*(1), 157–166. https://doi.org/10.1017/S1366728910000623

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*(12), 2400–2406. https://doi.org/10.1093/brain/123.12.2400

Seidenberg, M. S. (2005). Connectionist models of word reading. *Current Directions in Psychological Science*, *14*(5), 238–242.

Seidenberg, M. S., Farry-Thorn, M., & Zevin, J. D. (2022). Models of word reading: What have we learned? In *The Science of Reading* (pp. 36–59). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119705116.ch2

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523–568. https://doi.org/10.1037/0033-295X.96.4.523

Stark, C. E. L., & Squire, L. R. (2001). When zero is not zero: The problem of ambiguous baseline conditions in fMRI. *Proceedings of the National Academy of Sciences*, *98*(22), 12760–12766. https://doi.org/10.1073/pnas.221462998

Tarkiainen, A., Helenius, P., Hansen, P. C., Cornelissen, P. L., & Salmelin, R. (1999). Dynamics of letter string perception in the human occipitotemporal cortex. *Brain*, *122*(11), 2119–2132. https://doi.org/10.1093/brain/122.11.2119

Taylor, J. S. H., Davis, M. H., & Rastle, K. (2017). Comparing and validating methods of reading instruction using behavioural and neural findings in an artificial orthography. *Journal of Experimental Psychology: General*, *146*(6), 826.

Teichmann, L., Moerel, D., Baker, C., & Grootswagers, T. (2021). *An empirically-driven guide on using Bayes Factors for M/EEG decoding* (p. 2021.06.23.449663). bioRxiv. https://doi.org/10.1101/2021.06.23.449663

Teichmann, L., Moerel, D., Baker, C., & Grootswagers, T. (2022). An empirically-driven guide on using Bayes Factors for M/EEG decoding. *Aperture Neuro*, *2*, 2021.06.23.449663. https://doi.org/10.1101/2021.06.23.449663

Thesen, T., McDonald, C. R., Carlson, C., Doyle, W., Cash, S., Sherfey, J., Felsovalyi, O., Girard, H.,

Barr, W., Devinsky, O., Kuzniecky, R., & Halgren, E. (2012). Sequential then interactive processing of letters and words in the left fusiform gyrus. *Nature Communications*, *3*(1), Article 1. https://doi.org/10.1038/ncomms2220

Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words* (pp. xii, 274). Bureau of Publications, Teachers Co.

Tong, J., Binder, J. R., Humphries, C., Mazurchuk, S., Conant, L. L., & Fernandino, L. (2022). A Distributed Network for Multimodal Experiential Representation of Concepts. *Journal of Neuroscience*, *42*(37), 7121–7130. https://doi.org/10.1523/JNEUROSCI.1243-21.2022

Tremblay, P., & Gracco, V. L. (2009). Contribution of the pre-SMA to the production of words and non-speech oral motor gestures, as revealed by repetitive transcranial magnetic stimulation (rTMS). *Brain Research*, *1268*, 112–124. https://doi.org/10.1016/j.brainres.2009.02.076

Tremblay, P., & Gracco, V. L. (2010). On the selection of words and oral motor responses: Evidence of a response-independent fronto-parietal network. *Cortex*, *46*(1), 15–28. https://doi.org/10.1016/j.cortex.2009.03.003

Tsujimoto, S., Genovesio, A., & Wise, S. P. (2011). Frontal pole cortex: Encoding ends at the end of the endbrain. *Trends in Cognitive Sciences*, *15*(4), 169–176. https://doi.org/10.1016/j.tics.2011.02.001

Tulkens, S., Sandra, D., & Daelemans, W. (2018). WordKit: A Python Package for Orthographic and Phonological Featurization. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*, 2695–2703.

Turkeltaub, P. E., Goldberg, E. M., Postman-Caucheteux, W. A., Palovcak, M., Quinn, C., Cantor,

C., & Coslett, H. B. (2014). Alexia due to ischemic stroke of the visual word form area. *Neurocase*, *20*(2), 230–235. https://doi.org/10.1080/13554794.2013.770873

Umesh, P. (2012). Image processing in python. *CSI Communications*, *23*(2).

