

# Reversing the Standard Neural Signature of the Word–Nonword Distinction

William W. Graves<sup>1</sup>, Olga Boukrina<sup>2</sup>, Samantha R. Mattheiss<sup>1</sup>,  
Edward J. Alexander<sup>3</sup>, and Sylvain Baillet<sup>4</sup>

## Abstract

■ The distinction between letter strings that form words and those that look and sound plausible but are not meaningful is a basic one. Decades of functional neuroimaging experiments have used this distinction to isolate the neural basis of lexical (word level) semantics, associated with areas such as the middle temporal, angular, and posterior cingulate gyri that overlap the default mode network. In two fMRI experiments, a different set of findings emerged when word stimuli were used that were less familiar (measured by word frequency) than those typically used. Instead of activating default mode network areas often associated with semantic processing, words activated task-positive areas such as the inferior pFC and SMA, along with multifunctional ventral occipitotemporal cortices related to reading,

whereas nonwords activated default mode areas previously associated with semantics. Effective connectivity analyses of fMRI data on less familiar words showed activation driven by task-positive and multifunctional reading-related areas, whereas highly familiar words showed bottom-up activation flow from occipitotemporal cortex. These findings suggest that functional neuroimaging correlates of semantic processing are less stable than previously assumed, with factors such as word frequency influencing the balance between task-positive, reading-related, and default mode networks. More generally, this suggests that results of contrasts typically interpreted in terms of semantic content may be more influenced by factors related to task difficulty than is widely appreciated. ■

## INTRODUCTION

The ability to recognize a letter string as forming a word has been extensively investigated and is a critical step in the reading process. One major task used to investigate word recognition is lexical decision, where participants decide whether a string of letters is a valid word. Several decades of behavioral research using this task (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Rubenstein, Garfield, & Millikan, 1970) and over a decade of brain imaging research (McNorgan, Chabal, O'Young, Lukic, & Booth, 2015; Perani et al., 1999) have established some consistent findings. Behaviorally, when the nonword foils are well formed, such that they contain legal combinations of orthographic (letter combination) and phonological (sound combination) units, word recognition (1) takes longer to initiate and (2) is more influenced by semantic variables such as imageability than when the nonword foils do not contain legal constituents (Evans, Lambon Ralph, & Woollams, 2012). These findings, combined with the fact that by definition words are meaningful and nonwords are not, has led numerous functional brain imaging researchers to conclude that areas more active for words compared with well-formed nonwords are involved in processing word meanings or semantics (McNorgan

et al., 2015; Cattinelli, Borghese, Gallucci, & Paulesu, 2013; Taylor, Rastle, & Davis, 2013; Binder, Desai, Graves, & Conant, 2009).

Comparing neural responses to words and nonwords, however, is not entirely straightforward. Because lexical decisions to words are typically faster than to nonwords, this simple difference in time-on-task must be accounted for to draw conclusions about differences in semantic information processing, rather than differences in more domain-general process. Put another way, if one condition in a task is more difficult than the other, this difference could arise from many different sources and take many different forms. For example, differences between harder and easier conditions could result in differences in extent of visual attention between letter strings, recruitment of working memory resources, or level of effort. These differences, particularly when revealed by behavioral performance data, will necessarily manifest in the brain imaging data, possibly in the form of activation in areas that support domain-general processes. Some examples could include a response selection mechanism for mapping contents of working memory to a response (Rowe, Toni, Josephs, Frackowiak, & Passingham, 2000), a response inhibition system for preventing premature or prepotent responses from being made in error (Wager et al., 2005), and an error-monitoring system for adjusting response criteria (Ullsperger & von Cramon, 2004).

<sup>1</sup>Rutgers University, <sup>2</sup>Kessler Foundation, West Orange, NJ, <sup>3</sup>Tufts University, <sup>4</sup>McGill University

Under the right conditions, however, differences in behavioral performance data between conditions are clearly meaningful, as over a century of progress in experimental psychology has shown. Indeed, the choices of word frequency and imageability as factors to manipulate in the current experiments were motivated in large part by the extensive psycholinguistic literature on what performance differences that result from manipulating these factors tell us about cognitive processes involved in recognizing words. Effects of word frequency are among the most reliable and extensively studied in psycholinguistics. The basic finding is that lower-frequency words, compared with higher-frequency words, take longer to initiate responses to and are more error prone (Monsell, 1991). Effects of word frequency may arise throughout the lexical system (Monsell, Doyle, & Haggard, 1989) and are related to constructs such as word familiarity (Baayen, Feldman, & Schreuder, 2006; Colombo, Pasini, & Balota, 2006) and contextual diversity (Adelman, Brown, & Quesada, 2006). For the current experiments, we crossed levels of word frequency with levels of word imageability, with the aim of investigating the neurocognitive basis of lexical and semantic processing.

Although there is controversy regarding the exact role of semantics and imageability in reading aloud, semantic effects are clearly and consistently found for the lexical decision task used here (Balota et al., 2004). Values for imageability are obtained from humans rating the degree to which a word calls to mind an image. More highly imageable words generally have richer semantic representations (Paivio, 1991) and more semantic features (Plaut & Shallice, 1993), suggesting that imageability effects are straightforwardly interpreted in terms of semantics. Thus, we manipulated word imageability to reveal neural responses to semantic processing, as in several previous studies (Graves, Desai, Humphries, Seidenberg, & Binder, 2010; Hauk, Davis, Kherif, & Pulvermüller, 2008; Binder, 2007; Bedny & Thompson-Schill, 2006; Wise et al., 2000). We should then be able to compare this straightforwardly semantic manipulation with other contrasts related to semantics, such as words compared with nonwords.

A meta-analysis by Binder et al. (2009) examining the neural basis of semantic processing across 120 studies focused on studies contrasting semantically rich conditions to semantically poor conditions. For example, studies that compared neural responses to meaningful words against meaningless pronounceable nonwords (pseudo-words) or comparisons of high-imageability words against low-imageability words were included. They also used selection criteria to control for differences in “effort” or time-on-task between conditions by only including studies with either no differences in performance between conditions of interest or that had made an effort to account for those differences in some way, such as statistically modeling neural variance due to RTs. The consistent results across functional neuroimaging studies for semantically rich

conditions compared with semantically impoverished conditions revealed a set of brain areas that strikingly overlapped with what has come to be known as the “default mode” (DM) network (Buckner, Andrews-Hanna, & Schacter, 2008; Gusnard & Raichle, 2001). This prominently includes a largely bilateral set of regions such as the angular gyrus (AG), posterior cingulate (PC), precuneus, middle temporal gyrus (MTG), anterior temporal lobes (ATL), and dorso-medial pFC. It contrasts with and is spatially nonoverlapping with the “task-positive” network (Fox et al., 2005). This latter set of regions has been associated with a diverse array of resource-demanding functions, leading it to also be termed the “multiple-demand” (MD) network (Duncan, 2010). Like the DM network, the MD network is also largely bilateral and includes the inferior frontal junction (IFJ; centered at the junction of the inferior frontal and precentral sulci), intraparietal sulcus (IPS), and SMA. There was also a small third set of regions that showed some spatial overlap between these two networks, including the supramarginal gyrus (SMG) and ventral occipitotemporal cortex (vOT; Binder et al., 2009). Because words primarily differ from well-matched nonwords in that words are meaningful, areas more activated for words that largely correspond to the DM network have been interpreted as carrying out semantic processing (McNorgan et al., 2015; Cattinelli et al., 2013; Taylor et al., 2013; Binder & Desai, 2011; Binder et al., 1999, 2009).

