Distinct Neural Specializations for Learning to Read Words and Name Objects

J. S. H. Taylor[1,2], Kathleen Rastle[3], and Matthew H. Davis[1]

Abstract

■ Understanding the neural systems that underpin reading acquisition is key if neuroscientific findings are to inform educational practice. We provide a unique window into these systems by teaching 19 adults to read 24 novel words written in unfamiliar letters and to name 24 novel objects while in an MRI scanner. Behavioral performance on trained items was equivalent for the two stimulus types. However, componential letter-sound associations were extracted when learning to read, as shown by correct reading of untrained words, whereas object–name associations were holistic and arbitrary. Activity in bilateral anterior fusiform gyri was greater during object name learning than learning to read, and ROI analyses indicated that left mid-fusiform activity was predictive of success in object name learning but not in learning to read. In contrast, activity in bilateral parietal cortices was predictive of success for both stimulus types but was greater during learning and recall of written word pronunciations relative to object names. We argue that mid-to-anterior fusiform gyri preferentially process whole items and contribute to learning their spoken form associations, processes that are required for skilled reading. In contrast, parietal cortices preferentially process componential visual–verbal mappings, a process that is crucial for early reading development. ■

INTRODUCTION

Learning to read is arguably the most important skill children acquire in school. In their first few years of education, children learning to read alphabetic scripts must learn the letter-sound correspondences of their native writing system and how to break words down into their constituent letters and sound them out. This skill is what enables generalization—the ability to read unfamiliar words. Developmental psychology research has told us much about the underlying skills that help children to learn to read using letter-sound knowledge (e.g., Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001), and cognitive neuroscience research has revealed the brain regions that support skilled reading (Taylor, Rastle, & Davis, 2013). However, far less is known about the neural systems that are crucial for the earliest stages of reading development in which letter-sound correspondences are acquired.

The current study investigated this using an innovative "artificial orthography paradigm" in which adults learned to read new words written in unfamiliar symbols, while neural activity was measured with fMRI. This laboratory model of reading acquisition provides a unique window into the neural systems engaged when learning how to read words using letter-sound correspondences. This has never been achieved before; previous work has either studied adults or children processing familiar words relative to other stimuli such as objects (Ben-Shachar, Dougherty, Deutsch, & Wandell, 2011; Szwed et al., 2011; Kherif, Josse, & Price, 2010; Turkeltaub, Flowers, Lyon, & Eden, 2008), examined neural activity before and after learning (Mei et al., 2013; Yoncheva, Blau, Maurer, & McCandliss, 2010; Xue, Chen, Jin, & Dong, 2006; Frost et al., 2005; Hashimoto & Sakai, 2004), or, in rare cases, measured brain activity during paired-associate learning (Kumaran, Summerfield, Hassabis, & Maguire, 2009; Breitenstein et al., 2005), but not during learning to read.

Neural Systems for Skilled Reading in Left Occipitotemporal Cortex

As literacy is only a recent feature of human cognition, brain regions cannot have evolved specialization for reading. However, Dehaene and colleagues put forward the "neuronal recycling hypothesis," which suggests that left mid-fusiform gyrus neurons subserving object perception become specialized for recognizing written words through extended experience with text (Dehaene & Cohen, 2007, 2011). Dehaene and colleagues have shown that, under some conditions, this putative visual word form area (VWFA) responds more strongly to written words than to visually matched line drawings of objects (Szwed et al., 2011) and have proposed that the VWFA contains a posterior to anterior gradient representing

---

[1]Medical Research Council Cognition and Brain Sciences Unit, Cambridge, UK, [2]University of Cambridge, [3]Royal Holloway University of London

increasingly complex orthographic units, from letters, to frequently occurring bigrams, followed by quadrigrams (Vinckier et al., 2007; Dehaene, Cohen, Sigman, & Vinckier, 2005).

Whereas the work of Dehaene and colleagues suggests that the left mid-fusiform gyrus plays a role in analyzing orthographic forms in terms of their constituent parts, several authors argue that the VWFA represents whole words (Schurz et al., 2010; Glezer, Jiang, & Riesenhuber, 2009; Kronbichler et al., 2007) and is akin to an orthographic lexicon as proposed in some cognitive models (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). Converging evidence comes from patients with pure alexia who have lesions in left fusiform gyrus and typically have greater problems reading whole words than naming single letters (Tsapkini & Rapp, 2010; Cohen et al., 2003). Starrfelt, Habekost, and Leff (2009) also showed that such patients have an increased word length effect and suggested that they have difficulty with simultaneously processing and integrating multiple visual items.

Yet another contrasting view comes from Price and colleagues (Price & Devlin, 2003, 2011; Kherif et al., 2010), who argue that left ventral occipitotemporal cortex (vOT), including the VWFA, is not specialized for representing orthography. Instead, they suggest that this region sometimes responds more strongly to written words than objects (e.g., Szwed et al., 2011), because the former preferentially engage top–down influences from language processing regions. Some evidence for their proposal comes from a repetition suppression paradigm, in which activity in left vOT to a written word target (e.g., LION) was reduced when preceded by a picture of the same word (e.g., picture of a lion), relative to a picture of a different word (e.g., glove; Kherif et al., 2010). As the written word and picture shared semantic and phonological but not visual features, Kherif et al. argued that activity reduction in left vOT must have been driven by top–down influences from language processing regions. This view is supported by Mano et al.'s (2013) finding that task modulates activity in left vOT for written words versus visual objects. They showed that written words activate certain voxels in this region more than visual objects, but only during reading aloud and not in a brightness judgment task. Overall, this work suggests that left vOT is generally involved in visual object processing and that changes in its response profile may be driven by the formation of cross-modal associations between written and spoken language.

Two studies have examined how learning visual–verbal associations in an artificial writing system influences activity in fusiform gyri. Mei et al. (2013; using fMRI) and Yoncheva et al. (2010; using EEG) examined differences in fusiform activity following training on alphabetic items, which had one-to-one letter-sound mappings, versus logographic items, in which there was an arbitrary association between the individual letters in a word and its pronunciation. Mei et al. found that training induced increases in

fusiform activity were more left lateralized for alphabetic relative to logographic stimuli in posterior regions (± 40 -72 -18), but in anterior regions (± 40 -48 -18); changes were greatest in the left hemisphere for both stimulus types. Yoncheva et al. also reported that fusiform responses were more left lateralized for learning alphabetic relative to logographic stimuli. Overall, these studies provide some support for the idea that left posterior to mid-fusiform gyrus is sensitive to componentiality in spelling–sound relationships. However, these studies demonstrate the outcome of learning, rather than revealing the neural systems that contribute to learning, as is the goal of the current study. In a naturalistic longitudinal study, Ben-Shachar et al. (2011) tested children at multiple time points between the ages of 7 and 12 and found that left occipitotemporal sulcus (mean MNI coordinates across participants -49 -65 -9) became increasingly sensitive to the presence of written words in visual noise over this time period. This increasing cortical sensitivity predicted improvements in sight–word reading efficiency, but not pseudoword reading or phonological awareness skills. Unlike the studies using artificial writing systems, this supports the idea that left posterior to mid-fusiform regions play a role in the acquisition of whole-word, rather than letter-sound, reading skills.

Overall, cognitive neuroscience research demonstrates the importance of left occipitotemporal cortex, in particular posterior to mid-fusiform gyrus, for word reading. However, there is debate over the specificity of this region for representing written words versus visual objects and disagreement as to whether it is specialized for componential or whole-item processing. Furthermore, there is a paucity of research on the role of occipitotemporal cortices during reading acquisition. In the current experiment, we addressed these issues by comparing neural activity while adults learned to read words comprised of componential letter-sound mappings and learned holistic object–name associations.

Neural Systems for Skilled Reading in Left Parietal Cortex

Pugh et al. (2001) suggested that, whereas left vOT is involved in processing word identity, left temporoparietal cortex is involved in translating letters into sequences of sounds. This argument was supported by Carreiras et al. (2009), who observed structural changes in left temporoparietal cortex when ex-illiterates learned to read. However, although these studies specifically implicated left supramarginal gyrus (SMG) and posterior superior temporal gyrus (STG) in spelling–sound conversion, a recent meta-analysis of 36 neuroimaging studies of reading instead highlighted the importance of left inferior and superior parietal cortices in this process (Taylor et al., 2013). It was these more dorsal parietal regions, rather than the temporoparietal regions discussed by Pugh et al., which showed greater activity for pseudowords than

words, a key contrast that highlights letter-sound translation processes.

Further evidence that left inferior parietal cortex plays a role in componential cross-modal processing comes from Booth et al. (2003), who obtained performance correlated activity in this region when adults made judgments about whether spoken words were spelled similarly or written words sounded similar (cross-modal tasks), relative to making judgments about whether spoken words sounded similar or written words were spelled similarly (intramodal tasks). In children, this same region was active when making spelling similarity judgments about conflicting (jazz–has, pint–mint) relative to nonconflicting (dime–lime) spoken word pairs (Booth, Cho, Burman, & Bitan, 2007). Neuropsychological data also support the idea that left inferior parietal cortex is involved in letter-sound reading. Wilson et al. (2009) found that patients with surface dyslexia, who pronounce irregular words incorrectly (e.g., reading PINT so that it rhymes with MINT) because they have degraded lexical/semantic knowledge and rely on letter-sound rules, activated left inferior parietal cortex more than control participants when reading words.

Studies implicating left inferior and superior parietal cortex in letter-sound reading suggest that this region should play an important role in the early stages of reading acquisition when this skill is crucial. In support of this conjecture, Hashimoto and Sakai (2004) obtained activity in left superior parietal cortex when adults learned novel symbol–syllable associations versus processing familiar letter strings or symbol–tone associations. However, other studies implicate parietal cortex in whole-word learning. Lee et al. (2007) showed that gray matter density in bilateral inferior parietal cortices correlated with vocabulary size, and two studies have found that activation in left inferior parietal cortex changes as participants learn object names (Breitenstein et al., 2005; Cornelissen et al., 2004).

To summarize, there is evidence to suggest that left inferior parietal cortex may be important for employing and acquiring alphabetic letter-sound reading processes. However, as with left occipitotemporal cortex, it is somewhat unclear whether this region is particularly engaged when visual–spoken form mappings require componential processing or whether it is more generally involved in word learning. We distinguished between these possibilities by comparing neural activity during acquisition of componential letter-sound and holistic object–name associations.

The Current Study

Our goal was to shed new light on the roles that occipitotemporal and parietal cortices play in the earliest stages of reading acquisition. To this end, we considered the computational demands of learning to read words written in a novel alphabetic script versus learning novel object names. Both of these tasks involve paired-associate learning of novel visual–verbal mappings. However, the relationship between spelling and pronunciation in alphabetic languages is componential (breaks down into letters and sounds) and largely systematic (letters usually sound the same in different words), whereas the relationship between an object's visual form and its name is holistic and largely arbitrary, that is, similar sounding words (e.g., CAT, CAP) do not correspond to similar looking objects. By examining brain activity online, during acquisition, we were able to determine whether these distinct computational demands engage different neural systems.

Adults learned to read novel words written in an unfamiliar alphabet with systematic one-to-one letter-sound mappings and learned names for unfamiliar objects, which is an arbitrary and holistic association. We measured neural activity with fMRI throughout learning. It was imperative that we measured naming success online; first, to ensure that the two types of stimuli were equally learnable and, second, to enable us to relate performance to neural activity. Therefore, we interleaved training blocks, in which the visual–verbal pairings for words or objects were presented, with testing blocks, in which participants read words or named objects aloud. Analyses of performance during these test blocks showed that participants learned to read the words and name the objects with equivalent success but also confirmed that the two tasks did indeed have different computational demands because of their componential versus holistic visual–verbal associations.

Before we could use our fMRI data to address key questions about the differences between learning to read words and name objects, it was first necessary to validate our novel paradigm. We examined neural activity uniquely associated with learning and/or retrieving visual–verbal associations by comparing these trial types with trials in which participants attended to, but did not have to learn or retrieve, the novel stimuli. Multiple analyses using these comparisons converged on the finding that the neural systems that are activated during object naming and word reading in natural languages (Taylor et al., 2013; Price, Devlin, Moore, Morton, & Laird, 2005) were engaged when the participants in our study learned and retrieved novel visual–verbal associations. We next explored how the unique computational demands of learning to read words and name objects impacted on activity in this network of brain regions. We compared neural activity for word reading and object naming during both learning and retrieval trials, again relative to trials in which participants were exposed to but did not have to learn or retrieve, the novel stimuli. These analyses provided multiple lines of evidence to suggest unique contributions of occipitotemporal and parietal cortices to learning to read words and name objects.

METHODS

Participants

Twenty-two right-handed native English-speaking healthy adults aged 18–40 took part in the experiment. One

participant was excluded because he failed to learn how to read any of the words or name any of the objects, another because of excess head movement (>20 mm), and a third because of stimulus delivery problems. The remaining 19 participants (13 women) contributed to all analyses.

Materials

Three sets of 24 monosyllabic consonant–vowel–consonant pseudowords were constructed from 12 consonant (b, d, f, g, k, m, n, p, s, t, v, z) and four vowel (æ, ɛ, ɒ, ∧) phonemes. Within each set of pseudowords, consonants occurred twice in onset and twice in coda position, whereas vowels occurred six times each. Pseudowords were recorded by a female native English speaker and digitized at a sampling rate of 44.1 KHz. Each set of pseudowords was assigned to trained object, trained word, or untrained word categories across participants in a counterbalanced fashion. Twenty-four pictures of novel objects were randomly assigned a name from each of the three sets of 24 pseudowords. Sixteen unfamiliar alphabetic symbols were mapped to the 16 phonemes in a one-to-one manner and were used to construct the visual forms of trained and untrained

words. Figure 1 gives some examples of the experimental stimuli.