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B., & Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage*, *30*(4), 1414–1432. https://doi.org/10.1016/j.neuroimage.2005.11.002

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical Coding of Letter Strings in the Ventral Stream: Dissecting the Inner Organization of the Visual Word-Form System. *Neuron*, *55*(1), 143–156. https://doi.org/10.1016/j.neuron.2007.05.031

Walther, A., Nili, H., Ejaz, N., Alink, A., Kriegeskorte, N., & Diedrichsen, J. (2016). Reliability of dissimilarity measures for multi-voxel pattern analysis. *NeuroImage*, *137*, 188–200. https://doi.org/10.1016/j.neuroimage.2015.12.012

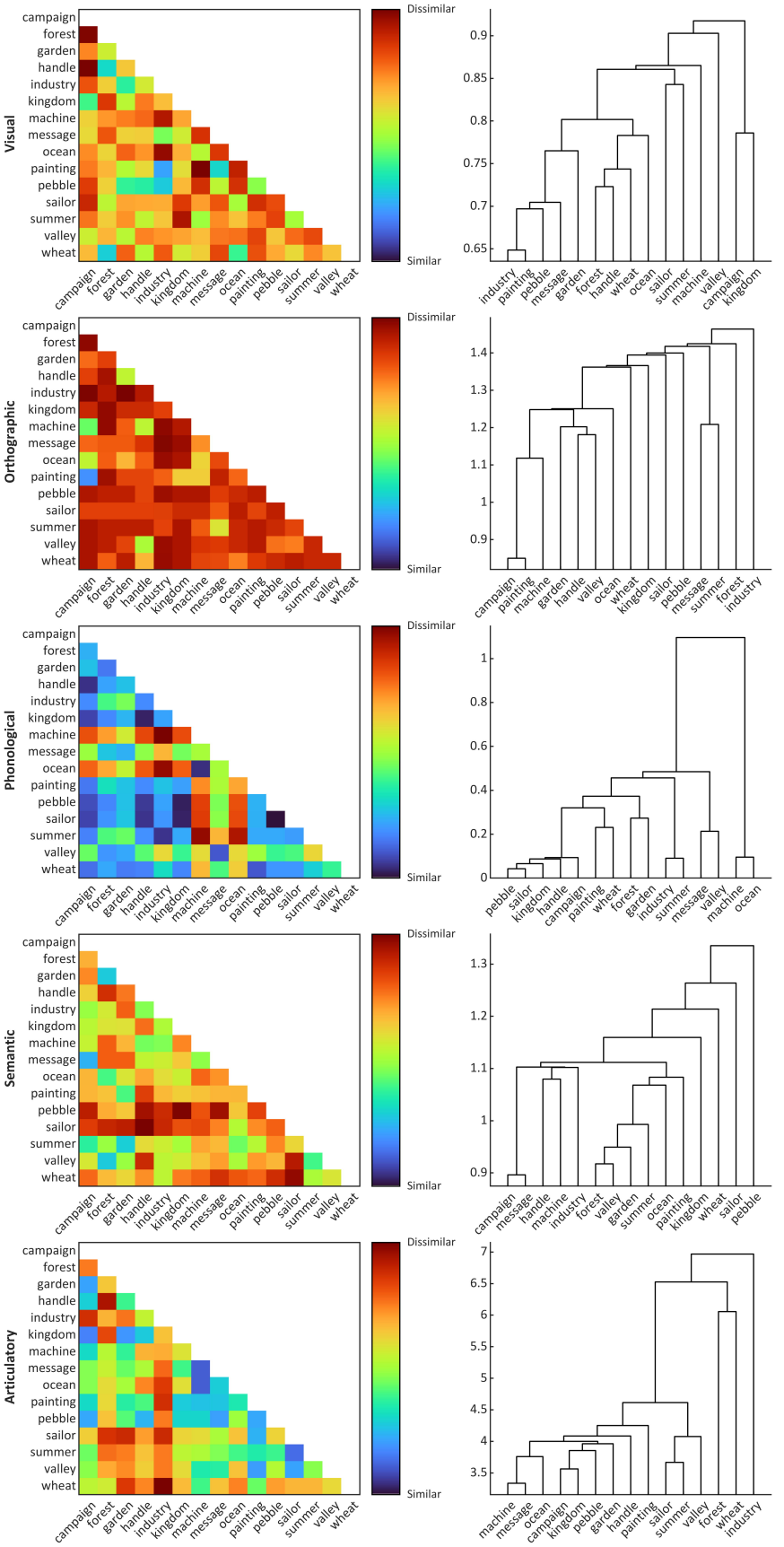
Wang, X., Xu, Y., Wang, Y., Zeng, Y., Zhang, J., Ling, Z., & Bi, Y. (2018). Representational similarity analysis reveals task-dependent semantic influence of the visual word form area. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-21062-0