The original goal of the first experiment reported here was to replicate in a single fMRI study the overall pattern of results from the basic semantic contrast in the Binder et al. (2009) meta-analysis. These data, however, showed a surprising pattern. We found activation in the MD network for words compared with nonwords and activation in the putative semantic network for nonwords compared with words. This was surprising considering that nonwords do not have meaning. To better understand the source of this finding, we turned again to the Binder et al. (2009) meta-analysis, which included numerous studies using the lexical decision task. A representative example is an earlier study by Binder, Westbury, McKiernan, Possing, and Medler (2005). As in the current study, their participants performed lexical decision on stimuli that were either words or pronounceable nonwords. The nonwords were well matched to the words in terms of phonological and orthographic measures. The words also varied in imageability and concreteness—semantic factors that were treated as essentially the same because of their high correlation of .94 based on ratings from the MRC Psycholinguistic Database (Coltheart, 1981). Words were matched for frequency across levels of imageability. In comparing the Binder et al. (2005) stimulus words with ours, we found that, although they did not differ in imageability, the median word frequency of their words was 16.8 occurrences per million (midpoint of the lowest quartile: 6.6), whereas ours were 12.4 occurrences per million (midpoint of the lowest quartile: 2.6), both according to the CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995). This

difference suggested word frequency as a possible source of the diverging patterns for the lexical contrast.

A follow-up fMRI experiment was then conducted to both check for replication of the unexpected results and further explore their source. Specifically, we tested the possibility that individual word characteristics (in this case, word frequency) can influence the neural correlates of a distinction as basic as lexicality (word/nonword status) to such an extent as to reverse its typical neural signature. A possible mechanism for this reversal is hypothesized to be the differential engagement of MD and DM networks because of asymmetries in difficulty between the words and nonwords.

## METHODS

### Experiment 1

#### *Stimulus Material*

A total of 312 words and 312 pseudowords were selected for lexical decision. Words were divided into high and low levels of frequency and imageability in a completely crossed  $2 \times 2$  factorial design, producing four unique conditions (high frequency, high imageability; high frequency, low imageability; low frequency, high imageability; low frequency, low imageability) with 78 words per condition. Log-transformed per-million values for word frequency were obtained from the CELEX lexical database (Baayen et al., 1995) and ranged from a minimum of 0.004 (“kelp”) to a maximum of 3.083 (“first”), with a mean of 1.155. Words were selected to have a clear bimodal distribution. Words with frequency greater than 1.2 were considered high frequency, and this category had a mode of 1.91 (in terms of occurrences per million, the median was 75.2). Words with frequency less than 0.7 were considered low frequency, and this category had a mode of 0.56 (and a median of 2.6 occurrences per million). Imageability values were derived from ratings studies in which participants were asked to rate words in terms of the degree to which they bring to mind an image (Clark & Paivio, 2004; Cortese & Fugett, 2004; Bird, Franklin, & Howard, 2001; Gilhooly & Logie, 1980; Toglia & Battig, 1978; Paivio, Yuille, & Madigan, 1968). Imageability ranged from a minimum of 1.8 (e.g., “guile”) to a maximum of 6.6 (“beach”), with a mean of 4.4. The distribution was again bimodal, with the low imageability category defined as less than or equal to 4 (mode: 2.9) and high imageability as greater than 4.7 (mode: 5.8). Words did not reliably differ across conditions in terms of number of letters, bigram frequency, trigram frequency, orthographic neighborhood size, or spelling–sound consistency. As expected, words categorized as high or low frequency showed reliable differences in word frequency values. Likewise, high- and low-imageability words differed reliably in terms of imageability. To help ensure surface similarity to words, pseudowords were generated to contain only trigram sequences valid in English. They also did not reliably differ from words

in terms of number of letters, bigram frequency, or trigram frequency.

#### *Participants*

A total of 20 participants (13 women), with mean age of 25.3 years, with mean years of education of 16.6, all right-handed ( $>70$  on the Oldfield handedness questionnaire), and with no history of neurological, psychiatric, or learning impairment diagnosis, underwent fMRI scanning. All had English as a first language. They gave written informed consent, and the study protocol was approved by the institutional review board of the Medical College of Wisconsin.

#### *Task and Imaging*

Participants were instructed in the lexical decision task in the scanner by being told they would see letter strings. Their task was to decide as quickly and accurately as possible whether or not each letter string was a valid English word. Each stimulus was displayed for 400 msec before being replaced with a single fixation cross. Fixation served as an implicit baseline. The scanning session was split into six runs, each containing 52 words and 52 pseudowords. These were intermixed with 52 fixation trials in a fully randomized rapid event-related design, with a mean inter-trial interval of 3.1 sec (*SD*: 2.0). E-Prime (Psychology Software Tools, Inc., Pittsburgh, PA; www.pstnet.com/eprime) was used for stimulus presentation and response recording.

MRI data were acquired using a 3.0-T Excite system (GE Healthcare Systems, Milwaukee, WI) with an eight-channel array head coil. To ensure high signal quality across the whole head, we obtained T1-weighted anatomical images in both the axial (180 slices,  $0.938 \times 0.938 \times 1.000$  mm) and sagittal (180 slices,  $1.000 \times 0.938 \times 0.938$  mm) planes using a spoiled gradient-echo sequence. Task-based functional scans were acquired using a gradient-echo echoplanar sequence with the following parameters: echo time (TE) = 25 msec, repetition time (TR) = 2 sec, field of view = 208 mm, matrix =  $64 \times 64$  pixel, in-plane voxel dimensions =  $3.25 \times 3.25$  mm, and slice thickness = 3.3 mm with no gap. Thirty-three interleaved axial slices were acquired, and each of the six functional runs consisted of 168 whole-brain image volumes. Resting-state functional scans were also acquired with a gradient-echo echoplanar sequence, but with the following parameters: TE = 25 msec, TR = 3 sec, field of view = 240 mm, matrix =  $128 \times 128$  pixel, in-plane voxel dimensions =  $1.875 \times 1.875$  mm, and slice thickness = 2.5 mm with no gap. Forty-one interleaved axial slices were acquired in a single run of 140 image volumes.

Image analyses were performed using AFNI (afni.nimh.nih.gov/afni; Cox, 1996). The one exception was the early step of B-field unwarping, in which the time series data were processed using the FSL (Smith et al., 2004) program

FUGUE, based on a field map acquired in the same dimensions as the task-based functional data. Subsequently, for each participant, the first six images in the time series were discarded to avoid saturation effects. The remaining images were spatially coregistered (motion-corrected; Cox & Jesmanowicz, 1999), and the resulting motion parameters were saved for use as noise covariates. Voxelwise multiple linear regression was performed using the AFNI program 3dDeconvolve. This included the following covariables of no interest: a third-order polynomial to model low-frequency drift, the six previously calculated motion parameters, and a term for signal in the ventricles used to model noise. Covariables of interest were modeled as impulse responses convolved with a gamma variate approximation of the hemodynamic response. They consisted of the following: (1–4) an indicator variable for each word type responded to correctly in the  $2 \times 2$  design (high frequency, high imageability; low frequency, low imageability; high frequency, low imageability; low frequency, high imageability), (5) an indicator variable for each nonword responded to correctly, (6) an indicator variable for each erroneous response, and (7) RT values for each correct trial. Because RT for words and pseudo-words would be highly correlated with the indicator variables for correct trials, the RT values for each trial were mean-centered by subtracting the overall mean RT of each participant's correct responses and dividing the result by the standard deviation for all correct trials (essentially z-scoring the RT values). The relevant contrast of words–nonwords was performed by first combining across the four word types and then performing the contrast. Similarly, contrasts between levels of word frequency were performed by collapsing across levels of imageability, and contrasts between levels of imageability were performed by collapsing across levels of word frequency. We also tested for multiplicative interaction of the effects of word frequency (high–low) and lexicality (word–nonword).