Procedure

Pre-exposure (Prior to Scanning)

Participants were first familiarized with the format of the training and testing trials using six pseudowords written in the Latin alphabet. The trial structure was otherwise identical to that used in the learning phase, which is described in the next section. Second, they were exposed to the visual forms of the stimuli described in the Materials section. They made same–different judgments about consecutively presented item pairs presented in a randomized order; two same and two different judgments were made about each item. Memory for visual forms was then tested in an old–new decision task in which all the experimental stimuli were presented along with 24 new unfamiliar objects and 24 new words constructed from 16 different unfamiliar symbols. Objects and words were presented in separate alternating blocks of six trials. Accuracy on this task was above chance for both words (mean d' = 3.18, $t(18)$ = 15.34, p < .001) and objects
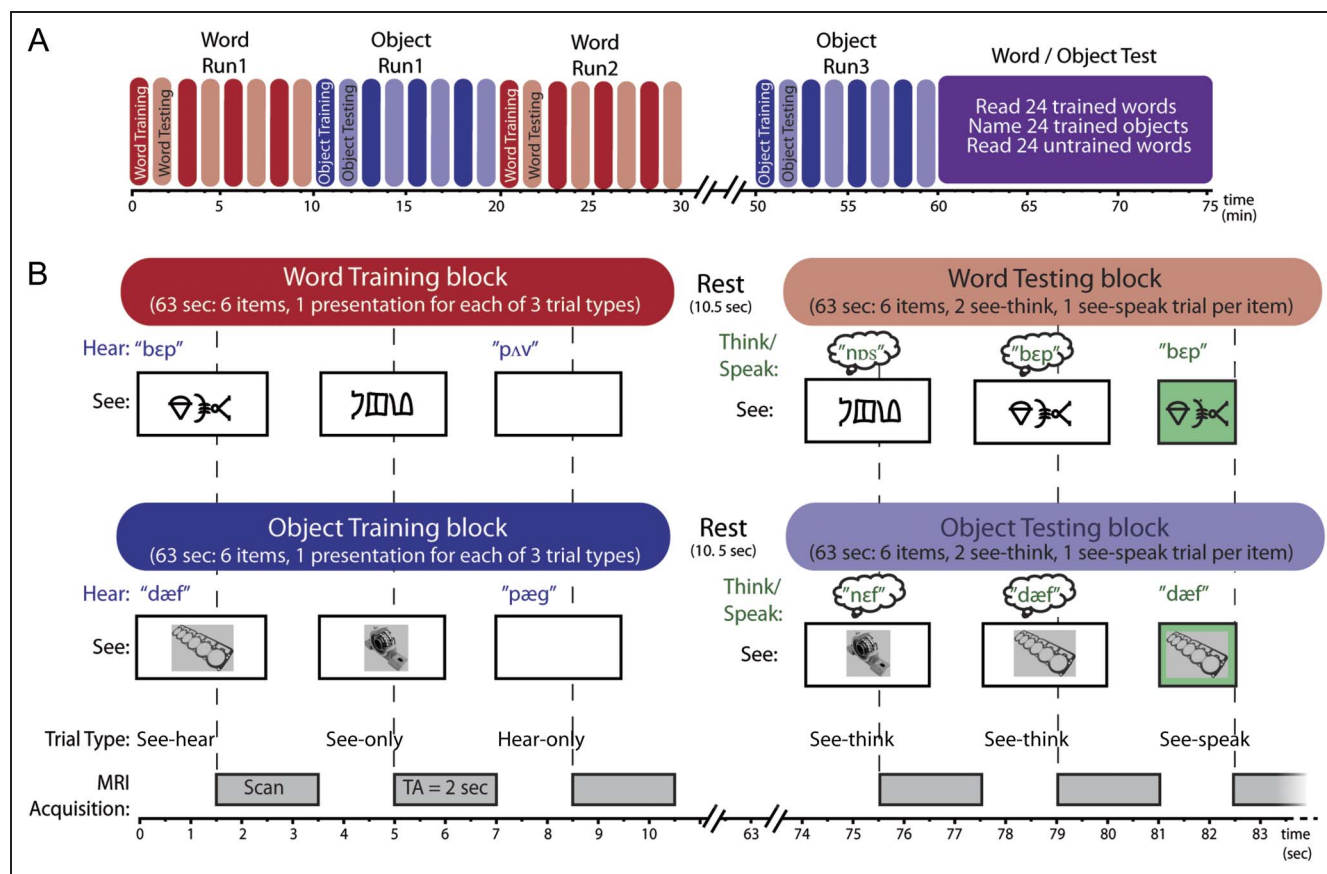

Figure 1. Structure and timings of fMRI session. (A) Experiment structure. (B) Trial format and timing during training and test blocks. Dotted lines indicate correspondence between stimulus presentation and scan onset. The order of items and trial types was randomized for each subject and each learning run. The trial format in the final test run was the same as trial format in the test blocks from the learning runs with fully randomized presentation of trained words, trained objects, and untrained words.

(mean d' = 3.19, $t(18)$ = 18.34, p < .001), and there was no difference between the two tasks, $t(18)$ < 1, *ns*, indicating that participants were equally familiar with the visual forms of the to-be-learned objects and symbols. Finally, preexposure to spoken forms involved listening to and repeating each of the 48 items four times.

Learning Phase (Figure 1)

Three object and three word runs were completed in the MRI scanner in an alternating order; half the participants started with a word run and half with an object run. All 24 words or objects were presented in a randomized order in each run. Runs were broken down into four training blocks (learn six items), each followed by a test block (retrieve pronunciations for these six items). Training blocks comprised 18 trials presented in a randomized order: six had concurrent presentation of an item's visual and spoken form (see-hear), six had isolated visual form presentation (see-only), and six had isolated spoken form presentation (hear-only). Contrasts between these different trial types enabled us to examine how activity differed when a trial afforded a learning opportunity (cross-modal presentation) relative to when it did not (unimodal presentation). Each training trial was 3500 msec in duration, with visual items presented for the first 2500 msec and spoken forms at the onset of the trial. Scan volume acquisition (1940 msec) commenced at 1560 msec. In testing blocks, participants retrieved the object names or written word pronunciations learned in the preceding training block. Testing blocks comprised 12 see-think trials, presented in a randomized order, in which participants were presented with an item's visual form and covertly retrieved its spoken form. Half of the see-think trials were immediately followed by a see-speak trial, in which the same item was presented and participants overtly articulated its pronunciation having retrieved it in the preceding see-think trial. Separating the retrieval (see-think) and articulation (see-speak) of pronunciations enabled us to dissociate activity related to each component and also allowed participants enough time to achieve both tasks. Each testing trial was 3500 msec in duration, with visual forms presented at the beginning of the trial and scanning acquisition (1940 msec) commencing at 1560 msec. Visual forms were presented for 2500 msec on see-think trials and 1500 msec on see-speak trials to encourage participants to generate spoken forms before the onset of scan volume acquisition. A 10.5-sec rest period followed each block.

Final Test Run

This was also conducted in the scanner (although fMRI data from this run do not contribute to the reported analyses) and involved participants reading aloud all the words they had learned, naming all the objects they had learned, and reading the 24 untrained words described

in the Materials section. Item presentation was fully randomized, and participants were not informed that they would see new items. The 72 items were presented in 12 blocks, each containing 12 see-think and 6 see-speak trials (as in the testing blocks from the learning runs). A 10.5-sec rest period followed each block.

Imaging Acquisition and Analysis

fMRI data were acquired on a 3T Siemens Trio scanner (Siemens Medical Systems, Erlangen, Germany) with a 12-channel head coil. BOLD fMRI images were acquired with fat saturation, 3 mm isotropic voxels and an interslice gap of .75 mm, flip angle of 78°, echo time = 30 msec, and a 64 × 64 data matrix. We used a sparse imaging design with a repetition time (3500 msec) longer than the acquisition time (1940 msec), which provided a 1560-msec period in which to present spoken words and record spoken responses in the absence of echoplanar scanner noise (Edmister, Talavage, Ledden, & Weisskoff, 1999; Hall et al., 1999). Stimuli were presented over high-quality electrostatic headphones built into ear defenders (Nordic-Neurolab, Bergen, Norway), and responses were recorded using a dual-channel MRI microphone (FOMRI II, Optoacoustics). One hundred seventy-four images were acquired in each of the six 10-min training–testing runs, and 258 images were acquired in the 15-min final test run. The acquisition was transverse oblique, angled to avoid the eyes and to achieve whole-brain coverage including the cerebellum. In a few cases, the very top of the parietal lobe was not covered. To assist in anatomical normalization, we also acquired a T_1-weighted structural volume using a magnetization prepared rapid acquisition gradient-echo protocol (repetition time = 2250 msec, echo time = 2.99 msec, flip angle = 9°, 1 mm slice thickness, 256 × 240 × 192 matrix, resolution = 1 mm isotropic).

Image processing and statistical analyses were performed using SPM8 software (Wellcome Trust Centre for Functional Neuroimaging, London, UK). The first six volumes of each scanning run were discarded to allow for equilibration effects. Images for each participant were realigned to the first image in the series (Friston et al., 1995) and coregistered to the structural image (Ashburner & Friston, 1997). The transformation required to bring a participant's structural T1 image into standard MNI space was calculated using tissue probability maps (Ashburner & Friston, 2005), and these warping parameters were then applied to all functional images for that participant. Normalized functional images were resampled to 2 mm isotropic voxels. The data were spatially smoothed with 8-mm FWHM isotropic Gaussian kernel prior to model estimation.

Data from each participant were entered into two general linear models for event-related analysis (Josephs & Henson, 1999). In both models, events were convolved with the SPM8 canonical hemodynamic response function. Movement parameters estimated at the realignment

stage of preprocessing were added as regressors of no interest. Low-frequency drifts were removed with a high-pass filter (128 sec) and AR1 correction for serial autocorrelation was made. Model 1 modeled all event types in each word and object run: see-only, hear-only, see-hear, see-think, and see-speak. We used this model to identify neural activity associated with learning (training blocks) and retrieving (testing blocks) visual–verbal associations.

From Model 1, two contrasts of interest were used to assess activity associated with learning and a third analysis examined activity associated with retrieval, all collapsed across stimulus type (words and objects). (1) To examine activity related to learning the associations between visual and verbal forms we contrasted see-hear trial activity with whichever of the unimodal trial types (see-only, hear-only) had the maximum activity. This [cross-modal – max(unimodal)] contrast was implemented using the imcalc function in SPM8; at each voxel, the mean activity for the unimodal trial with the greatest activity (averaged across run and stimulus type) was subtracted from the mean activity for cross-modal trials (averaged across run and stimulus type). This contrast is equivalent to requiring cross-modal activity to be greater than activity during both types of unimodal trial, that is, the intersection of [see-hear – see-only] ∩ [see-hear – hear-only]. (2) To examine how cross-modal learning activity changed over the course of training, we conducted a one-way ANOVA to compare the three training runs (Run 1 vs. Run 2 vs. Run 3) using the contrast [cross-modal – max(unimodal)] as the dependent measure for each run. (3) We assessed activity associated with retrieval of verbal from visual forms using the contrast [see-think – see-speak]. The same visual item was presented on each of these two trial types, enabling us to subtract activity related purely to perceptual processing. However, whereas see-think trials required participants to covertly generate an item's pronunciation from its visual form, see-speak trials always immediately followed see-think trials and thus only required participants to overtly articulate an item's pronunciation that they had retrieved on the previous trial.

Following these analyses collapsed across stimulus type, Model 1 was also used to conduct two sets of analyses that compared word and object activation during training and testing. Training analyses used both simple contrasts [word see-hear – object see-hear] and the reverse, and interaction contrasts [word [cross-modal – max(unimodal)] – object [cross-modal – max(unimodal)]] and the reverse. The same was true of testing analyses, in which the simple contrasts were [word see-think – object see-think] and the reverse, and the interaction contrasts were [word [see-think – see-speak] – object [see-think – see-speak]] and the reverse. Interaction contrasts had the advantage of subtracting out neural responses driven by visual differences between the stimulus types. However, as training progressed, participants may have started to covertly recall item pronunciations on see-only trials. Interaction contrasts

therefore also had the potential to conceal differences in neural activity that were in fact related to learning. It was thus important to additionally examine simple contrasts, particularly for the training analyses.

Model 2 specifically examined how neural activity during training differed as a function of learning success. To derive a success measure, we took an item's accuracy in the test block of the current run (1 = *correct*, 0 = *incorrect*) and subtracted its accuracy in the previous run. Note that for Run 1 we assumed that all items were incorrect on the "previous run." This gave us values of 1 for learned items, which were correct this run and incorrect on the previous run (Words: Run 1 = 34%, Run 2 = 39%, Run 3 = 15%; Objects: Run 1 = 46%, Run 2 = 29%, Run 3 = 16%), values of 0 for not-learned items, which were either correct or incorrect on both the current and previous run (Words: Run 1 = 66%, Run 2 = 56%, Run 3 = 78%; Objects: Run 1 = 54%, Run 2 = 66%, Run 3 = 76%), and values of −1 for forgotten items, which were incorrect on this run and correct on the previous run (Words: Run 1 = 0%, Run 2 = 5%, Run 3 = 7%; Objects: Run 1 = 0%, Run 2 = 5%, Run 3 = 8%). Model 2 thus contained 15 event types in each word and object run: see-only, hear-only, see-hear, see-think, and see-speak; each factorially crossed with the three levels of success for the corresponding item: learned, not-learned, forgotten. Using this model, we conducted the following contrasts on activity during see-hear trials: [learned – not-learned] collapsed across stimulus type, [objects [learned – not-learned] – words [learned – not-learned]], and [words [learned – not-learned] – [objects [learned – not-learned]]. We hypothesized that regions that were functionally involved in learning would show greater activity for learned than not-learned items. Furthermore, we reasoned that a more pronounced difference in activity for learned relative to not-learned items, for one stimulus type versus the other, would provide evidence that a region differentially contributed to learning to read versus learning object names. As we did not have prior hypotheses about brain activity associated with decreases in accuracy, trials in which an item was forgotten were modeled but did not contribute to any contrasts.

Contrasts of parameter estimates were taken forward to second-level group analyses (one-sample and paired-sample *t* tests, one-way ANOVA) using Participants as a random effect. All comparisons were assessed using a voxelwise uncorrected threshold of $p < .001$. After thresholding, only activations exceeding a cluster extent family-wise error (FWE)-corrected threshold of $p < .05$, obtained using the nonstationarity toolbox in SPM8 (Hayasaka, Phan, Liberzon, Worsley, & Nichols, 2004), were further considered for interpretation. Figures show results at this cluster extent-corrected threshold, displayed on a canonical brain image. Plots show mean parameter estimates (for the canonical hemodynamic response function) at specific voxels with zero reflecting activity following unmodeled null events (rest blocks). Cluster coordinates are reported in the space of the MNI 152 average brain template, and

anatomical labels were generated by MRICron (Rorden, Karnath, & Bonilha, 2007), which uses the automated anatomical labeling (AAL) template (Tzourio-Mazoyer et al., 2002).