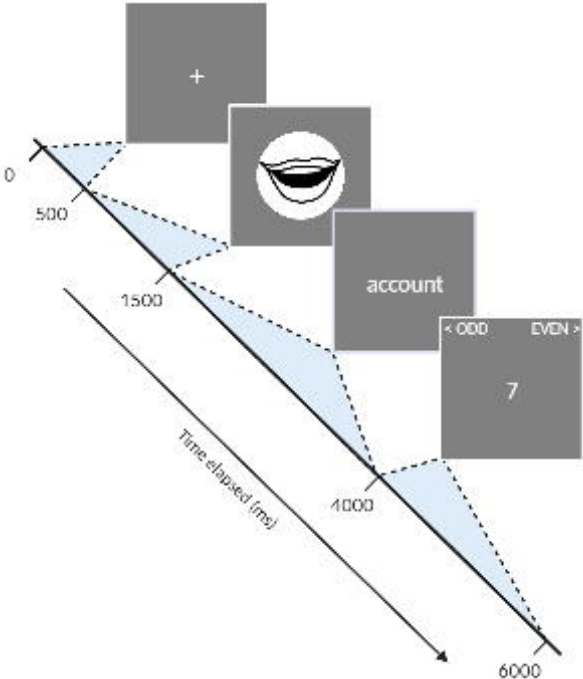
Warrington, E. K., & Shallice, T. (1980). Word-form dyslexia. *Brain: A Journal of Neurology*, *103*(1), 99–112. https://doi.org/10.1093/brain/103.1.99