The resulting contrast coefficient maps for each participant were linearly resampled into Talairach space (Talairach & Tournoux, 1988) with a voxel size of  $1 \text{ mm}^3$  and spatially smoothed with a 6-mm FWHM Gaussian kernel. These smoothed coefficient maps were then passed to a random effects analysis comparing the coefficient values to a null hypothesis of zero across participants. The resulting group activation maps were thresholded at a voxelwise  $p < .01$ , uncorrected. A cluster extent threshold was then calculated using the AFNI program 3dClustSim to perform Monte Carlo simulations estimating the chance probability of spatially contiguous voxels passing this threshold. Clusters smaller than  $812 \mu\text{l}$  were removed, resulting in a whole-brain corrected threshold of  $p < .05$ .

A functional connectivity analysis of the resting state (fixation baseline) was also performed using a seed in the PC to verify that participants showed the typical pattern of resting-state functional connectivity and to provide ROIs for the effective connectivity analyses derived from independent data on this group of participants to

avoid overfitting from multiple analyses of the same data. Methods for this analysis and the results are provided as supplemental material.

## Experiment 2

### Stimuli and Task

Stimuli were identical to the previous experiments, with the exception that all high-frequency words were presented for lexical decision in the first half of the experiment, followed by all low-frequency words. This was done to isolate effects of word frequency context and provide an optimal design for the planned effective connectivity analyses.

### Participants

The 11 participants (eight women) had a mean age of 21.6 years, with 14.2 mean years of education. This new set of participants had attained fewer years of education than the previous sample ( $t = 4.1, p < .001$ ), presumably as a consequence of their being younger ( $t = 2.9, p < .01$ ). All participants met inclusion criteria as described above and gave written informed consent for the study as approved by the Rutgers University institutional review board.

### Image Acquisition and Analysis

The MRI data for this experiment were acquired in the Rutgers University Brain Imaging Center, using a 3T MAGNETOM Trio (Siemens Healthcare, Flanders, NJ) with a 12-channel array head RF receive coil. High-resolution, T1-weighted anatomical reference images were acquired as a set of 176 contiguous sagittal slices ( $1 \text{ mm}^3$  voxels) using a magnetization prepared rapid gradient echo sequence for whole-head coverage. Functional scans were acquired using a gradient-echo EPI sequence with the following parameters: TE = 25 msec, TR = 2 sec, field of view = 192 mm, matrix =  $64 \times 64$  pixel, and  $3 \text{ mm}^3$  voxels. Thirty-five interleaved axial slices (no gap) were acquired for whole-brain coverage, and each of the six functional runs consisted of 168 image volumes.

Analysis of the MRI data were as described for Experiment 1 above, with the following exceptions: (1) nonwords appearing in the first half of the experiment as foils for high-frequency words were modeled separately from those appearing in the last half of the experiment as foils for low-frequency words (note that the order of nonwords was fully rerandomized across the entire experiment for each participant), (2) because of the slightly smaller voxel size in this data set compared with Experiment 1, the mapwise cluster correction to  $p < .05$  resulted in excluding clusters smaller than  $805 \mu\text{l}$ , and (3) effective connectivity analyses were performed on the minimally processed fMRI task data using Independent Multiple-sample and Greedy Equivalence Search (IMAGES; Ramsey

et al., 2010), as implemented in the Tetrad (version 4.3.10-7) software environment. In an approach similar to that described previously (Boukrina & Graves, 2013), we provided ROIs as priors for the Bayesian framework used in IMAges. The ROIs shown in Figure 4 were generated from the resting-state functional connectivity results from Experiment 1 (Figure S1), but kept in volume space for 3-D functional image analysis. Unlike dynamic causal modeling (Friston, Harrison, & Penny, 2003), where the exact connections must be specified beforehand, IMAges is a search algorithm that discovers the maximally likely set of connections. This is made computationally tractable by only considering as separate those connections that fall into different Markov equivalence classes, estimated using conditional independence relations, and do not form cyclic loops. Once the significant connections are found, directionality is determined using the LOFS algorithm (stands for LiNG Orientation, Fixed Structure, where LiNG stands for Linear, Non-Gaussian). Its general approach is to test, using linear models, directionality of flow among groups of ROIs, with the most likely direction corresponding to the model whose residual error is least Gaussian (Ramsey, Hanson, & Glymour, 2011).

## RESULTS

### Experiment 1

#### *Behavioral Results*

Means for all conditions are shown in Table 1. Analyses with RT as the dependent variable showed lexical decisions to words being faster than nonwords (itemwise  $t(622) = 16.7, p < .0001$ ). Effects of word frequency and imageability were also significant, with high-frequency and high-imageability words eliciting faster responses than low-frequency and low-imageability words. RT showed reliable effects of word frequency ( $F(1, 19) = 134.5, p < .0001$ ) and imageability ( $F(1, 19) = 111.0, p < .0001$ ). There was also a reliable interaction between word frequency and imageability ( $F(1, 19) = 43.1, p < .0001$ ), such that effects of imageability were greater for low- compared with high-frequency words and effects of word frequency were greater for low- compared with high-imageability words.

Analyses with accuracy (1 – error rate) as the dependent variable showed lexical decisions to nonwords being more accurate than those to words ( $t(18) = 15.9, p < .0001$ ). The combination of faster and less accurate responses to words compared with nonwords suggests the possibility of a speed–accuracy tradeoff. Word frequency and imageability significantly influenced error rates, with high-frequency and high-imageability words eliciting more accurate responses than low-frequency and low-imageability words, resulting in reliable effects of word frequency ( $F(1, 19) = 80.1, p < .0001$ ) and image-

ability ( $F(1, 19) = 124.3, p < .0001$ ). Unlike the word–nonword contrast, this pattern is consistent with the RT results, suggesting that speed–accuracy tradeoff was not an issue for words. Also consistent with RT, there was a significant interaction ( $F(1, 19) = 90.3, p < .0001$ ), with greater effects of imageability for low- compared with high-frequency words and greater effects of word frequency for low- compared with high-imageability words.

Because of the apparent speed–accuracy trade-offs between the word and nonword conditions and to get an overall sense of which conditions are the “hardest,” we calculated the inverse efficiency score (IES; Bruyer & Brysbaert, 2011; Townsend & Ashby, 1978). The lexicality effect was significant, with words showing a higher IES than pseudowords ( $t(622) = 2.611, p < .01$ ). ANOVA with Items (words) as the random variable showed main effects of Word frequency ( $F(308) = 14.6, p < .001$ ) and Imageability ( $F(308) = 14.3, p < .001$ ) in the expected direction and a significant interaction ( $F(308) = 10.3, p < .001$ ).

#### *Imaging Results*

A direct contrast of word (warm colors in Figure 1A) compared with nonword (cool colors in Figure 1A) trials showed primarily left-lateralized activation for words in the IFJ, bilateral occipital cortices and vOT (with a larger spatial extent on the left), left SMA, and left BG (coordinates in Table S1). Activation for nonwords compared with words was primarily in bilateral AG, dorsal and medial prefrontal, ATL, PC, cuneus, and precuneus.

To examine the extent to which the areas that activated for words compared with nonwords corresponded to known networks previously identified in terms of functional connectivity (Fox et al., 2005), analysis of resting-state data was performed on the same 20 participants using a seed placed in the PC, as described in Methods. Areas with a resting-state time series significantly positively correlated with PC are shown in cool colors in Figure S1, and areas with a time series significantly negatively correlated with PC are shown in warm colors (for coordinates, see Table S2). This color scheme was chosen for ease of comparison with the word–nonword contrast in Figure 1A. Visual comparison of resting-state data (Figure S1) with the lexicality contrast (Figure 1A) shows areas activated for words corresponding to areas anticorrelated with the PC seed and areas activated for nonwords corresponding to the PC-correlated resting state/DM network.