RESULTS

Behavioral Data (Figure 2)

Adults learned to read the words and name the objects with equal success. Accuracy increased on each training run, $F(2, 36) = 80.92$, $\eta^2 = .82$, $p < .001$, and did not differ between the two stimulus types (main effect, $F(1, 18) < 1$, stimulus type × run interaction, $F(2, 36) = 2.16$, $p > .1$; Figure 2A). Figure 2B shows that final test run accuracy was equally good for trained words and objects, $t(18) < 1$, and that participants could also read untrained words, although somewhat less accurately than trained words, $t(18) = 3.96$, $\eta^2 = .68$, $p = .001$. It is worth noting that 70–80% items correct constitutes good performance on this difficult retrieval and production task. Participants' ability to read untrained words confirms that they extracted the sounds of the individual symbols from which the trained items were systematically constructed.

Figure 2C illustrates that, in the final test run, object naming accuracy was higher the more times an item had been named correctly during training. In contrast, final test accuracy for word reading did not differ for items named once, twice, or thrice correctly during training (ANOVA on final test accuracy: interaction between number of times correct in training [0 vs. 1 vs. 2 vs. 3] and stimulus type [words vs. objects], $F(3, 54) = 6.16$, $p = .001$, $\eta^2 = .26$). Thus, although overall performance was matched for the two types of stimuli, the holistic and arbitrary nature of the object–name associations meant that each item had to be learned individually and did not support learning of others, whereas the componen-

tial and systematic nature of the symbol–sound mappings that comprised the words meant that learning to read one word supported learning of others containing the same symbols. These behavioral differences further confirm that learning to read the words did indeed involve acquiring componential systematic symbol-to-sound mappings, whereas the arbitrary object–name associations were learned holistically. We can therefore consider how these differential learning demands were instantiated in our neuroimaging data.

Brain Regions Supporting Learning and Recall of Visual–Verbal Associations

Training Blocks

To determine which brain regions were activated during visual–verbal association learning, we contrasted cross-modal (see-hear) trial activity with the maximum activity observed during unimodal (see-only, hear-only) trials, averaged across all runs and both stimulus types (objects, words). The justification for using this contrast was that unimodal trials afforded little opportunity for forming visual–verbal associations, whereas cross-modal trials afforded maximal opportunity for forming such links. We subtracted the maximum unimodal activity because some regions may be deactive (relative to rest) during either see-only or hear-only trials, in which case, the mean or summed unimodal activity would not accurately reflect that region's activity profile (Beauchamp, 2005). Details on how we computed this contrast were provided in the Methods section. Cross-modal activity was greater than unimodal activity in bilateral occipitotemporal cortices (extending from middle occipital cortices to anterior fusiform gyri), bilateral inferior and superior parietal cortices, and left hippocampus (Figure 3, yellow overlay; Appendix 1).
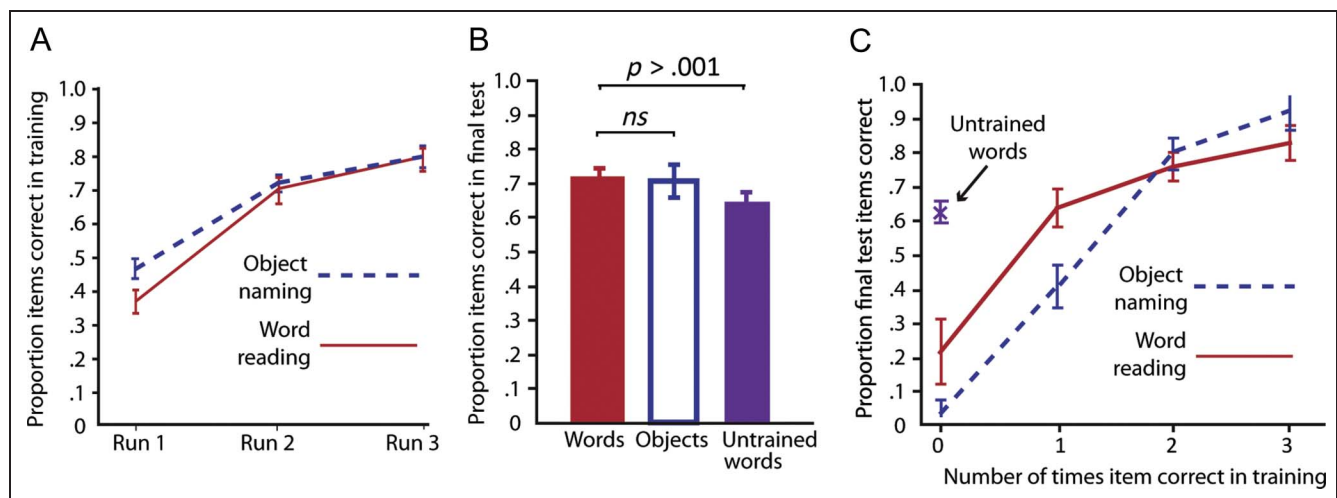

Figure 2. Behavioral performance. (A) Word reading and object naming accuracy during training. (B) Word reading (trained and untrained items) and object naming accuracy in the final test run. (C) Relationship between training and test performance. All error bars in this and subsequent figures use standard error appropriate for within-participant designs (Loftus & Masson, 1994).

Figure 3. Brain regions showing greater cross-modal than unimodal activity, averaged across stimulus type. In this and all subsequent figures, left and right hemisphere slices show whole-brain activations at $p <$ .001 voxelwise uncorrected and $p <$ .05 FWE cluster-corrected for 19 participants. Yellow = [see-hear – maximum (see-only, hear-only)], averaged across run and stimulus type; green = regions showing a main effect of run (1 vs. 2 vs. 3) in an ANOVA using the contrast [see-hear – maximum (see-only, hear-only)] as the dependent measure, averaged across stimulus type; orange = intersection of these two maps. Plots show activity (mean BOLD parameter estimate, arbitrary units) for each trial type in each run at peak voxels from the whole-brain contrast [see-hear – maximum (see-only, hear-only)] (yellow boxes: mid-fusiform gyrus, superior parietal cortex, hippocampus) or from the whole-brain main effect of run on this contrast (green box: precentral gyrus).
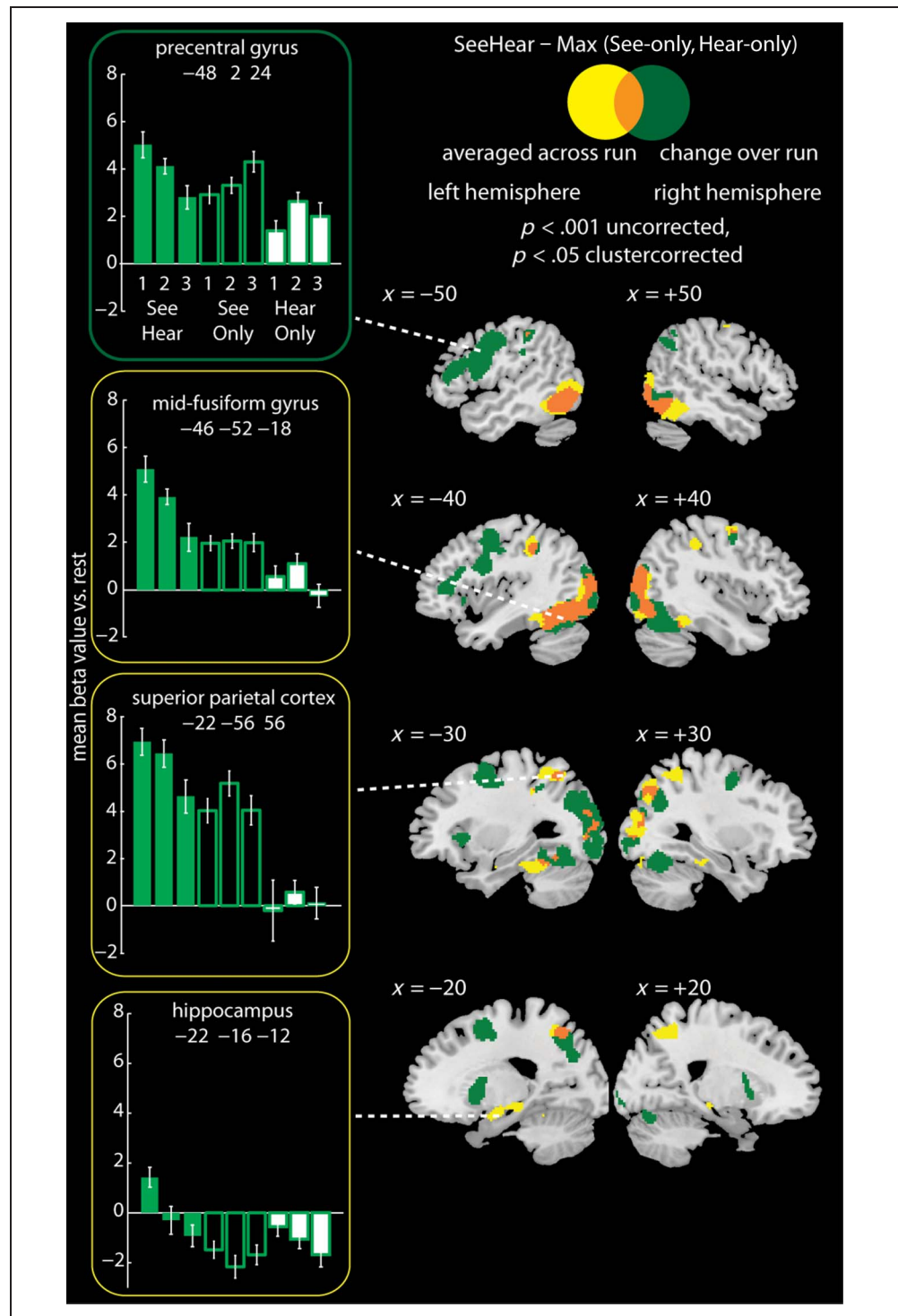
If activity during cross-modal trials was functionally related to the acquisition of visual–verbal associations, it should have declined over the course of training, as the learning demands of the task decreased. In contrast, as unimodal trials did not have these learning demands, activity during these trials should not have changed over the course of training. To determine whether this was the case, we conducted a one-way ANOVA (Run 1 vs. Run 2 vs. Run 3) using the contrast [cross-modal – max(unimodal)] as the dependent measure for each run. This showed that the difference in activity between these trial types decreased over the course of training in bilateral occipito-temporal and parietal cortices, left pFC, right precentral gyrus, left putamen and insula, and right pallidum (Figure 3, green overlay; Appendix 2).

The plots and orange overlay in Figure 3 illustrate that, as we expected, in regions showing greater cross-modal than unimodal activity overall, cross-modal activity

decreased over the three training runs, whereas unimodal activity was maintained. Within left pFC, a region which did not show greater cross-modal than unimodal activity overall, a profile of declining cross-modal activity and increasing see-only activity was observed. This may reflect participants starting to covertly recall item pronunciations during see-only trials as training progressed. Overall, this analysis confirmed our expectation that the decreasing learning demands should lead to declining cross-modal activity over training. Maintenance or increasing unimodal trial activity over the course of training demonstrates that declining cross-modal activity did not simply reflect a general reduction in attention.

Relationship between Activity and Performance

Further evidence for the functional involvement of bilateral occipitotemporal and parietal cortices and left pFC in acquiring visual–verbal associations was provided by ROI analyses, comparing see-hear trial activity for items that were learned with items that were not-learned, as described in the Methods section. Six 5-mm radius spherical ROIs were defined in left and right occipitotemporal cortices (inferior occipital, mid-fusiform, anterior fusiform), centered on peak coordinates from the contrast [cross-modal – max(unimodal)] activity, except those in mid-fusiform, which were centered on the VWFA (Cohen et al., 2002) and its right hemisphere homologue. As the peak coordinates for the VWFA are reproducible across individuals and writing systems (Dehaene & Cohen, 2011), we felt confident in defining the VWFA ROI in this way, rather than using a localizer of English word reading relative to rest, for example. Multiple spherical ROIs were constructed in occipitotemporal cortices because [cross-modal – max(unimodal)] activation clusters in these regions were extremely large. Three further ROIs were defined in left and right parietal cortex and left hippocampus; each encompassed the cluster extent-corrected activity in these regions in this [cross-modal – max(unimodal)] contrast. Finally, three ROIs were defined in left pFC (superior frontal, precentral, inferior frontal), centered on peak coordinates from the one-way ANOVA demonstrating change in [cross-modal – max(unimodal)] activity over training. Using these functional ROIs does not constitute "double dipping" as they were defined on the basis of an orthogonal contrast from an unbiased design in which all items (learned and not-learned) were included in both sides of the contrast [cross-modal – max(unimodal)] (see Kriegeskorte, Simmons, Bellgowan, & Baker, 2009, supplementary materials). In each ROI, we conducted a paired t test comparing see-hear trial activity for learned versus not-learned items. Table 1 shows that activity was significantly greater for learned than not-learned items in all ROIs, except left hippocampus, left anterior fusiform, and right anterior fusiform. These analyses support the idea that bilateral occipitotemporal cortices extending to midfusiform, bilateral inferior, and superior parietal cortices

and left pFC were functionally involved in visual–verbal associative learning.

Testing Blocks

Brain regions involved in word reading and object naming are often investigated using silent reading/naming tasks to minimize "task induced effects and avoid the activation of temporal regions caused by subjects processing their own voice" (Mechelli, Gorno-Tempini, & Price, 2003, p. 269). In our study, it was imperative that participants read/named aloud so that we could measure learning success. To examine neural activity during recall, independent of articulation, testing blocks included both see-think trials (covert retrieval) and see-speak trials, which followed the corresponding see-think trial and required participants to overtly articulate the item they had retrieved on the previous trial. Thus, see-think trials made greater demands on the neural systems supporting retrieval of visual–verbal associations, whereas see-speak trials made greater demands on the neural systems for articulation. We therefore used the contrast [see-think – see-speak] to determine which neural systems were involved in retrieving verbal forms from visual forms. As the same item was presented on adjacent see-think and see-speak trials, this contrast also revealed retrieval-related activity over and above visual processing activity.