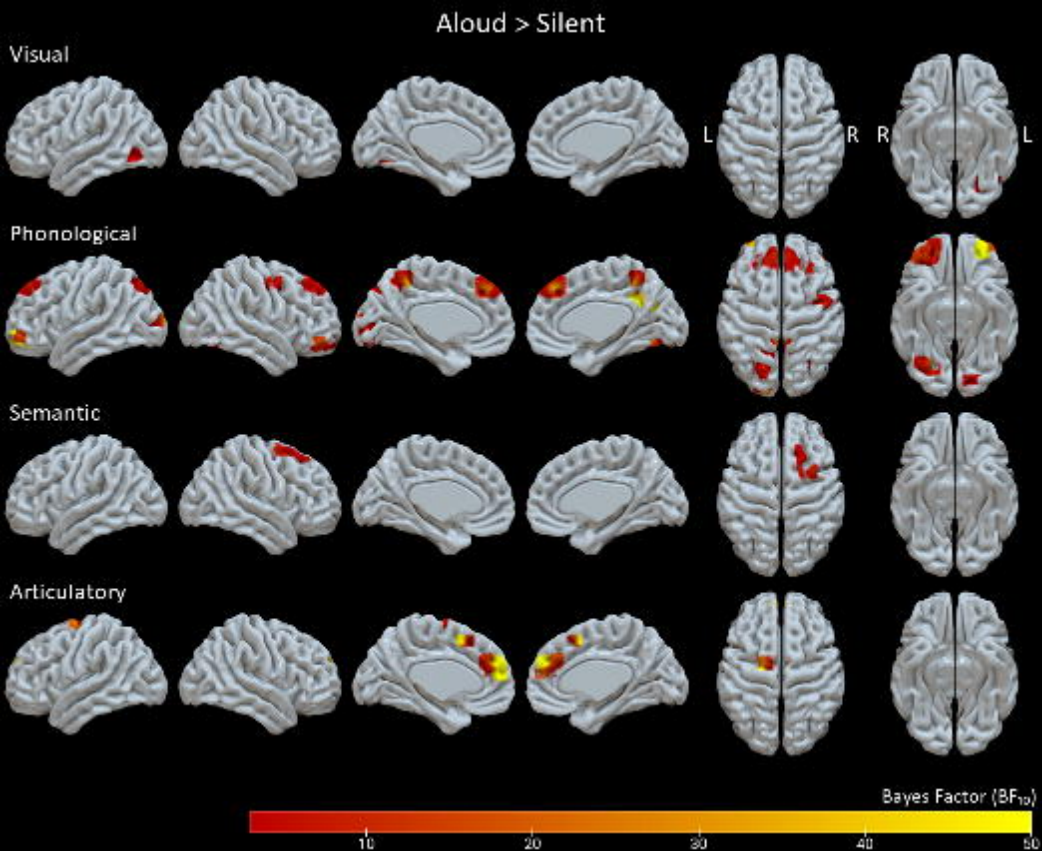
Zhang, W., Liu, Y., Wang, X., & Tian, X. (2020). The dynamic and task-dependent representational transformation between the motor and sensory systems during speech production. *Cognitive Neuroscience*, *11*(4), 194–204.

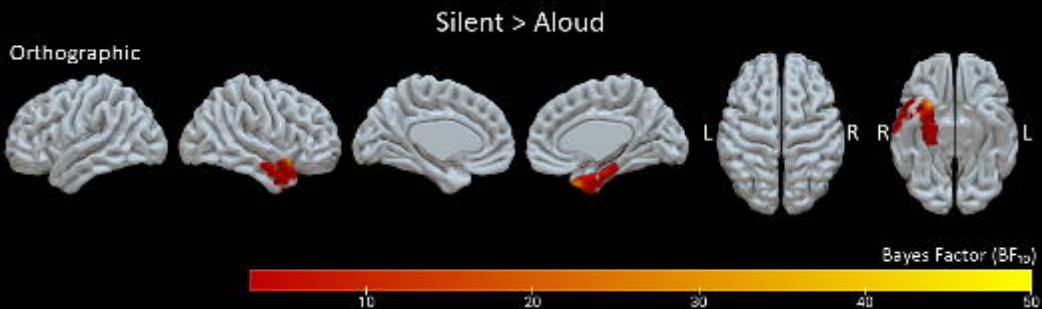
https://doi.org/10.1080/17588928.2020.1792868

Zhao, L., Chen, C., Shao, L., Wang, Y., Xiao, X., Chen, C., Yang, J., Zevin, J., & Xue, G. (2017). Orthographic and Phonological Representations in the Fusiform Cortex. *Cerebral Cortex*, *27*(11), 5197–5210. https://doi.org/10.1093/cercor/bhw300


Visual
Orthographic
Phonological
Semantic
Articulatory
campaign
forest
garden
handle
industry
kingdom
machine
message
ocean
painting
pebble
sailor
summer
valley
wheat
Dissimilar
Similar


+
account
< ODD
EVEN >
7
0
500
1500
4000
6000
Time elapsed (ms)


Aloud > Silent
Visual
Phonological
Semantic
Articulatory
L
R
R
L
Bayes Factor (BF_{10})
10
20
30
40
50


Silent > Aloud
Orthographic
L
R
R
L
Bayes Factor (BF_{10})
10
20
30
40
50