Degree of overlap between activation for nonwords in the lexicality contrast and the putative semantic network from Binder et al. (2009) was examined by mapping both results onto the same brain in atlas space (Figure 1B). Activations shown in cool colors in Figure 1A are shown in red in Figure 1B, and its overlap with the putative semantic network is in yellow. These overlaps occur most prominently in the AG, dorsal and ventromedial pFC, PC, precuneus, and ATL. Areas of the putative semantic system were generally more extensive than those activated for

**Table 1.** Means of Performance Measures for Each Condition in Both Experiments

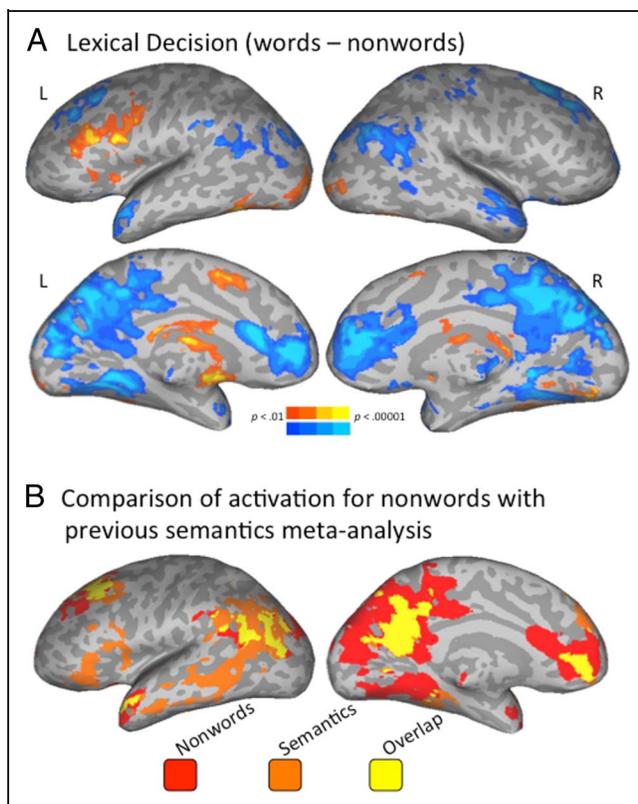
| <i>Experiment 1, RT (msec)</i>    |      |                       |            |             | <i>Experiment 2, RT (msec)</i>    |      |                       |            |             |
|-----------------------------------|------|-----------------------|------------|-------------|-----------------------------------|------|-----------------------|------------|-------------|
|                                   |      | <i>Word Frequency</i> |            |             |                                   |      | <i>Word Frequency</i> |            |             |
|                                   |      | <i>High</i>           | <i>Low</i> | <i>Mean</i> |                                   |      | <i>High</i>           | <i>Low</i> | <i>Mean</i> |
| Imageability                      | High | 695.7                 | 754.0      | 724.8       | Imageability                      | High | 1210.0                | 1342.5     | 1276.3      |
|                                   | Low  | 727.8                 | 859.2      | 793.5       |                                   | Low  | 1231.7                | 1444.1     | 1337.9      |
|                                   | Mean | 711.8                 | 806.6      |             |                                   | Mean | 1220.9                | 1393.3     |             |
| Overall mean for words            |      |                       |            | 748.5       | Overall mean for words            |      |                       |            | 1287.2      |
| <i>Experiment 1, Accuracy (%)</i> |      |                       |            |             | <i>Experiment 2, Accuracy (%)</i> |      |                       |            |             |
|                                   |      | <i>Word Frequency</i> |            |             |                                   |      | <i>Word Frequency</i> |            |             |
|                                   |      | <i>High</i>           | <i>Low</i> | <i>Mean</i> |                                   |      | <i>High</i>           | <i>Low</i> | <i>Mean</i> |
| Imageability                      | High | 96.3                  | 91.5       | 93.9        | Imageability                      | High | 93.7                  | 82.4       | 88.1        |
|                                   | Low  | 91.4                  | 64.6       | 78.0        |                                   | Low  | 89.0                  | 44.6       | 66.8        |
|                                   | Mean | 93.8                  | 78.0       |             |                                   | Mean | 91.4                  | 63.5       |             |
| Overall mean for words            |      |                       |            | 85.9        | Overall mean for words            |      |                       |            | 77.4        |
| <i>Experiment 1, IES</i>          |      |                       |            |             | <i>Experiment 2, IES</i>          |      |                       |            |             |
|                                   |      | <i>Word Frequency</i> |            |             |                                   |      | <i>Word Frequency</i> |            |             |
|                                   |      | <i>High</i>           | <i>Low</i> | <i>Mean</i> |                                   |      | <i>High</i>           | <i>Low</i> | <i>Mean</i> |
| Imageability                      | High | 724.0                 | 834.4      | 779.2       | Imageability                      | High | 1301.9                | 1810.4     | 1556.1      |
|                                   | Low  | 828.3                 | 2116.0     | 1472.1      |                                   | Low  | 1520.6                | 5302.7     | 3361.9      |
|                                   | Mean | 776.2                 | 1475.2     |             |                                   | Mean | 1411.2                | 3510.6     |             |
| Overall mean for words            |      |                       |            | 1125.7      | Overall mean for words            |      |                       |            | 2447.3      |
| <i>Experiment 1, Pseudowords</i>  |      |                       |            |             | <i>Experiment 2, Pseudowords</i>  |      |                       |            |             |
| Mean RT (msec)                    |      |                       |            | 811.0       | Mean RT (msec)                    |      |                       |            | 1350.2      |
| Mean accuracy (%)                 |      |                       |            | 94.4        | Mean accuracy (%)                 |      |                       |            | 91.4        |
| Mean IES                          |      |                       |            | 872.0       | Mean IES                          |      |                       |            | 1506.1      |

Main effects of lexicality, word frequency, and imageability were all significant at  $p < .001$ , as described in the Results section. The interaction of word frequency and imageability was also significant at  $p < .001$  by all three performance measures in both experiments. The sole exception was the lexicality effect for Experiment 1 as measured by IES, which was significant at  $p < .01$ .

nonwords on the lateral surface, whereas the opposite was true (a larger spatial extent of activation for nonwords than the semantic system) on the medial surface. Areas associated with semantics but not activated for nonwords extended along the MTG and included the orbital and triangular parts of the inferior frontal gyrus (IFG).

Analysis of the behavioral data in terms of IES suggested that the word condition was more difficult than the nonword condition when taking into account both RT and accuracy. Similarly, high-frequency words were responded to more accurately and quickly than low-

frequency words, suggesting that lexical decision to high-frequency words was less difficult than for low-frequency words. If the relative balance between engagement of task-positive and DM networks is being determined, at least in part, by relative difficulty of the stimuli, then areas activated for high- compared with low-frequency words should show a similar pattern to that seen for nonwords compared with words. Indeed, this is largely the pattern we observed (cf. Figures 1A and 2A). In each case, the more difficult or “harder” condition (words in the lexical contrast, low-frequency words in the frequency contrast) activated



**Figure 1.** Correspondence between task-based and resting-state fMRI data. (A) Direct contrast between words (warm colors) and nonwords (cool colors). (B) Overlapping (yellow) and separate maps of significant findings from the semantics meta-analysis (orange; Binder et al., 2009) and activations for nonwords from Experiment 1 (red). Left hemisphere only shown here.

left IFJ, whereas the less difficult or “easier” condition (nonwords in the lexical contrast, high-frequency words in the frequency contrast) activated bilateral ATL, PC, dorsal and medial prefrontal cortices, precuneus, and right AG. There were also some reading-related areas where activation did not overlap across the lexicality and frequency analyses. In the left inferior parietal lobule, the AG activated more for words than nonwords, whereas the SMG activated more for high- than low-frequency words. Additionally, the left vOT activated for words compared with nonwords, whereas there were no reliable differences in this area for the word frequency analysis. Overall, there was a great deal of overlap for areas activated when contrasting hard and easy conditions, whether that contrast was between words and nonwords or between high- and low-frequency words.

We also contrasted activations for high-imageability words with those of low-imageability words (Figure 2B). High- compared with low-imageability words showed activation in several areas including right superior frontal gyrus, left parahippocampal gyrus, and bilateral precuneus and PC cortices. Low- compared with high-imageability words showed activation only in left IFG.