The contrast [see-think – see-speak] revealed activity in bilateral occipitotemporal, inferior parietal, middle frontal, and inferior frontal cortices (Figure 4, hot colors; Appendix 3).[1] The reverse contrast [see-speak – see-think] activated bilateral STG, bilateral insulae, bilateral postcentral gyri, cuneus, and middle and anterior cingulate, which, as predicted, are regions involved in articulation and processing the sound of one's own voice (Figure 4, cool colors; Appendix 3). Comparing Figures 3 and 4, it is clear that there was a great deal of overlap between activation for retrieving (see-think – see-speak) and acquiring (cross-modal – max(unimodal)) visual–spoken form associations, particularly within bilateral occipitotemporal and parietal cortices. Our analyses therefore provide converging evidence for the importance of these regions in learning to read words and name objects.

Differences in Neural Activity between Learning to Read Words and Name Objects

Having established the brain regions activated when learning visual–verbal associations, we conducted further whole-brain analyses to determine whether these (or other) regions showed differential activity during learning and retrieval for word reading versus object naming.

Training Blocks

We first used whole-brain paired t tests to compare cross-modal (see-hear trial) activity (relative to rest) when

Table 1. ROI Analyses Contrasting See-hear Trial Activity for Learned–Not-learned Items

										Contrast		
											[Object [Learned – Not-learned] – [Word [Learned – Not-learned]]]	
		Center of Mass			*Size*		*[Learned – Not-learned]*					
ROI	*Hemisphere*	*X*	*Y*	*Z*	*Voxels*	*Vol (mm)*	*t(18)*	*p*	*η²*	*t(18)*	*p*	*η²*
Parietal cortex	L	−34	−45	49	768	6128	2.48	.01	.50	<1	*ns*	
Parietal cortex	R	28	−58	49	795	6360	2.05	<.05	.44	<1	*ns*	
Occipitotemporal cortex	L											
Inferior occipital		−46	−76	−10	81	648	3.1	<.01	.59	<1	*ns*	
Mid-fusiform (VWFA)		−42	−56	−12	81	648	2.41	.01	.49	2.27	.02	.47
Anterior fusiform		−30	−34	−24	81	648	1.29	.11	.29	<1	*ns*	
Occipitotemporal cortex	R											
Inferior occipital		44	−78	−10	81	648	4.28	<.001	.71	<1	*ns*	
Mid-fusiform		42	−56	−12	81	648	1.85	<.05	.40	<1	*ns*	
Anterior fusiform		28	−26	−22	81	648	<1	*ns*		<1	*ns*	
Hippocampus	L	−19	−18	−11	318	2544	<1	*ns*		−1.05	*ns*	
Prefrontal cortex	L											
Superior frontal		−22	2	54	81	648	2.85	.005	.56	−1.4	.09	
Precentral		−46	2	54	81	648	3.12	<.01	.59	1.28	.11	
Inferior frontal (triangularis)		−44	36	10	81	648	3.36	<.01	.62	<1	*ns*	
Mid-fusiform defined using peaks from:	L											
Group												
1. [Cross-modal – Unimodal]		−46	−52	−18	81	648	Not tested			2.02	.02	.43
2. [See-think – Rest]		−50	−46	−16	81	648	Not tested			2.02	.02	.43
Individual subjects		*Mean (SD) peak location*										
3. [See-think – Rest]		−44 (4.9)	−52 (6.2)	−15 (4.1)	81	648	Not tested			1.97	.03	.42

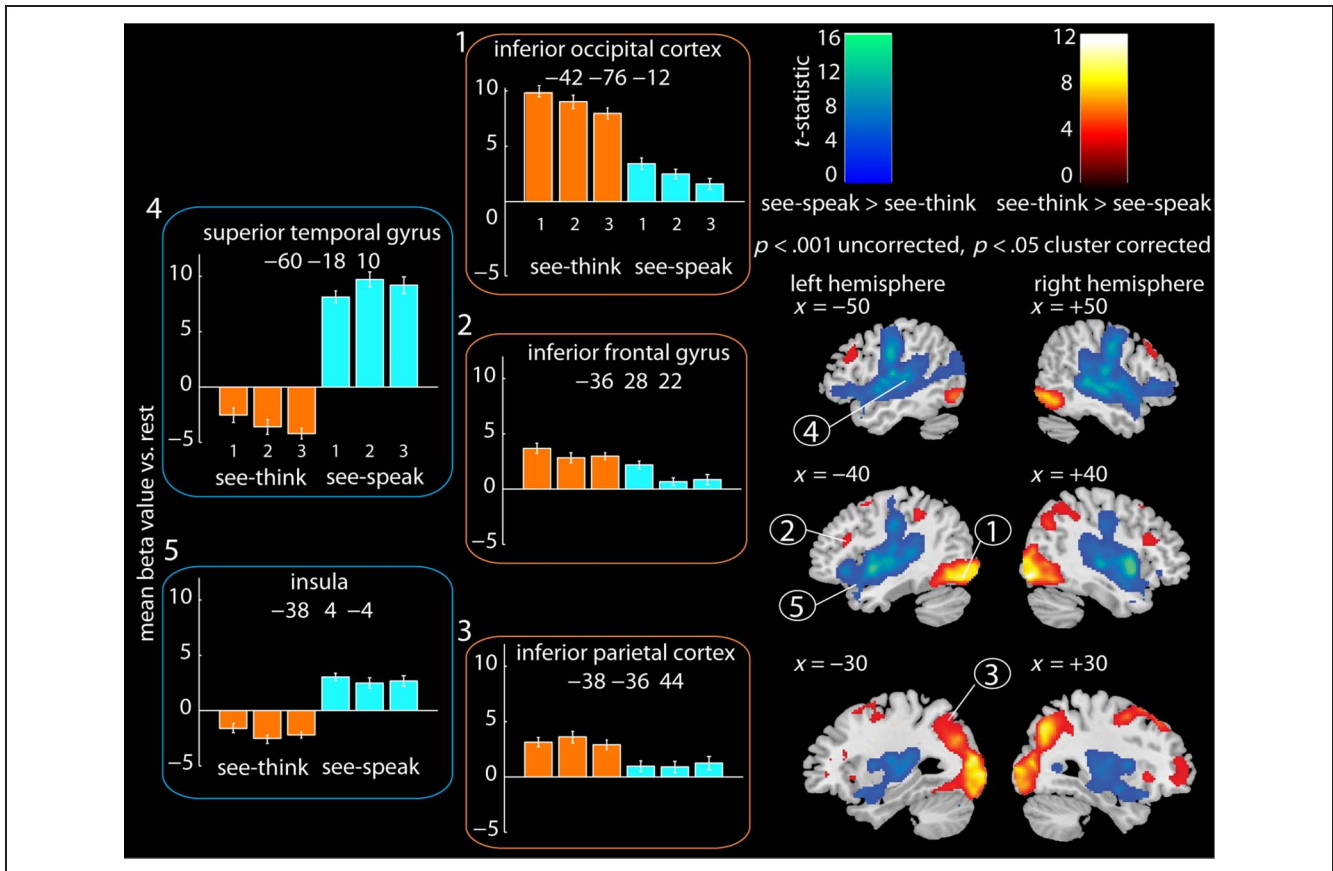*p* values are not corrected for multiple comparisons.

Figure 4. Brain regions involved in recalling spoken from visual forms (hot colors [see-think – see-speak]) and in motoric articulation of spoken forms (cool colors [see-speak – see-think]). Plots show activity for each trial type at peak coordinates obtained from these whole-brain contrasts.

participants learned to read words versus learned to name objects. Natural language experiments that have investigated visual processing of words and objects have had difficulty controlling for differences in the extent to, and ease with, which written word and object stimuli evoke phonological/semantic associations (Price & Devlin, 2011). In contrast, in our experiment, none of the items were given a meaning, phonological associations were always present on cross-modal trials, and these were learned with equal success for words and objects. Activity was greater during cross-modal object trials than cross-modal word trials in mid-to-anterior regions of bilateral medial fusiform gyri, extending into parahippocampal gyrus, and in left inferior frontal gyrus (pars orbitalis) (Figure 5, pale blue overlay; Appendix 4). In addition, left angular gyrus was less deactive during cross-modal object trials than cross-modal word trials. For the reverse contrast, bilateral inferior and superior parietal cortices and left superior frontal gyrus were more active during cross-modal word than cross-modal object trials (Figure 5, pale red overlay; Appendix 4). Thus, whereas bilateral medial fusiform gyri and left inferior frontal gyrus (pars orbitalis) were more engaged during cross-modal processing of whole-item visual–verbal associations (object naming), bilateral parietal cortices and left superior frontal gyrus were more engaged during cross-modal processing of visual–verbal

associations that are componential and systematic (word reading).

While better controlling for automatic activation of linguistic associations, these simple contrasts leave open the possibility that visual differences between the word and object stimuli may have contributed to differential cross-modal activation. We therefore conducted two paired t tests, again across the whole brain, to look for interaction effects reflecting differential cross-modal relative to unimodal activity for objects versus words [object [cross-modal – max(unimodal)] – word [cross-modal – max(unimodal)]] and the reverse. Anterior regions of left and right fusiform gyri and left inferior frontal gyrus (pars orbitalis) were relatively more active for cross-modal than unimodal trials for objects compared with words (Figure 5, bold blue overlay; Appendix 4). The reverse word > object interaction contrast revealed clusters in precuneus, left anterior superior frontal gyrus, right inferior frontal gyrus, and calcarine cortex (Figure 5, bold red overlay; Appendix 4). However, as none of these regions were in fact more active during cross-modal word than object trials and some were deactive relative to rest, we cannot conclude that they were more involved in learning to read words than in learning to name objects. For completeness, Figure 6 (purple overlay) shows that, as well as the differences just reported, there was also a great deal

of overlap between [cross-modal – max(unimodal)] activity for words and objects.

Differential Relationships between Activity and Performance

To further investigate differences in the functional involvement of brain regions activated during learning of the two stimulus types, we conducted an interaction contrast [object [learned – not-learned] – word [learned – not-learned]] in each of the 12 ROIs described earlier. This contrast was nonsignificant in all ROIs except the VWFA, in which we obtained an interaction effect ($p = .02$, not corrected for 12 multiple comparisons) reflecting a bigger difference in see-hear trial activity for learned relative to not-learned items, for objects compared with words (Table 1). To ensure that this was not just a chance finding because of our particular ROI selection, we confirmed this result in several ways. First, we constructed two further minimally overlapping 5-mm spherical ROIs centered on peak coordinates nearest to the VWFA obtained from the overall contrasts: [cross-modal – max (unimodal)] (Mid-fusiform 1), and [see-think – rest] (Mid-fusiform 2). Both of these additional ROIs showed the same interaction effect in favor of greater [learned – not-learned] activity on see-hear trials for objects relative to words ($p = .02$, reaches Bonferroni-corrected significance for two comparisons). Second, the same result was obtained using 5-mm spherical ROIs centered on subjects' individual peaks nearest to the VWFA (defined as -42 -56 -12) from the contrast [see-think – rest] ($p = .03$). These additional ROI analyses are reported in Table 1.

To summarize, many regions did not show differential activity for object–name learning versus learning to read, for example, bilateral occipitotemporal cortices extending to mid-fusiform gyri, dorsal regions of left inferior frontal and precentral gyri, and left hippocampus. However, in left and right anterior fusiform gyri and left inferior frontal gyrus (orbitalis), we observed greater activity during object–name learning than learning to read, and

Figure 5. Brain regions showing differential cross-modal relative to unimodal activity and cross-modal activity relative to rest when learning to read words versus learning to name objects. Pale blue = object [[cross-modal – rest] – word [cross-modal – rest]], pale red = [word [cross-modal – rest] – object [cross-modal – rest]], bold blue = [object [cross-modal – unimodal] – word [cross-modal – unimodal]], bold red = [word [cross-modal – unimodal] – object [cross-modal – unimodal]]. Where bold and light colors overlap, bold colors only are shown. (A) Sagittal slices. (B) Axial slice. (C) Plots showing activity for each trial type at peak coordinates obtained from these whole-brain contrasts.
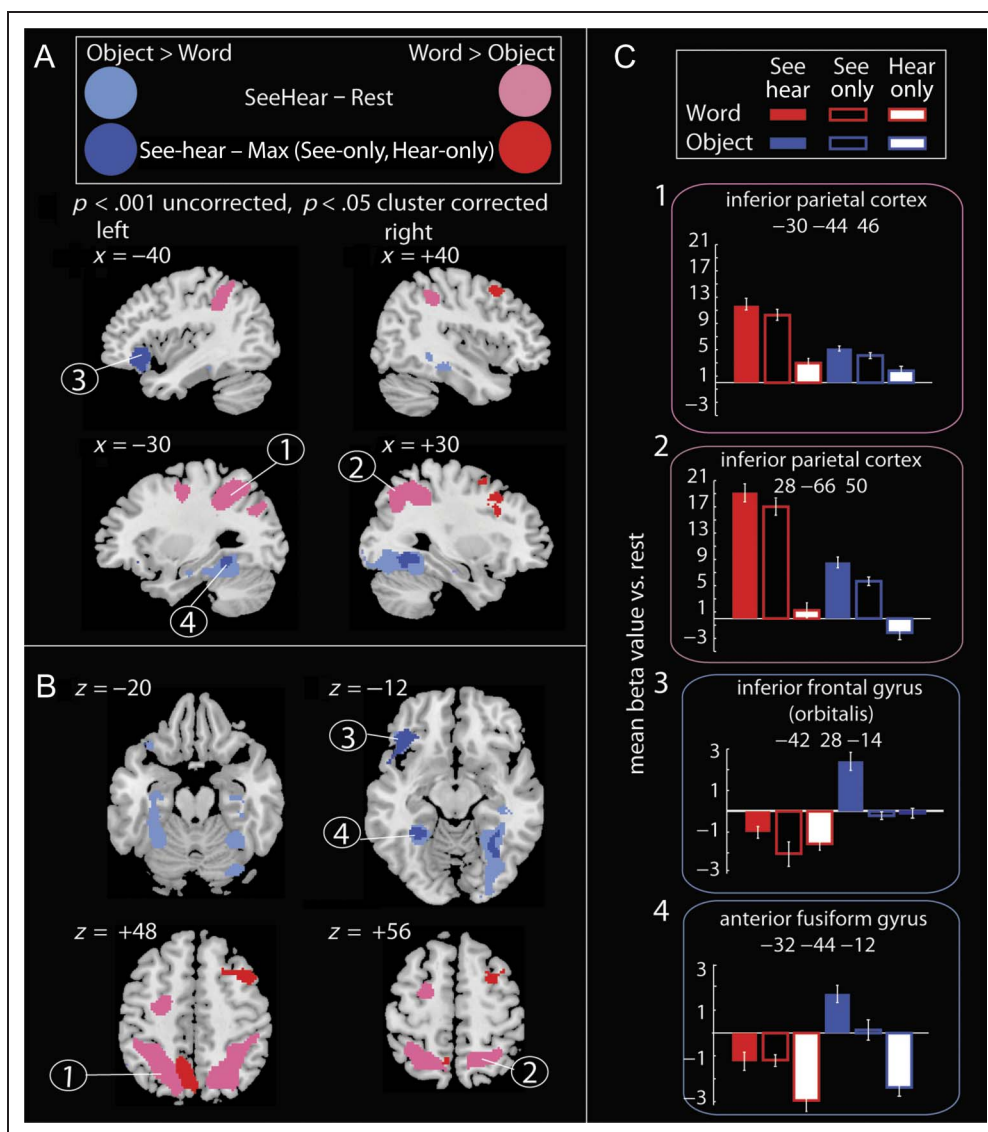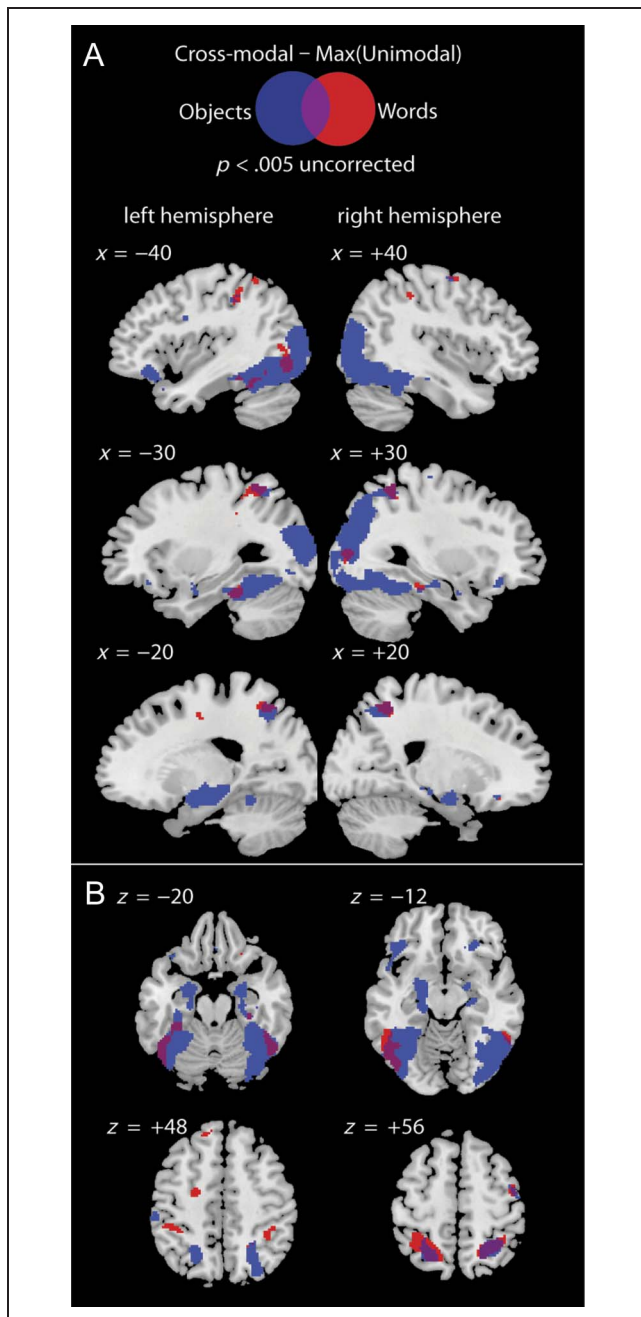
Figure 6. Brain regions showing overlap between cross-modal relative to unimodal activity when learning to read words and when learning to name objects. Blue = object [see-hear – max(see-only, hear-only)], Red = word [see-hear – max(see-only, hear-only)], Purple = overlap between these two contrasts. (A) Sagittal slices and (B) axial slices show whole-brain activations at $p < .005$ voxelwise uncorrected for 19 participants.

this was not driven by visual differences between the stimulus types. Furthermore, ROI analyses indicated that activity in left mid-fusiform gyrus was modulated by performance to a greater extent for object–name learning than learning to read. These results suggest that mid-to-anterior fusiform regions are more engaged when learning holistic, relative to componential, visual–verbal associations.

The opposite profile, greater activity during learning to read than during object name learning, was observed in bilateral inferior and superior parietal cortices. As these regions did not show an interaction effect (greater cross-modal relative to unimodal activity for words vs. objects), we cannot be certain that greater cross-modal activity was driven by the componential nature of the associative learning task, rather than visual differences between the stimuli. However, it is possible that we failed to obtain this interaction for the word > object contrast because participants started to covertly retrieve item pronunciations during see-only trials and that they did this to a greater extent for word reading than object naming. We were motivated to consider this possibility by Price and Devlin's (2011) assertion that, during passive viewing tasks, written words evoke their phonological associations more automatically than pictures of objects. If this were the case, we should obtain activity in the same parietal regions during retrieval of written word pronunciations relative to object names in the test blocks, a prediction we tested in the final set of analyses.

Testing Blocks

We first conducted simple paired t tests between see-think activity (relative to rest) for word reading versus object naming. Only right angular gyrus was highlighted by the [object see-think – word see-think] contrast (Figure 7, pale blue overlay; Appendix 5); however, this region was deactive relative to rest. The reverse contrast revealed that bilateral inferior and superior parietal cortices, middle occipital cortices, precentral gyri, and cerebellum, left superior frontal gyrus, and right inferior frontal gyrus were more active during word see-think than object see-think trials (Figure 7, pale red overlay; Appendix 5).

As with the training data, we next conducted analyses to look for interaction effects to confirm that greater activation during word or object see-think trials was not driven by visual differences between the stimuli. Two paired t tests were conducted across the whole brain: [object naming [see-think – see-speak] – word reading [see-think – see-speak]] and the reverse. As described earlier, demands on the neural systems involved in retrieving spoken forms from visual forms were greater during see-think than see-speak trials, but the same visual item was presented on consecutive see-think and see-speak trials. The object > word contrast revealed an activation cluster in right angular gyrus (Figure 7, bold blue overlay; Appendix 5); however, this region was deactive during all trial types and we cannot therefore make claims about its involvement in retrieving pronunciations for either stimulus type. The reverse contrast revealed that bilateral inferior and superior parietal cortices, right middle frontal gyrus, left precentral gyrus, right inferior temporal gyrus, cerebellum, and left middle occipital cortex were relatively more active during see-think than see-speak trials for word reading compared with object naming (Figure 7, bold

red overlay; Appendix 5). As with the training data, Figure 8 (purple overlay) shows that in addition to the differences just reported there was also extensive overlap between [see-think – see-speak] activity for the two stimulus types.

To summarize, many regions that were active during retrieval of visual–verbal associations, such as occipito-temporal cortices, were not differentially active for object naming versus word reading. In fact, unlike the bilateral anterior fusiform and left inferior frontal (orbitalis) activity we observed during object name learning, no regions were relatively more engaged in retrieving object names than in reading words. However, bilateral inferior and superior parietal cortices and left superior frontal gyrus were more active when reading the words than when retrieving object names, and interaction contrasts confirmed that this effect was not driven by visual differences between the stimulus types. Furthermore, Figure 9 shows that bilateral parietal activation for word > object retrieval overlapped with activation that was greater for word relative to object cross-modal training trials. This consistency across analyses provides converging evidence that bilateral inferior and superior parietal cortices are more engaged when learning and retrieving the componential symbol–sound mappings from which the written words were constructed than when retrieving holistic object–name associations.

Figure 7. Brain regions showing differential see-think relative to see-speak activity and see-think activity relative to rest, when reading words versus naming objects. Pale blue = [object [see-think – rest] – word [see-think – rest]], pale red = [word [see-think – rest] – object [see-think – rest]], bold blue = [object [see-think – see-speak] – word [see-think – see-speak]], bold red = [word [see-think – see-speak] – object [see-think – see-speak]]. (A) Sagittal slices. (B) Axial slices. (C) Plots showing activity for each trial type at peak coordinates obtained from these whole-brain contrasts.
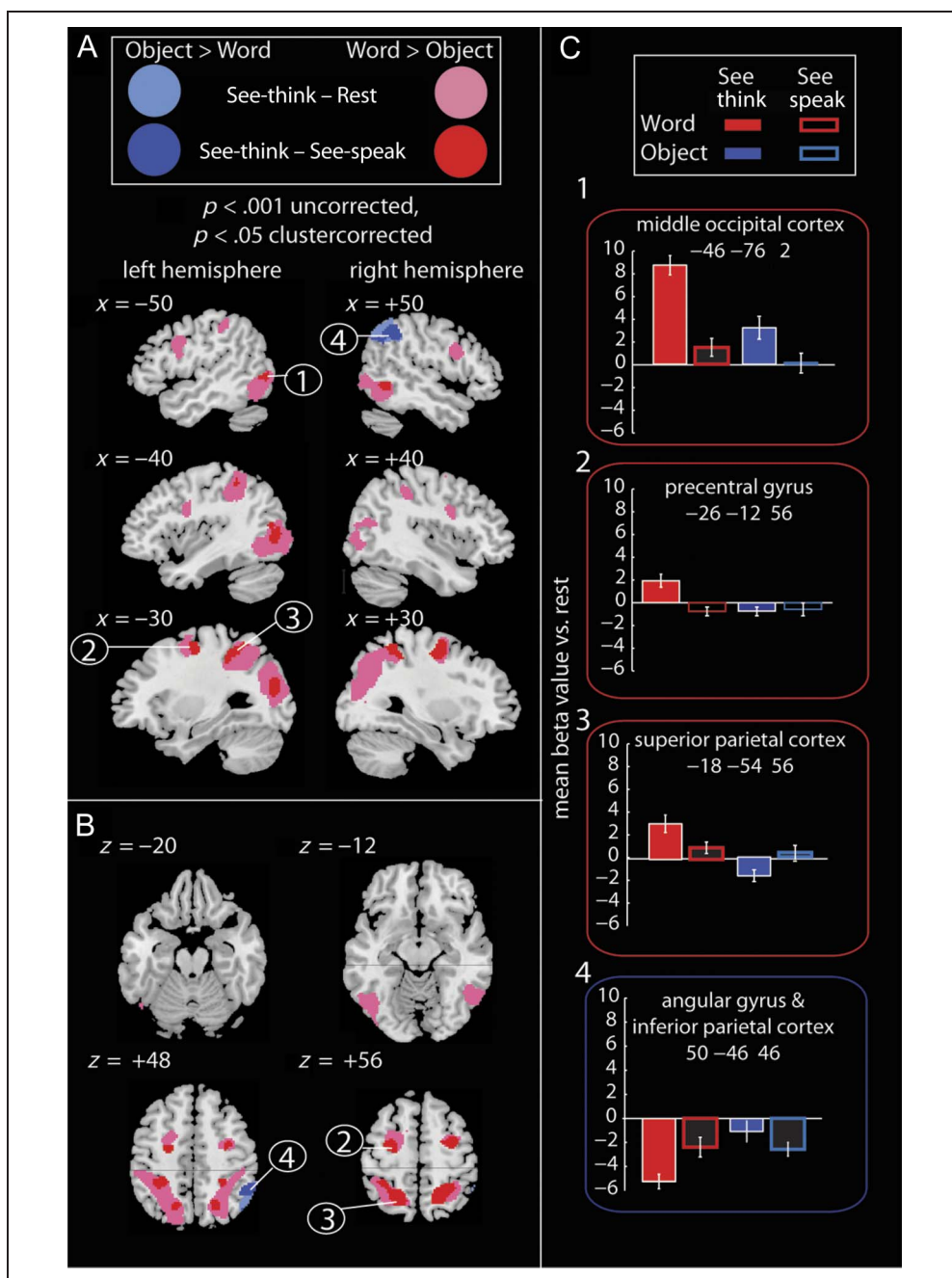
DISCUSSION