A final question was whether the pattern seen in the lexicality contrast would differ depending on whether

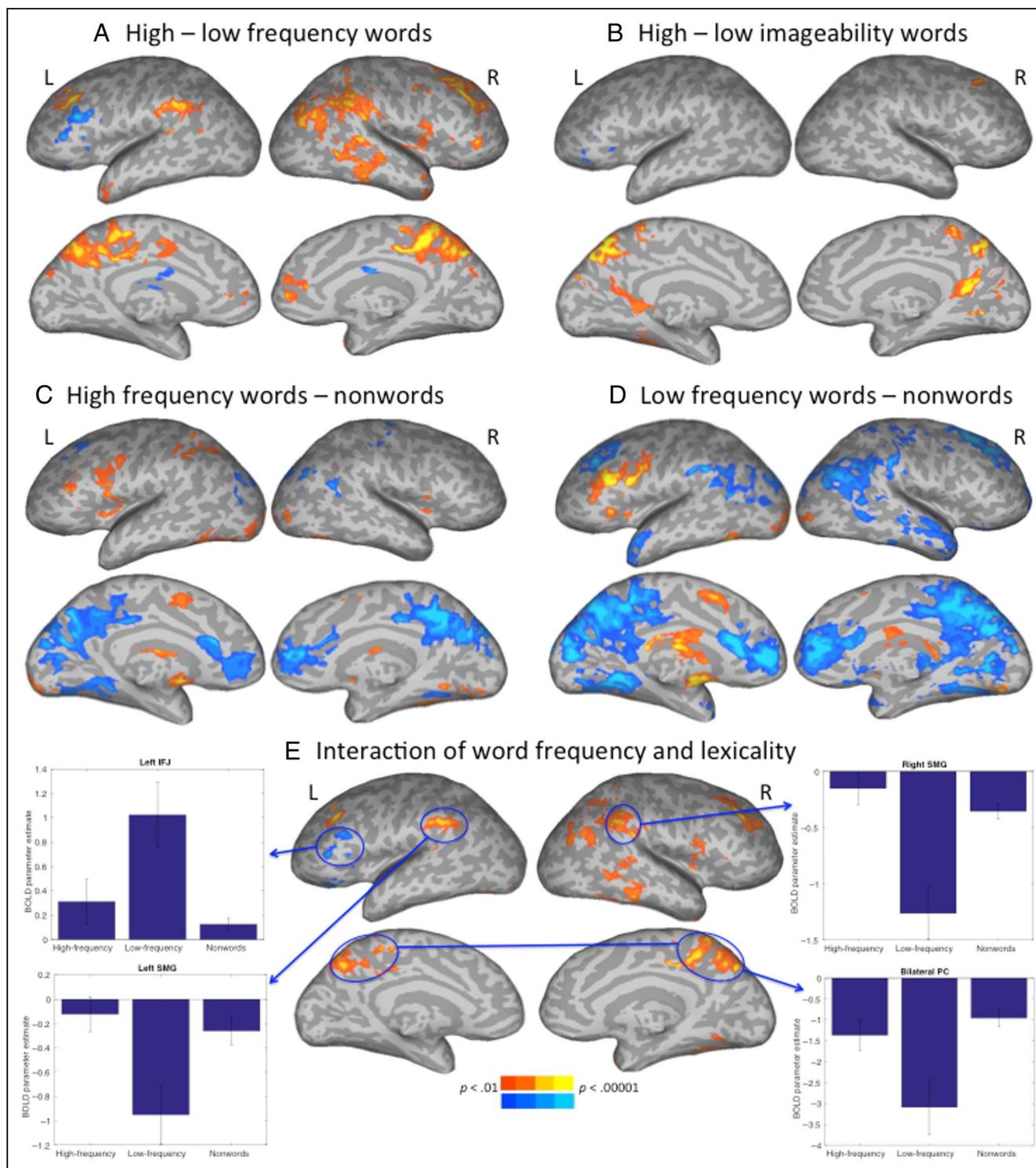
the nonwords were being contrasted with either high- or low-frequency words. Qualitatively, although the high-frequency words minus nonwords contrast (Figure 2C) yielded results similar to those from the low-frequency words minus nonwords contrast (Figure 2D), the low-frequency words–nonwords contrast looked more similar to the lexicality contrast (Figure 1A). We checked for statistical reliability of this pattern by testing for areas showing an interaction between the word frequency contrast and the lexicality contrast (Figure 2E). Graphs for four representative regions are shown in Figure 2E, where activation levels for high-frequency words, low-frequency words, and nonwords are graphed relative to fixation baseline. The left IFJ, which showed activation for words compared with nonwords, here shows activation for low-frequency words compared with either high-frequency words or nonwords. The other regions in warm colors all show a different pattern. Bilateral SMG is adjacent to the bilateral AG regions that showed more activation for nonwords than words, and the interaction analysis shows that rather than nonwords being more activated than words relative to baseline, they are instead less deactivated than low-frequency words compared with baseline. High-frequency words showed a very different pattern from low-frequency words, instead closely following the pattern seen for nonwords.

## Experiment 2

To test for replication of the results from Experiment 1 and to further explore the possibility that inclusion of relatively unfamiliar, low-frequency words may have driven the unexpected result in Figure 1A, a new sample of participants ( $N = 11$ ) was scanned in a different scanner at a different institution. Word stimuli were blocked by frequency, with all high-frequency words appearing in the first half of the session and all low-frequency words appearing in the last half. This was done to determine the influence of word frequency list context and to optimize the design for effective connectivity analysis, as described in the Methods section.

### Behavioral Results

Means for each condition are given in Table 1. This new set of participants showed an RT pattern similar to the previous participants in that words were responded to more quickly than nonwords (itemwise  $t(622) = 11.4$ ,  $p < .0001$ ). The factorial manipulation of word properties also yielded a pattern of results similar to those from the previous set of participants. High-frequency words were responded to more quickly than low-frequency words ( $F(1, 10) = 153.5$ ,  $p < .0001$ ), and high-imageability words were responded to more quickly than low-imageability words ( $F(1, 10) = 23.4$ ,  $p < .001$ ). The interaction of word frequency and imageability also followed the same pattern as the previous participants, with the effect of imageability being greater for low- compared with high-frequency



**Figure 2.** Experiment 1 contrasts for high- minus low-frequency words (A) and high- minus low imageability words (B). Separate contrasts are also shown for high-frequency words minus nonwords (C) and low-frequency words minus nonwords (D). (E) Interaction of word block type (high- or low-frequency words) with the lexicality contrast, with parameter estimates graphed for four representative ROIs.

words and the effect of word frequency being greater for low- compared with high-imageability words ( $F(1, 10) = 19.9, p < .001$ ).

Accuracy rates were also similar to the previous group, with less accurate responses to words than nonwords

( $t(10) = 16.3, p < .0001$ ). Factorial manipulations for words revealed a higher accuracy for high-frequency compared with low-frequency words ( $F(1, 10) = 100.9, p < .0001$ ) and greater accuracy for high-imageability compared with low-imageability words ( $F(1, 10) = 187.9, p < .0001$ ). There

was also a significant interaction between word frequency and imageability ( $F(1, 10) = 102.7, p < .0001$ ), following the pattern described for the RT data.

As in Experiment 1, we calculated the IES as a way of combining RT and accuracy to determine conditions under which participants found lexical decision most difficult. Analysis by items revealed the IES for words to be significantly higher than for pseudowords ( $t(618) = 5.3, p < .001$ ). ANOVA was performed with median replacement of four words in the low-frequency, low-imageability condition, because they were not responded to correctly by any of the 11 participants. As with RT and Accuracy, there were significant main effects in the expected direction, with performance on high-frequency words being better than low-frequency words ( $F(1, 308) = 44.1, p < .001$ ) and performance on high-imageability words being better than low-imageability words ( $F(1, 308) = 32.4, p < .001$ ). The interaction of Word frequency and Imageability was also significant in the expected direction ( $F(1, 308) = 24.9, p < .001$ ).