Our fMRI training study revealed that bilateral occipito-temporal and parietal cortices, left hippocampus, and left pFC were activated during the acquisition of visual–verbal associations that are important both for learning to read words and name objects. Activity in these brain regions was greater when visual and spoken forms were
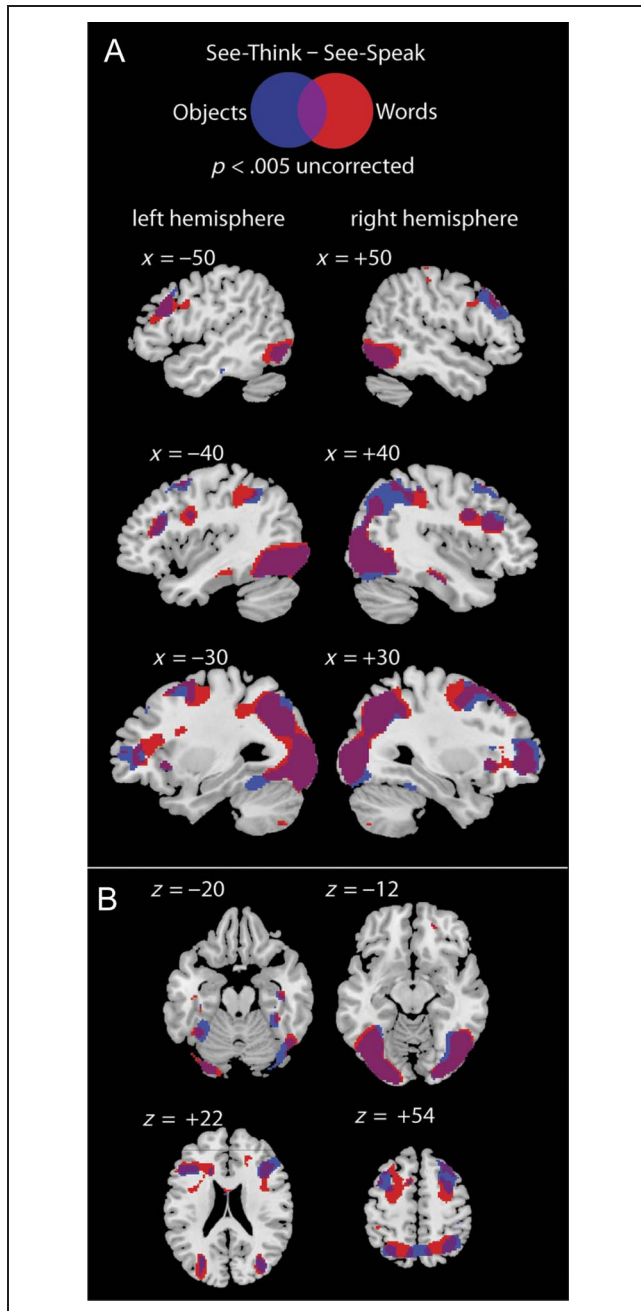

Figure 8. Brain regions showing overlap between see-think relative to see-speak activity when reading words and when naming objects. Blue = object [see-think – see-speak], Red = word [see-think – see-speak], Purple = overlap between these two contrasts. (A) Sagittal slices and (B) axial slices show whole-brain activations at $p < .005$ voxel-wise uncorrected for 19 participants.
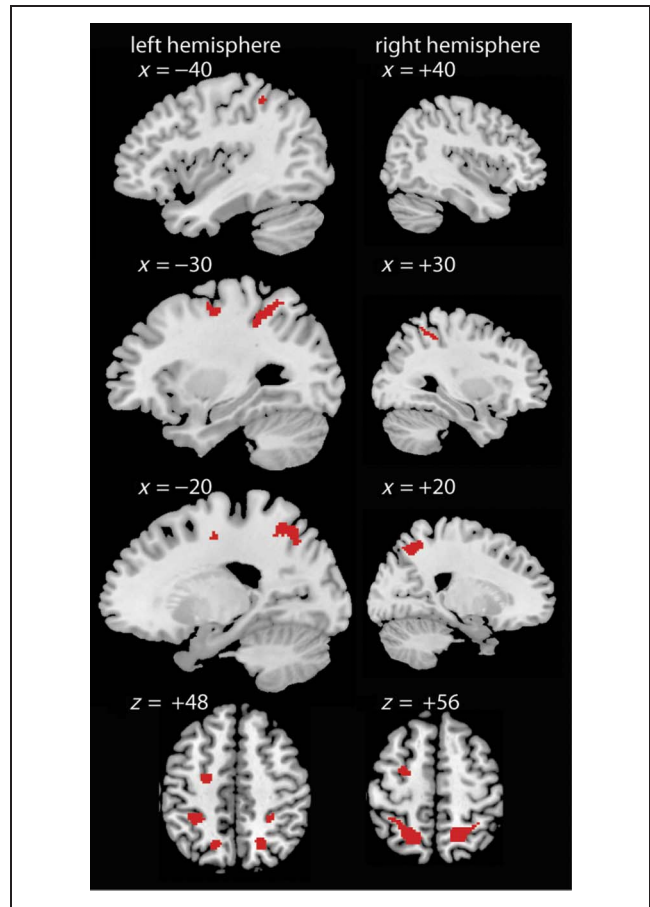
Figure 9. Overlap between activity for two contrasts (both $p < .001$ voxelwise uncorrected and $p < .05$ FWE cluster-corrected): (1) [word [cross-modal – rest] – object [cross-modal – rest]] and (2) [word [see-think – see-speak] – object [see-think – see-speak]].

presented in conjunction, providing a learning opportunity, than when they were presented in isolation. In addition, this cross-modal activity declined over the course of training as participants acquired the correct associations and the learning demands decreased. Finally, cross-modal activity in all of these regions, except the hippocampus, was greater for visual–spoken form associations that were learned than for those that were not learned. The similarity between the network of regions activated during this experiment and in previous neuroimaging studies of reading and object naming (Taylor et al., 2013; Price et al., 2005) confirms the validity of our laboratory model of learning to read and learning to name objects. More specifically, our results confirm that occipitotemporal cortex activity is influenced by the linguistic associations of visual stimuli (Dehaene & Cohen, 2011; Price & Devlin, 2011) and support research implicating left parietal cortex in learning visual–verbal associations (Breitenstein et al., 2005; Cornelissen et al., 2004; Hashimoto & Sakai, 2004).

Despite participants' equal overall success at learning to read words and name objects, there were important behavioral differences between the two stimulus types. The likelihood of an object being named successfully at

final test was predicted by the number of times that object was named correctly during training, whereas final test success for word reading was equally good for words read once, twice, or thrice correctly during training. This reflects the fact that participants learned to read the words by associating the symbols from which they were composed with individual sounds, and thus learning about one word helped learning about others, whereas objects had a holistic and arbitrary mapping to their name. Further evidence that participants extracted componential, systematic symbol sounds when learning to read the words comes from their success at reading untrained words in the final test task. Our laboratory model of cross-modal learning thus captured the key distinctions between learning to read words and name objects. Using this model, we conducted several tests to determine whether these different computational demands resulted in neural specializations for learning to read words and name objects.

Specialization for Object–Name Learning in Mid-to-anterior Fusiform Gyri

Cross-modal associative learning of object names activated bilateral anterior fusiform gyri and left inferior frontal gyrus (pars orbitalis) more than learning to read. More posterior fusiform regions (including the VWFA) were equivalently engaged during learning and recall of object names and written word pronunciations, and ROI analyses revealed that VWFA activity during cross-modal learning predicted subsequent object naming but not word reading success. In combination, these results seem somewhat at odds with the idea that the left mid-fusiform gyrus is specialized for analyzing componential orthographic form (Vinckier et al., 2007; Dehaene et al., 2005). Instead, our data suggest that bilateral mid-to-anterior fusiform gyri preferentially process whole-object visual forms and contribute to object name learning.

One potential objection to this conclusion is that left vOT may contain subpopulations of neurons showing opposing response profiles, that is, some more responsive to written words than objects and some, the reverse. Dehaene and Cohen (2011) levied such a criticism against reports of equivalent left vOT activation for words and faces (Mei et al., 2010) or objects (Price & Devlin, 2003). In the current study, rather than obtaining equivalent left vOT activity for word reading and object naming, we obtained greater left anterior fusiform activation for objects relative to words, and both group and single-subject ROI analyses indicated that left mid-fusiform activity was more predictive of object naming than word reading success. Nonetheless, our group analysis of smoothed fMRI data may still have overlooked the opposite response profile if it was seen in subpopulations of vOT neurons in nonoverlapping locations across participants. Hence, future work should seek to confirm these findings with high-resolution fMRI and single-subject analyses, as, for example, used by Mano et al. (2013).

What implications do our findings have for the role of the left fusiform in reading? The putative VWFA is consistently activated when participants read words from a variety of languages and orthographies (Baker et al., 2007; Bolger, Perfetti, & Schneider, 2005; Cohen et al., 2000). Furthermore, the sensitivity of this region to printed words increases as children learn to read (Ben-Shachar et al., 2011). However, these findings implicate the VWFA in the recognition of word forms, not in letter-sound decoding. Further evidence for a role for the left mid-anterior fusiform gyrus in whole-word rather than letter-sound reading comes from Mechelli et al. (2005), who found that coupling between this region and left inferior frontal gyrus (triangularis) was greater during irregular word relative to regular word and pseudoword reading. Irregular words (e.g., YACHT, THYME) cannot be read correctly using letter-sound rules and are therefore more dependent on whole-word knowledge than regular words or pseudowords (Coltheart et al., 2001; Plaut, McClelland, Seidenberg, & Patterson, 1996). In light of this background literature, our findings that (a) anterior fusiform gyri are more engaged when learning object names than when learning to read and (b) ROI analyses indicated that VWFA activity is modulated by learning success for object naming but not word reading lead us to suggest that the role of the left mid-to-anterior fusiform gyrus in reading is to process the visual forms of whole items and associate them with their spoken forms.

One possible interpretation of this conclusion, with regards to reading, is that the left fusiform gyrus contains whole-word representations akin to those in the orthographic lexicon of the dual route cascaded (DRC) model of reading (Glezer et al., 2009; Kronbichler et al., 2007; Coltheart et al., 2001). However, it is also possible that what drives responses in this region is not bottom–up activation of stored visual representations of whole-item forms, but top–down interactions with brain regions processing the linguistic information associated with these visual forms, in line with Price and Devlin's (2011) interactive account of vOT processing. Our results move this debate forward by showing that (1) anterior fusiform gyri are more engaged by object naming than word reading even when the necessity and difficulty of visual–spoken form mappings is equated, and thus, the extent of top–down feedback should be equivalent, (2) that this pattern is present from the very earliest stages of learning, and (3) that activity in the left mid-fusiform region is related to the success of whole object–name learning. Our results therefore imply that these mid-anterior regions of fusiform gyri preferentially contribute to processing whole-item visual forms and associating them with linguistic representations.

Specialization for Learning to Read Words in Parietal Cortices

Unlike in left mid-fusiform gyrus, activity in bilateral parietal cortices predicted word reading and object naming

success. This suggests a role for these regions in learning both componential and holistic visual–verbal associations, supporting evidence for their involvement in cross-modal processing in multiple domains (Behrmann, Geng, & Shomstein, 2004). However, the same regions in bilateral inferior and superior parietal cortices were more active during both cross-modal learning and retrieval for word reading relative to object naming. An interaction analysis demonstrated that this was not driven by perceptual differences between the stimuli, but rather by the particular processes involved in covert retrieval of phonological forms. Thus, parietal cortices were engaged when participants learned and retrieved systematic and componential letter sounds, relative to arbitrary and holistic object names. A role for these left parietal regions in letter-sound translation processes is supported by a variety of evidence outlined in the Introduction. For example, left inferior and superior parietal cortices are active during pseudoword relative to word reading (Taylor et al., 2013), and left inferior parietal cortex is active when participants manipulate (Booth et al., 2003, 2007) and learn spelling–sound mappings (Hashimoto & Sakai, 2004). Surface dyslexic patients, who rely on letter-sound rather than whole-word reading strategies, also activate this region to a greater extent than control subjects during word reading (Wilson et al., 2009).

A question remains, however, as to why left parietal cortex is more engaged during tasks that require componential relative to holistic visual–verbal mappings. Several studies implicate a particular subregion of parietal cortex, the left SMG, in tasks that require phonological manipulation. For example, left SMG is active when participants make rhyme judgments about visually presented words (Sliwinska, Khadilkar, Campbell-Ratcliffe, Quevenco, & Devlin, 2012; Seghier et al., 2004; Devlin, Matthews, & Rushworth, 2003). Left posterior STG has also been suggested to be important for linking visual and verbal stimuli; in particular, a review by Blomert (2011) suggests that this region automatically integrates letters and sounds. However, the word > object activation we observed was superior to the SMG and STG. This is also the case for the parietal activation reported by Taylor et al. (2013), Wilson et al. (2009), Booth et al., (2003, 2007), and Hashimoto and Sakai (2004). We therefore suggest that, whereas left SMG seems to play a role in manipulating verbal information and left STG may be a key region for the convergence of visual and auditory information, our and others' results suggest that left inferior parietal cortex is particularly engaged when information has to be componentially or sequentially mapped between the visual and verbal modality.

Parietal cortices are part of the dorsal visual processing stream that has been proposed to be involved in using visual information to guide actions (Goodale & Milner, 2010; Milner & Goodale, 2008) or in visually guided attention (Gottlieb, 2007; Husain & Nachev, 2007) including saccadic eye movements (Simon, Mangin, Cohen, Le Bihan, & Dehaene, 2002). When learning to read the novel words in our study, participants mapped each letter to its corresponding sound, which involved allocating attention to the portion of the visual input currently being translated. Furthermore, although reading short words and pseudowords in a familiar alphabet does not typically induce multiple saccades (Rayner, Raney, & Sereno, 1996), reading words written in unfamiliar symbols may have done so. Therefore, learning to read the novel words may have made greater demands on the serial allocation of attention and potentially on mechanisms controlling eye movements than learning to map whole objects to their names.

Cohen, Dehaene, Vinckier, Jobert, and Montavont (2008) obtained superior parietal cortex activity for reading words presented in visual noise as opposed to on a clear background. As this should induce a serial letter-sound rather than whole-word reading strategy, this activity was thought to reflect increased serial visual attention demands. Deficits in visual attention (Lobier, Zoubrinetzky, & Valdois, 2012; Bosse, Tainturier, & Valdois, 2007) and reduced activation in left inferior and superior parietal cortex (Reilhac, Peyrin, Démonet, & Valdois, 2013) have also been argued to play a role in some forms of developmental dyslexia. However, these studies either used words or letters (Cohen et al., 2008; Bosse et al., 2007), which automatically activate their linguistic associations (Ziegler, Pech-Georgel, Dufau, & Grainger, 2010), or involved categorizing visual stimuli (e.g., indicating how many elements of a particular family of stimuli were present in the array; Reilhac et al., 2013; Lobier et al., 2012), which may instigate an implicit naming strategy. As these studies did not separate purely visual from visual–verbal processing, their results are not inconsistent with our proposal that left inferior and superior parietal cortices are particularly involved when componential/serial processing involves mapping between the visual and verbal domains.

A serial letter-to-sound translation process forms part of most computational models of reading. In the DRC model, the sublexical route translates spellings to sounds from left to right using a set of grapheme-to-phoneme conversion rules (Coltheart et al., 2001; Rastle & Coltheart, 1999). Similarly, the sublexical route of the connectionist dual process model (Perry, Ziegler, & Zorzi, 2007) translates spellings to sounds via a two-layer network that incorporates a serial left-to-right process whereby the letter string is first parsed into graphemes (e.g., church → CH, UR, CH). Perry et al. (2007, p. 300) state that this graphemic parsing process "is likely to involve *left-to-right shifts of spatial attention* over the letter string." With regards to the triangle model, Plaut (1999) developed a recurrent network version that resolves grapheme–phoneme correspondences serially, from left-to-right across the letter string. As the network learns about spelling-to-sound correspondences, it is able to pronounce more than one grapheme in a single fixation. However,

when the network encounters difficulty with a grapheme–phoneme mapping, the problematic grapheme is refixated, thus allowing the network to make use of its more extensive experience reading graphemes at fixation. Although adult skilled readers typically do not refixate monosyllabic words, Plaut et al. suggested that refixations could instead be interpreted as covert allocations of attention.

In all three of these computational models, serial processing is more important for pseudoword than word reading. In the DRC and connectionist dual process models, this is because words are represented as wholes in the lexical route, whereas pseudoword reading must be accomplished by the sublexical route. In Plaut's (1999) "refixation" version of the triangle model, more frequently experienced combinations of letters (e.g., words relative to pseudowords) require fewer fixations because the model learns about statistical regularities between spelling and sound. Thus, in these models, the process of mapping visual stimuli to their verbal output componentially, rather than as a whole, necessitates the serial allocation of attention. This concurs with our finding that parietal cortices were more engaged when retrieving pronunciations of, relative to viewing, written words as compared with objects. It is the process of having to componentially map vision to language that drives serial visual attention.

We therefore suggest that learning to read words increased activity in parietal cortices relative to learning object names because letters in the novel words had to be translated to their corresponding sound, necessitating sequential allocation of attention. With more extensive training on the novel words in the current study, it is probable that a whole-word rather than letter-sound reading strategy would be favored. Given the current results, this should drive greater reliance on occipitotemporal rather than parietal cortices to accomplish the visual–verbal mapping. More generally, the division of labor between parietal and occipitotemporal cortices during reading could provide an important marker of the extent to which children or adult second language learners are reliant on letter-sound versus whole-word reading processes.

Conclusions

Our fMRI paradigm provides a unique window into the processes that support the earliest stages of learning to read words and name objects. Our data shed new light on debates about neural specialization for reading acquisition in two ways. First, we showed that bilateral mid-to-anterior fusiform gyri were preferentially engaged in learning about whole forms of visual stimuli and their linguistic associations, something that should be crucial for skilled efficient reading. Second, we demonstrated that inferior and superior parietal cortices were more engaged when visual forms had to be broken down into parts to be associated with spoken forms, something crucial in the beginning stages of reading development. Current best practice in literacy acquisition seeks to teach children such componential letter-sound knowledge in the earliest stages of learning to read alphabetic scripts. We thus suggest that parietal cortices deserve greater attention in research programs aiming to elucidate the neural mechanisms underpinning reading development.

APPENDICES

Appendix 1. Brain Regions Showing Greater Activity during Cross-modal (See-hear) Training Trials than the Maximum Activity Observed during Unimodal (See-only, Hear-only) Training Trials

Brain Region (AAL)	Hemisphere	x	y	z	Voxels	Z Value	Cluster-level p Value
See-hear – Max (See-only, Hear-only), Collapsed across Run and Task							
Superior parietal cortex	R	26	−50	58	795	6.04	.001
Supramarginal gyrus		42	−36	44			
Superior occipital cortex		28	−72	44			
Inferior occipital ortex	L	−46	−76	−10	2614	5.77	<.001
Fusiform gyrus		−46	−52	−18			
Fusiform gyrus		−38	−38	−18			
Middle occipital cortex		−44	−78	12			
Middle temporal gyrus		−50	−68	4			
Middle occipital cortex		−26	−78	36			
Middle occipital cortex		−34	−80	0			

Appendix 1. (*continued*)

Brain Region (AAL)	Hemisphere	x	y	z	Voxels	Z Value	Cluster-level p Value
Middle occipital cortex		−38	−90	14			
Middle occipital cortex		−34	−84	24			
Superior parietal cortex	L	−22	−56	56	766	5.33	<.001
Inferior parietal cortex		−38	−36	40			
Supramarginal gyrus		−58	−24	40			
Middle occipital cortex	R	44	−78	10	2096	5.11	<.001
Inferior temporal gyrus		54	−58	−16			
Fusiform gyrus		40	−44	−20			
Fusiform gyrus		28	−26	−22			
Middle occipital cortex		28	−82	8			
Inferior occipital cortex		48	−78	−10			
Lingual gyrus	R	32	−78	−16			
Inferior temporal gyrus		56	−60	−4			
Lingual gyrus	L	−12	−32	−4	318	4.65	<.05
Amygdala		−26	−4	−18			
Hippocampus		−22	−16	−12			
Precentral gyrus	R	46	−6	56	99	4.59	.05
Middle frontal gyrus		32	−2	54			

$p < .001$ and cluster-level FWE-corrected at $p < .05$. All peaks > 12 mm apart are reported.

Appendix 2. Brain Regions in which the Difference in Activity between Cross-modal and Unimodal Training Trials Changed over the Three Training Runs

Brain Region (AAL)	Hemisphere	x	y	z	Voxels	Z Value	Cluster-level p Value
One-way ANOVA Examining the Effect of Run (1 vs. 2 vs. 3) on the Contrast: See-hear – Max (See-only, Hear-only)							
Middle occipital cortex	L	−28	−78	26	4525	6.01	<.001
Inferior occipital cortex		−42	−78	−12			
Inferior occipital cortex		−46	−66	−4			
Fusiform gyrus		−44	−54	−18			
Lingual gyrus		−36	−90	−12			
Inferior parietal cortex		−56	−30	46			
Superior parietal cortex		−20	−64	50			
Fusiform gyrus		−28	−60	−12			
Middle occipital cortex		−36	−84	4			
Fusiform gyrus		−32	−46	−14			
Inferior parietal cortex		−42	−40	42			
Superior parietal cortex		−24	−48	46			
Supramarginal gyrus		−50	−28	32			

Appendix 2. (*continued*)

Brain Region (AAL)	Hemisphere	x	y	z	Voxels	Z Value	Cluster-level p Value
Middle frontal gyrus	L	−24	4	54	3995	5.71	<.001
Precentral gyrus		−48	2	24			
Precentral gyrus		−42	2	50			
Supplementary motor		−10	12	50			
Supplementary motor		−6	4	64			
Supplementary motor		6	6	56			
Inferior frontal gyrus		−48	34	16			
Postcentral gyrus		−48	−6	38			
Mid-cingulate		8	16	44			
Postcentral gyrus		−54	−10	48			
Inferior frontal gyrus		−50	20	24			
Precentral gyrus	R	34	−2	50	269	5.62	.001
Middle occipital cortex	R	32	−62	30	3379	5.21	<.001
Cerebellum		40	−64	−20			
Inferior temporal gyrus		56	−62	−12			
Inferior occipital cortex		38	−72	−8			
Middle occipital cortex		40	−84	16			
Cerebellum		16	−76	−20			
Inferior occipital cortex		48	−76	−2			
Inferior occipital cortex		34	−88	−2			
Fusiform gyrus		42	−50	−22			
Inferior occipital cortex		24	−94	−10			
Superior occipital cortex		28	−78	40			
Calcarine cortex		20	−100	0			
Fusiform gyrus		24	−64	−12			
Cerebellum		−2	−72	−18			
Cerebellum		4	−82	−22			
Putamen	L	−20	10	2	555	4.89	<.001
Insula		−28	22	2			
Inferior frontal gyrus		−40	22	−2			
White matter	R	12	6	2	244	4.19	<.01
White matter		20	6	16			

$p < .001$ and cluster-level FWE-corrected at $p < .05$. All peaks > 12 mm apart are reported.

Appendix 3. Brain Regions Activated When Covertly Recalling Spoken Forms from Visual Forms [See-think – See-speak] and When Articulating Spoken Forms [See-speak – See-think]

Brain Region (AAL)	Hemisphere	x	y	z	Voxels	Z Value	Cluster-level p Value
See-think – See-speak							
Inferior occipital cortex	bilateral	40	−82	−10	9617	6.34	<.001
Middle occipital cortex		−28	−88	14			
Middle occipital cortex		36	−78	10			
Lingual gyrus		−34	−86	−14			
Superior occipital cortex		30	−70	42			
Middle occipital cortex		−26	−72	32			
Lingual gyrus		22	−92	−6			
Precueus		4	−64	46			
Fusiform gyrus		−36	−54	−16			
Inferior parietal cortex		36	−44	42			
Inferior parietal cortex		−38	−36	44			
Middle frontal gyrus	R	28	34	48	660	4.76	.001
Middle frontal gyrus		28	6	52			
Insula	L	−24	26	0	782	4.35	.005
Inferior frontal gyrus		−36	28	22			
Inferior frontal gyrus		−54	20	32			
Superior frontal gyrus		−28	46	0			
Middle frontal gyrus	R	28	24	−2	538	4.12	<.05
Middle frontal gyrus		28	58	4			
Middle frontal gyrus	L	−24	22	58	334	4.06	<.05
Middle frontal gyrus	L	−28	0	52			
Insula	bilateral	40	10	−6	42041	7.08	<.001
Superior temporal gyrus		58	−30	8			
Superior temporal gyrus		62	−6	8			
Insula		−38	4	−4			
Superior temporal gyrus		−60	−18	10			
Rolandic operculum		48	−16	14			
Superior temporal gyrus		48	−20	−2			
Postcentral gyrus		−50	−12	32			
Cuneus		16	−82	34			
Superior temporal gyrus		60	−10	−6			
Superior temporal gyrus		−58	−6	−6			
Cuneus		6	−94	14			
Rolandic operculum		−46	−4	6			
Superior temporal pole		58	6	0			
Superior temporal gyrus		−64	−28	16			
Lingual gyrus		4	−74	0			

Appendix 3. (*continued*)

Brain Region (AAL)	Hemisphere	x	y	z	Voxels	Z Value	Cluster-level p Value
See-speak – See-think							
Lingual gyrus		20	−52	2			
Postcentral gyrus		50	−8	36			
Lingual gyrus		2	−66	10			
Mid-cingulate	bilateral	4	14	34	9343	6.42	<.001
Superior frontal gyrus		−10	60	24			
Supplementary motor		−6	20	60			
Supplementary motor		8	12	64			
Mid-cingulate		−4	−12	40			
Mid-cingulate		14	−30	46			
Precentral gyrus		20	−30	64			
Anterior cingulate		−8	50	10			
Supplementary motor		4	0	64			
White matter		−16	−26	58			
Mid-cingulate		−12	−28	42			
Anterior cingulate		2	38	4			
Olfactory bulb		0	18	−8			
Superior frontal gyrus		18	56	26			
Superior frontal gyrus		−8	46	42			
Superior frontal gyrus		14	30	50			

$p < .001$ and cluster-level FWE-corrected at $p < .05$. All peaks > 12 mm apart are reported.

Appendix 4. Differences in Cross-modal Activity for Learning to Read Words versus Learning to Name Objects

Brain Region (AAL)	Hemisphere	x	y	z	Voxels	Z Value	Cluster-level p Value
Objects [See-hear – Rest] > Words [See-hear – Rest]							
Inferior frontal gyrus	L	−48	28	−12	406	8.84	<.01
Inferior frontal gyrus		−36	26	−20			
Fusiform gyrus	R	34	−52	−16	1495	7.9	<.001
Fusiform gyrus		30	−76	−14			
Parahippocampal gyrus		36	−26	−16			
Inferior occipital cortex		22	−98	−8			
Fusiform gyrus	L	−30	−50	−14	576	4.59	<.01
Fusiform gyrus		−34	−16	−22			
Fusiform gyrus		−34	−34	−22			

Brain Region (AAL)	Hemisphere	x	y	z	Voxels	Z Value	Cluster-level p Value
Words [See-hear – Rest] > Objects [See-hear – Rest]							
Precuneus	R	12	−56	54	1533	5.39	<.01
Inferior parietal cortex		32	−44	46			
Superior parietal cortex		28	−66	50			
Inferior parietal cortex		46	−34	48			
Inferior parietal cortex	L	−30	−44	46	2104	4.82	.001
Superior parietal cortex		−24	−56	50			
Middle occipital cortex		−18	−66	52			
Precuneus		−12	−62	52			
Inferior parietal cortex		−46	−36	42			
Middle occipital cortex		−28	−76	36			
Precentral gyrus	L	−26	−10	46	395	4.52	<.05
Objects [See-hear – Max (See-only, Hear-only)] > Words [See-hear – Max (See-only, Hear-only)]							
Fusiform gyrus	L	−32	−44	−12	126	5.4	<.05
Inferior frontal gyrus	L	−42	28	−14	562	4.49	<.01
Inferior frontal gyrus		−52	20	−8			
Rolandic operculum		−52	8	2			
Fusiform gyrus	R	28	−56	−8	190	4.29	<.05
Inferior frontal cortex	R	30	30	−8	99	4.18	<.01
White matter		22	42	−6			
Words [See-hear – Max (See-only, Hear-only)] > Objects [See-hear – Max (See-only, Hear-only)]							
Precuneus	Bilateral	−10	−58	46	810	5.89	.001
Precuneus		−8	−66	48			
Precuneus		8	−64	34			
Superior frontal gyrus	L	−22	56	4			
Inferior frontal gyrus	R	34	22	32	368	4.28	.001
Middle frontal gyrus		32	20	44			
Middle frontal gyrus		44	14	50			
Middle frontal gyrus		28	6	56			
Calcarine cortex	Bilateral	−8	−84	4	639	4.22	<.001
Calcarine cortex		16	−84	8			
Lingual gyrus		6	−78	−2			

$p < .001$ and cluster-level FWE-corrected at $p < .05$. Top 5 peaks > 8 mm are reported.