Overall, the main effects and interactions are consistent across both participant groups. Although error rates were higher overall for words compared with nonwords, this pattern appears to have been driven by word frequency. When considered separately, high-frequency words had error rates comparable to nonwords, whereas error rates to low-frequency words were considerably higher.

### Imaging Results

In spite of this experiment being conducted on a different scanner, in a different region of the United States, with different participants compared with Experiment 1, they produced broadly similar results, particularly for the lexicality contrast. Words activated bilateral vOT and lateral occipital regions, along with subcortical regions and SMA (warm colors in Figure 3A). Nonwords activated bilateral ATL, AG, medial pFC, PC, precuneus, and parahippocampal gyrus (cool colors in Figure 3A, full coordinates in Table S3).

Contrasts of high- compared with low-frequency words revealed activation for high-frequency words in bilateral superior temporal gyrus, MTG, SMG, SMA, medial occipital, mid-cingulate, and dorsolateral prefrontal cortices, and IPS (Figure 3B). No areas were activated for low-compared with high-frequency words. Contrasts of high-compared with low-imageability words revealed activation for high-imageability words in areas including bilateral SMG and dorsolateral prefrontal cortex, and right STG and MTG (Figure S2). No areas were activated for low-compared with high-imageability words.

Blocking the word stimuli by frequency allowed us to test the hypothesis that inclusion of very low frequency, unfamiliar words was driving the engagement of the dorsal attention network or task-positive network for words. Contrasting high-frequency words to nonwords

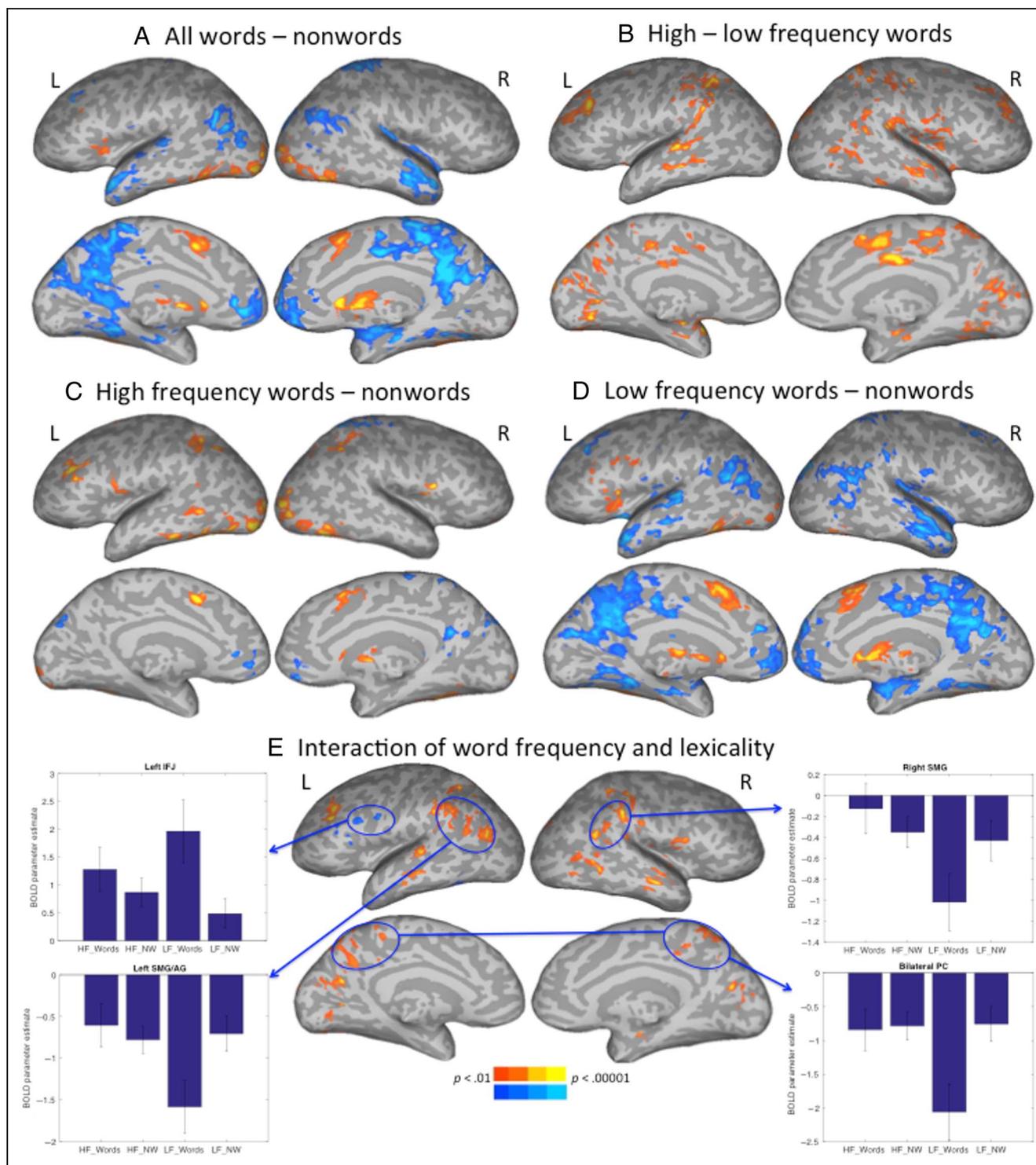
revealed activation for high-frequency words in left middle frontal gyrus, opercular IFG, and MTG, along with bilateral activations in vOT and IPS. Activations for nonwords included bilateral ventromedial prefrontal cortices and right PC (Figure 3C). The contrast of low-frequency words with nonwords revealed a strikingly different pattern (Figure 3D). Low-frequency words activated several areas including left IFG and anterior insula, vOT, bilateral subcortical structures, and SMA. Activations for nonwords included bilateral AG, ATL, dorsal and medial prefrontal cortices, parahippocampal cortex, PC, and precuneus.

To determine areas where these lexicality results are reliably different across levels of word frequency, we tested for areas showing an interaction between levels of word frequency and the word–nonword (lexicality) contrast. Activation for each condition is graphed relative to the fixation baseline in Figure 3E. Areas showing an interaction included left-sided IFJ, an activation spanning SMG and AG, IPS, right SMG, and bilateral precuneus. The pattern of activation is broadly similar to that seen in the fully randomized design of Experiment 1 (Figure 2E), with the exception of an interaction effect in the IPS that was not present in Experiment 1 (cf. Figures 2E and 3E). The pattern of activation in IPS is graphed for both experiments in Figure S3. Unlike the other graphs, this shows greater activation for high-frequency words than for low-frequency words and nonwords. However, the prevailing pattern across the other areas where word frequency and lexicality interact is that activation for low-frequency words differs from nonwords more than does activation of high-frequency words compared with nonwords.

The final analysis reported here examined the role of six major brain areas that are both highlighted by the lexicality contrasts reported here and are key components of the task-positive network (IFJ, IPS), DM network (AG, PC), or possibly both (vOT, SMG). We performed a Bayesian effective connectivity analysis using these ROIs. They were generated based on the resting state data from Experiment 1, as described in the Methods and shown in Figure 4A. In general, responses to high-frequency words were dominated by bottom–up activation from vOT to areas previously associated with phonological (SMG), semantic (AG), and attentional (IPS) processing (Figure 4B). Responses to low-frequency words (Figure 4C), on the other hand, were characterized by relatively more top–down activation from IFJ, PC, and IPS to vOT.