Appendix 5. Differences in Activity for Recalling Spoken from Visual Forms when Reading Words versus Naming Objects

Brain Region (AAL)	Hemisphere	x	y	z	Voxels	Z Value	Cluster-level p Value
Objects [See-think – Rest] > Words [See-think – Rest]							
Angular gyrus	R	46	−66	44	708	3.98	<.001
Inferior parietal cortex		56	−46	42			
Inferior parietal cortex		52	−50	54			
Words [See-think – Rest] > Objects [See-think – Rest]							
Middle occipital cortex	L	−40	−82	4	4850	6.14	<.001
Superior parietal cortex		−22	−64	52			
Middle occipital cortex		−30	−80	18			
Middle occipital cortex		−26	−76	30			
Inferior temporal gyrus		−46	−62	−8			
Inferior parietal cortex		−40	−40	50			
Middle occipital cortex	R	32	−50	54	3006	5.51	<.001
Superior parietal cortex		20	−56	56			
Inferior temporal gyrus		48	−58	−6			
Middle occipital cortex		28	−70	34			
Middle occipital cortex		48	−74	0			
Middle temporal gyrus		40	−68	18			
Supramarginal gyrus		40	−34	44			
Middle occipital cortex		40	−82	6			
Calcarine cortex		26	−72	6			
Precentral gyrus	R	28	−4	56	379	4.9	<.01
Cerebellum	Bilateral	−6	−76	−42	278	4.65	<.05
Cerebellum		10	−76	−40			
Superior frontal gyrus	L	−26	−4	56	607	4.61	<.01
Superior frontal gyrus		−20	−4	56			
Inferior frontal gyrus	R	46	6	26	330	4.54	<.05
Precentral gyrus	L	−52	6	36	392	4.09	.05
Inferior frontal gyrus		−44	4	26			
Objects [See-think – See-speak] > Words [See-think – See-speak]							
Inferior parietal cortex	R	50	−46	46	343	4.32	<.01
Supramarginal gyrus		58	−44	46			
Angular gyrus		50	−58	36			
Words ([See-think – See-speak] > Objects [See-think – See-speak]							
Middle frontal gyrus	R	32	−4	54	240	5.25	<.01
Inferior parietal cortex	R	26	−46	54	440	5.22	.001
Superior parietal cortex		16	−60	56			

Appendix 5. (*continued*)

Brain Region (AAL)	Hemisphere	x	y	z	Voxels	Z Value	Cluster-level p Value
Words ([See-think – See-speak] > Objects [See-think – See-speak]							
Superior parietal cortex	L	−18	−54	56	498	5.16	<.01
Inferior parietal cortex		−32	−42	52			
Inferior parietal cortex		−32	−34	40			
White matter		−22	−44	44			
Precentral gyrus	L	−26	−12	56	213	4.29	<.05
Inferior temporal gyrus	R	52	−54	−6	175	4.2	<.05
Cerebellum	Bilateral	14	−72	−42	221	4.16	<.05
Cerebellum		−4	−76	−38			
Middle occipital cortex	L	−40	−76	2	530	4.12	.01
Middle occipital cortex		−36	−72	14			

$p < .001$ and cluster-level FWE-corrected at $p < .05$. Top 5 peaks > 8 mm are reported.

Acknowledgments

Note

1. Reduced activity for see-speak relative to see-think trials in early visual areas may in part arise from repetition suppression effects. See-think trials are the first presentation and see-speak trials are the second presentation of a specific item on successive trials (cf. Kherif et al., 2010; Andrews & Ewbank, 2004).

REFERENCES

Andrews, T. J., & Ewbank, M. P. (2004). Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *Neuroimage, 23,* 905–913.

Ashburner, J., & Friston, K. J. (1997). Multimodal image coregistration and partitioning–A unified framework. *Neuroimage, 6,* 209–217.

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *Neuroimage, 26,* 839–851.

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, U.S.A., 104,* 9087–9092.

Beauchamp, M. S. (2005). Statistical criteria in fMRI studies of multisensory integration. *Neuroinformatics, 3,* 93–113.

Behrmann, M., Geng, J. J., & Shomstein, S. (2004). Parietal cortex and attention. *Current Opinion in Neurobiology, 14,* 212–217.

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

Blomert, L. (2011). The neural signature of orthographic-phonological binding in successful and failing reading development. *Neuroimage, 57,* 695–703.

Bolger, D. J., Perfetti, C. A., & Schneider, W. (2005). Cross-cultural effect on the brain revisited: Universal structures plus writing system variation. *Human Brain Mapping, 25,* 92–104.

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2003). Relation between brain activation and lexical performance. *Human Brain Mapping, 19,* 155–169.

Booth, J. R., Cho, S., Burman, D. D., & Bitan, T. (2007). Neural correlates of mapping from phonology to orthography in children performing an auditory spelling task. *Developmental Science, 10,* 441–451.

Bosse, M. L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition, 104,* 198–230.

Breitenstein, C., Jansen, A., Deppe, M., Foerster, A. F., Sommer, J., Wolbers, T., et al. (2005). Hippocampus activity differentiates good from poor learners of a novel lexicon. *Neuroimage, 25,* 958–968.

Carreiras, M., Seghier, M. L., Baquero, S., Estévez, A., Lozano, A., Devlin, J. T., et al. (2009). An anatomical signature for literacy. *Nature, 461,* 983–986.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., et al. (2000). The visual word form area—Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123,* 291–307.

Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. *Neuroimage, 40,* 353–366.

Cohen, L., Lehericy, 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,* 1054–1069.

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

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.

Cornelissen, K., Laine, M., Renvall, K., Saarinen, T., Martin, N., & Salmelin, R. (2004). Learning new names for new objects: Cortical effects as measured by magnetoencephalography. *Brain and Language, 89,* 617–622.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron, 56,* 384–398.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences, 15,* 254–262.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences, 9,* 335–341.

Devlin, J. T., Matthews, P. M., & Rushworth, M. F. S. (2003). Semantic processing in the left inferior prefrontal cortex: A combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience, 15,* 71–84.

Edmister, W. B., Talavage, T. M., Ledden, P. J., & Weisskoff, R. M. (1999). Improved auditory cortex imaging using clustered volume acquisitions. *Human Brain Mapping, 7,* 89–97.

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J. B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. *Human Brain Mapping, 2,* 165–189.

Frost, S. J., Mencl, W. E., Sandak, R., Moore, D. L., Rueckl, J. G., Katz, L., et al. (2005). A functional magnetic resonance imaging study of the tradeoff between semantics and phonology in reading aloud [Article]. *NeuroReport, 16,* 621–624.

Glezer, L. S., Jiang, X., & Riesenhuber, M. (2009). Evidence for highly selective neuronal tuning to whole words in the "visual word form area." *Neuron, 62,* 199–204.

Goodale, M. A., & Milner, A. D. (2010). Two visual streams: Interconnections do not imply duplication of function. *Cognitive Neuroscience, 1,* 65–68.

Gottlieb, J. (2007). From thought to action: The parietal cortex as a bridge between perception, action, and cognition. *Neuron, 53,* 9–16.

Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., et al. (1999). "Sparse" temporal sampling in auditory fMRI. *Human Brain Mapping, 7,* 213–223.

Hashimoto, R., & Sakai, K. L. (2004). Learning letters in adulthood: Direct visualization of cortical plasticity for forming a new link between orthography and phonology. *Neuron, 42,* 311–322.

Hayasaka, S., Phan, K. L., Liberzon, I., Worsley, K. J., & Nichols, T. E. (2004). Nonstationary cluster-size inference with random field and permutation methods. *Neuroimage, 22,* 676–687.

Husain, M., & Nachev, P. (2007). Space and the parietal cortex. *Trends in Cognitive Sciences, 11,* 30–36.

Josephs, O., & Henson, R. N. A. (1999). Event-related fMRI: Modelling, inference and optimisation. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 354,* 1215–1228.

Kherif, F., Josse, G., & Price, C. J. (2010). Automatic top–down processing explains common left occipito-temporal responses to visual words and objects. *Cerebral Cortex, 21,* 103–114.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience, 12,* 535–540.

Kronbichler, M., Bergmann, J., Hutzler, F., Staffen, W., Mair, A., Ladurner, G., et al. (2007). Taxi vs. taksi: On orthographic word recognition in the left ventral occipitotemporal cortex. *Journal of Cognitive Neuroscience, 19,* 1584–1594.

Kumaran, D., Summerfield, J. J., Hassabis, D., & Maguire, E. A. (2009). Tracking the emergence of conceptual knowledge during human decision making. *Neuron, 63,* 889–901.

Lee, H., Devlin, J. T., Shakeshaft, C., Stewart, L. H., Brennan, A., Glensman, J., et al. (2007). Anatomical traces of vocabulary acquisition in the adolescent brain. *The Journal of Neuroscience, 27,* 1184–1189.

Lobier, M., Zoubrinetzky, R., & Valdois, S. (2012). The visual attention span deficit in dyslexia is visual and not verbal. *Cortex, 48,* 768–773.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin and Review, 1,* 476–490.

Mano, Q. R., Humphries, C., Desai, R. H., Seidenberg, M. S., Osmon, D. C., Stengel, B. C., et al. (2013). The role of left occipitotemporal cortex in reading: Reconciling stimulus, task, and lexicality effects. *Cerebral Cortex, 23,* 988–1001.

Mechelli, A., Crinion, J. T., Long, S., Friston, K. J., Ralph, M. A. L., Patterson, K., et al. (2005). Dissociating reading processes on the basis of neuronal interactions. *Journal of Cognitive Neuroscience, 17,* 1753–1765.

Mechelli, A., Gorno-Tempini, M. L., & Price, C. J. (2003). Neuroimaging studies of word and pseudoword reading: Consistencies, inconsistencies, and limitations. *Journal of Cognitive Neuroscience, 15,* 260–271.

Mei, L., Xue, G., Chen, C., Xue, F., Zhang, M., & Dong, Q. (2010). The "visual word form area" is involved in successful memory encoding of both words and faces. *Neuroimage, 52,* 371–378.

Mei, L., Xue, G., Lu, Z. L., He, Q., Zhang, M., Xue, F., et al. (2013). Orthographic transparency modulates the functional asymmetry in the fusiform cortex: An artificial language training study. *Brain and Language, 125,* 165–172.

Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia, 46,* 774–785.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review, 114,* 273–315.

Plaut, D. C. (1999). A connectionist approach to word reading and acquired dyslexia: Extension to sequential processing. *Cognitive Science, 23,* 543–568.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *Neuroimage, 19,* 473–481.

Price, C. J., & Devlin, J. T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences, 15,* 246–253.

Price, C. J., Devlin, J. T., Moore, C. J., Morton, C., & Laird, A. R. (2005). Meta-analyses of object naming: Effect of baseline. *Human Brain Mapping, 25,* 70–82.

Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., et al. (2001). Neurobiological studies of reading and reading disability. *Journal of Communication Disorders, 34,* 479–492.

Rastle, K., & Coltheart, M. (1999). Serial and strategic effects in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 482–503.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science, 2(2 Suppl.),* 31–74.

Rayner, K., Raney, G. E., & Sereno, S. C. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1188–1200.

Reilhac, C., Peyrin, C., Démonet, J. F., & Valdois, S. (2013). Role of the superior parietal lobules in letter-identity processing within strings: fMRI evidence from skilled and dyslexic readers. *Neuropsychologia, 51,* 601–612.

Rorden, C., Karnath, H., & Bonilha, L. (2007). Improving lesion-symptom mapping. *Journal of Cognitive Neuroscience, 19,* 1081–1088.

Schurz, M., Sturm, D., Richlan, F., Kronbichler, M., Ladurner, G., & Wimmer, H. (2010). A dual-route perspective on brain activation in response to visual words: Evidence for a length by lexicality interaction in the visual word form area (VWFA). *Neuroimage, 49,* 2649–2661.

Seghier, M. L., Lazeyras, F., Pegna, A. J., Annoni, J. M., Zimine, I., Mayer, E., et al. (2004). Variability of fMRI activation during a phonological and semantic language task in healthy subjects. *Human Brain Mapping, 23,* 140–155.

Simon, O., Mangin, J. F., Cohen, L., Le Bihan, D., & Dehaene, S. (2002). Topographical layout of hand, eye, calculation, and language-related areas in the human parietal lobe. *Neuron, 33,* 475–487.

Sliwinska, M. W., Khadilkar, M., Campbell-Ratcliffe, J., Quevenco, F., & Devlin, J. T. (2012). Early and sustained supramarginal gyrus contributions to phonological processing. *Frontiers in Psychology, 3.* Article 161.

Starrfelt, R., Habekost, T., & Leff, A. P. (2009). Too little, too late: Reduced visual span and speed characterize pure alexia. *Cerebral Cortex, 19,* 2880–2890.

Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabrègue, R., Amadon, A., et al. (2011). Specialization for written words over objects in the visual cortex. *Neuroimage, 56,* 330–344.

Taylor, J. S. H., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin, 139,* 766–791.

Tsapkini, K., & Rapp, B. (2010). The orthography-specific functions of the left fusiform gyrus: Evidence of modality and category specificity. *Cortex, 46,* 185–205.

Turkeltaub, P. E., Flowers, D. L., Lyon, L. G., & Eden, G. F. (2008). Development of ventral stream representations for single letters. *Annals of the New York Academy of Science, 1145,* 13–29.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage, 15,* 273–289.

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,* 143–156.

Wilson, S. M., Brambati, S. M., Henry, R. G., Handwerker, D. A., Agosta, F., Miller, B. L., et al. (2009). The neural basis of surface dyslexia in semantic dementia. *Brain, 132,* 71–86.

Xue, G., Chen, C., Jin, Z., & Dong, Q. (2006). Language experience shapes fusiform activation when processing a logographic artificial language: An fMRI training study. *Neuroimage, 31,* 1315–1326.

Yoncheva, Y. N., Blau, V. C., Maurer, U., & McCandliss, B. D. (2010). Attentional focus during learning impacts N170 ERP responses to an artificial script. *Developmental Neuropsychology, 35,* 423–445.

Ziegler, J. C., Pech-Georgel, C., Dufau, S., & Grainger, J. (2010). Rapid processing of letters, digits and symbols: What purely visual-attentional deficit in developmental dyslexia? *Developmental Science, 13,* F8–F14.