## DISCUSSION

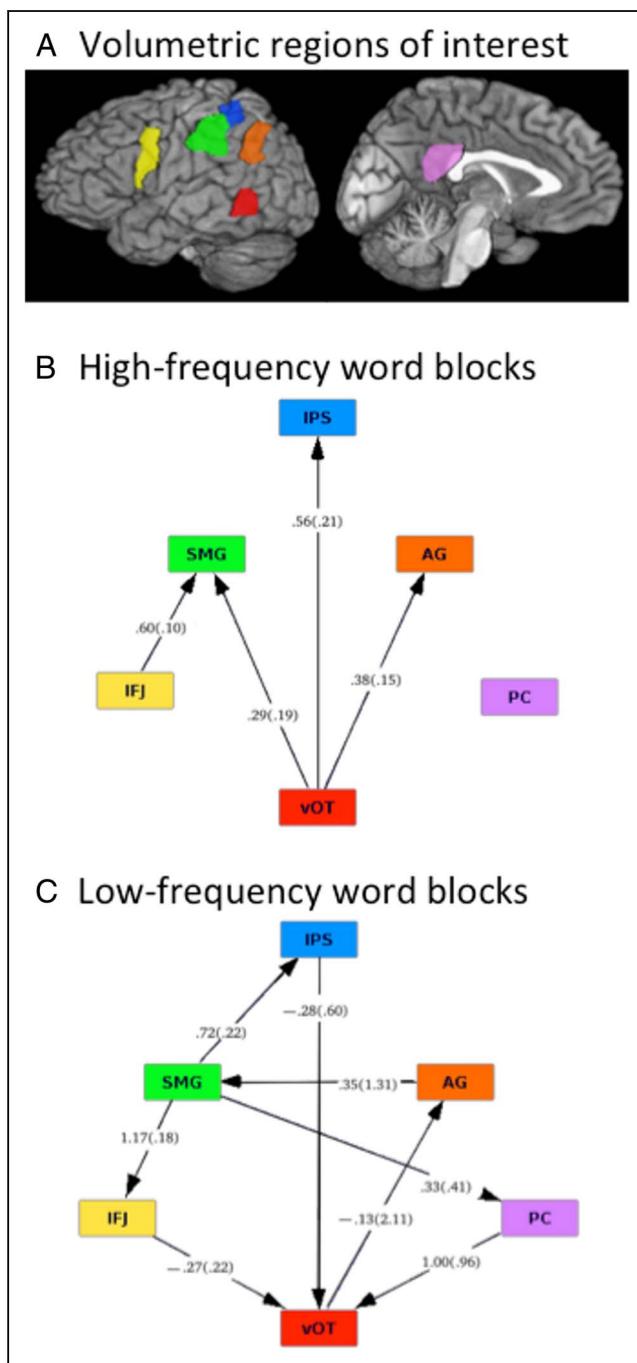
An initial fMRI experiment showed an unexpected result when contrasting words with nonwords, in that nonwords activated regions associated in numerous other studies with semantic processing. Words activated a combination of areas broadly corresponding to what has been described as the dorsal attention, task-positive, or MD network (including IFJ and SMA), along with vOT, which has been associated both with the dorsal attention



**Figure 3.** Experiment 2 lexicality (word-nonword) contrasts for all stimuli together (A), high- minus low-frequency words (B), high-frequency words and the nonword background against which they appeared (C), and low-frequency words with their nonword background (D). (E) Interaction of word block type (high- or low-frequency words) with the lexicality contrast, with parameter estimates graphed for four representative ROIs.

network (Vogel, Miezin, Petersen, & Schlaggar, 2012) and reading-specific processes (Dehaene et al., 2010). The set of areas activated for nonwords, on the other hand, broadly corresponded to the DM network (Fox et al., 2005). This pattern was confirmed in a follow-up experiment with dif-

ferent participants in a different scanner but with the same stimuli. To test the hypothesis that the reversal of the expected pattern was driven by effects of familiarity or word frequency, high- and low-frequency words were presented in separate halves of the follow-up fMRI experiment. The



**Figure 4.** Effective connectivity analysis results for ROIs (A) within high-frequency word trials (B) and low-frequency word trials (C). Note the generally top-down (anterior to posterior) direction of connectivity for the low-frequency words compared with the primarily bottom-up (posterior to anterior) direction of connectivity for the high-frequency words.

presence of low-frequency words did appear to be a major factor determining engagement of areas in the DM or putative semantic network. This was shown by the pattern of greater activation for low-frequency words compared with both high-frequency words and nonwords in the task-positive regions of the IFJ (Figure 3E) and IFG (Figure 2E). High-frequency words and nonwords, how-

ever, showed similar levels of modest activation in these task-positive areas and less deactivation compared with low-frequency words in DM areas such as the bilateral PC. The bilateral SMG, which is adjacent to but does not entirely overlap either the DM or the task-positive network (Humphreys & Lambon Ralph, 2015; Nelson et al., 2010), showed a similar pattern to the one seen in the precuneus, in that across both experiments, low-frequency words showed more deactivation relative to high-frequency words and nonwords.

Contrasts between levels of word frequency and imageability were performed with the aim of revealing neural components of lexical and semantic processing. Although one of us (W.G.) has previously performed similar contrasts and interpreted activation for high-frequency and high-imageability words in terms of lexical and semantic processing, it is striking that the areas activated for high-frequency words in the word frequency contrast (Figure 2A) are largely the same ones activated for meaningless nonwords when contrasted with low-frequency words (Figure 2D). These areas include many that are often interpreted as carrying out semantic processing, such as the ATL, PC, and precuneus. Activation in the precuneus was also found for high- compared with low-imageability words, again pointing to a diverging pattern of activation in that this putative semantic area activated for both semantically rich (high-imageability words) and semantically impoverished (nonwords) conditions.

One way of interpreting the fMRI data is with respect to the behavioral data (Table 1). Overall, lexical decisions to words seemed to be harder than to nonwords, in that when RT was weighted by proportion of correct responses according to the IES, the weighted RT was significantly longer for words than nonwords. Low-frequency words also had longer RTs and were less accurate overall than high-frequency words. Thus, lexical decisions to low-frequency words were the most difficult. Corresponding patterns in the fMRI data showed the most activation for the most difficult condition in regions of the task-positive network and the most deactivation for the most difficult condition in regions of the DM network. Altogether, this suggests these areas are responding to difficulty of the task condition rather than, for example, the detailed semantic content associated with the words.

### Reconciling Difficulty Effects with Neurocognitive Components of Reading

A recent study by Taylor, Rastle, and Davis (2014) presented a thoughtful and detailed analysis of the issue of how to treat the combination of domain-general and information-specific effects of RT in functional neuroimaging data. Although their focus on RT only covers part of what went into the combination of RT and accuracy for the difficulty effects seen here, their treatment of general effort compared with engagement of particular cognitive processes is highly relevant. They used a reading aloud

task with regular words, irregular words, and pseudowords. RT was modeled in their fMRI analyses essentially the same way we have modeled it here. Contrasting words with pseudowords revealed activation for words in left AG, both before and after RT was included as a regressor. This is in line with what is typically seen for this contrast, whereas this area activated for pseudowords in the current study. Pseudowords were also the more difficult condition in their study, whereas words were more difficult in ours. Their basic argument is that accounting for variance in fMRI data because of RT accounts for differences in effort between conditions, which may result from multiple sources both specific and non-specific to reading, and what remains is the neural basis of any cognitive process, or representation, differentially engaged between the conditions being compared.

This framework from Taylor et al. (2014) may be useful for interpreting the current results. Although our results are counter to theirs and to what is typically seen when contrasting words with well-formed, pronounceable nonwords, it is in line with previous results across studies for reading words compared with various low-level baselines (Turkeltaub, Eden, Jones, & Zeffiro, 2002; Fiez & Petersen, 1998). Those meta-analyses highlight areas for word reading such as the IFG and vOT, areas found here to be particularly activated for low-frequency words compared with nonwords (Figures 2D and 3D). Therefore, areas known to be involved in word reading may be more highly recruited for more difficult words.

What, then, makes lexical decisions to low-frequency words so difficult? Although we do not yet have a clear answer, we can rule out one possibility. Perhaps low-frequency words are simply unknown to the participants who get them wrong and are therefore being treated like nonwords. If this were the case, low-frequency words would be expected to have behavioral and neural responses similar to nonwords. That is not the pattern seen in our data. Instead, it is the high-frequency words that were more like nonwords. The sole exception to this is in the IPS (Figure S3), but this was only significant in Experiment 2, so the dependability of the IPS pattern remains to be determined. The overall pattern strongly suggests predominance of difficulty effects, because essentially the only factor high-frequency words and nonwords have in common is that they were less difficult than low-frequency words.

As for why participants have such difficulty with low-frequency words in particular, the two-factor model of lexical decision from Balota and Chumbley (1984) may offer a clue. They considered word frequency as an objective approximation of familiarity. Using familiarity as a basis for setting the lexical decision criterion, high-frequency words are relatively easily classified because they are familiar, and pseudowords are classified as nonwords because they are clearly unfamiliar. In between are low-frequency words. On the basis of a familiarity criterion alone, responses to low-frequency words should

be intermediate between high-frequency words and nonwords, as typically shown by RT data. However, because of their intermediate status between high-frequency words and nonwords, additional nonlexical factors may come into play, perhaps related to the decision component of the lexical decision task. Indeed, the requirement to make a binary decision is common across numerous tasks that may or may not require lexical processing (go/no-go, *n*-back, etc.), consistent with low-frequency words engaging the MD network (Duncan, 2010). Thus, the insufficiency of familiarity as a basis for making rapid lexical decisions to low-frequency words may lead to engagement of domain-general decision-related components of the type supported by the MD network.

In addition to checking for replication of the main effects and interactions, we also conducted the second experiment so we could separate high- from low-frequency words to optimize for effective connectivity analyses. The comparisons described above highlight activation differences across different putative networks; they do not speak to the flow of information between networks. Considering the role of the IFJ and IPS in the task-positive network, AG and PC in the DM network, and the potentially intermediate role of vOT and SMG, these regions were chosen for our effective connectivity analyses. Lexical decisions for high-frequency words elicited a generally bottom-up, ventral to dorsal pattern, with activation flowing from vOT to SMG, AG, and IPS. There was, however, a generally top-down, anterior and dorsal to ventral pattern of activation for the block of low-frequency words, with activation flowing from IPS, PC, and IFJ to vOT. These results provide converging neural evidence supporting explanations of lexical decision, such as by Balota and Chumbley (1984), that an effortful, decision-related component is more involved for low- than high-frequency words, due presumably to the words being relatively unfamiliar. Overall, these patterns suggest the neural basis of the word-nonword distinction in lexical decision can be fundamentally influenced by word frequency. More generally, these findings suggest that results of contrasts typically interpreted in terms of relative semantic content may be more influenced by additional factors related to task difficulty than is widely appreciated.

This apparent preponderance of task difficulty effects leaves open the question of what areas are processing orthography, phonology, and semantics. Although this study was not designed to distinguish the neural basis of these cognitive components of reading, we noted above that many areas that are active for word reading compared with baseline, such as the IFG and vOT, are also part of the MD network and activated here for words (particularly those of low frequency) compared with nonwords. Many studies have also shown the IFG to be involved in phonological processing (Vigneau et al., 2006; Poldrack et al., 1999) and the vOT in orthographic processing (Dehaene, Cohen, Sigman, & Vinckier, 2005), or perhaps some combination of orthographic and phonological processing

(Mano et al., 2013; Yoncheva, Zevin, Maurer, & McCandliss, 2010). No doubt the question of what cognitive components are reflected in the activations for the hard compared with easy condition is related to the question of what makes the low-frequency words in our set more difficult to accept as words than it is to reject the nonwords. Because the high-frequency words and meaningless nonwords showed more similar activation than did the low-frequency words and nonwords, semantic information is unlikely to be the primary contributor to this difference. We also ensured the nonwords were pronounceable by constructing them to consist of valid English trigrams, and they did not differ from the words on multiple measures of orthographic typicality. Detailed matching on phonological variables, however, is not straightforward, as there may be multiple acceptable pronunciations for nonwords. If it were the case that lexical decisions could be made using phonological information or information about orthography–phonology mapping that would not be inconsistent with an interpretation of the recruitment of IFG and vOT components of the task-positive network for the more difficult condition.

The results call into question the idea that the network of areas found in meta-analyses of lexical decision (McNorgan et al., 2015) and semantics (Binder et al., 2009) to be activated across studies is specifically related to semantic processing. Instead, our findings are more consistent with an interpretation in which activation of task-positive regions such as the IFJ and IPS are balanced against activation of areas associated with the DM network such as the AG and PC. A recent study by Humphreys, Hoffman, Visser, Binney, and Lambon Ralph (2015), however, suggested the semantic and DM networks are at least partially distinct, with the ATL more associated with semantic processing and the AG more with the DM network. Additional work from the same group suggests the ATL plays a central role in semantic processing (Lambon Ralph, 2014). For example, although relatively dorsal areas of the ATL were found here to activate for nonwords, more ventral areas activated in conditions specific to semantics across multiple tasks (Humphreys et al., 2015), and during reading under conditions in which semantics is predicted to be most relevant (Hoffman, Lambon Ralph, & Woollams, 2015). The more ventral part of the ATL also suffers from signal dropout in standard EPI (Visser, Embleton, Jefferies, Parker, & Lambon Ralph, 2010), which is another reason why the current study does not directly speak to the role of the ventral ATL in semantics.

Our finding of both ATL and AG activation for meaningless nonwords does, however, suggest a trade-off between the MD and DM networks based on relative task demands. Other groups have also hypothesized demand-based trade-offs between the MD and DM networks (Fox & Raichle, 2007). For example, when working memory demands and RT increase during parametric modulation, interregional correlations within medial frontal nodes of

the DM network increase (Hampson, Driesen, Skudlarski, Gore, & Constable, 2006), along with an overall shift away from modularity and toward more coherence within the DM network (Vatansver, Menon, Manktelow, Sahakian, & Stamatakis, 2015). Our results suggest a similar shift away from modularity and toward network level processing with heightened difficulty effects during word recognition.

### Potential Limitations

One concern with dividing the words into two large blocks based on frequency is the potential for nonspecific order effects, rather than findings being related to word frequency per se. However, to the extent that activity in the IFJ reflects engagement of attention, general engagement of attention would be expected to be greater at the beginning of the scan than at the end. The fact that task-positive effects from IFJ emerged only for the low-frequency words, in the second half of the experiment, argues against any concerns related to disengagement of attention due to fatigue.

One specific point about the fMRI analyses: The RT regressor values for each trial were *z*-scored to avoid correlation with the binary regressors for each stimulus type. Indeed, this resulted in RT regressors that were never correlated with any other regressor at more than  $r = .35$ . Although still modestly correlated, this is well within the acceptable range for multicollinearity in multiple linear regression models (Kutner, Nachtsheim, Neter, & Li, 2005), so that the RT regressor will account for variance most associated with RT, whereas the regressors for the various stimulus conditions will account for variance most associated with those conditions.

Regarding the task, we should acknowledge that participants were responding essentially at chance for the low-frequency/low-imageability condition in Experiment 2. This makes neural activity for correct responses in this condition difficult to interpret, as it is unclear whether these were trials on which participants recognized the words or just happened to select the YES response. Additionally, use of a single lexical decision task leaves open the possibility that effects of semantic processing would be different in tasks that, for example, require an explicit semantic judgment. This certainly bears future investigation. There is, however, ample evidence for the influence of semantic variables in lexical decision tasks using well-formed nonwords. Studies have shown effects not only of the imageability variable used here but also of semantic priming (Evans et al., 2012), lexical semantic ambiguity (Rodd, Gaskell, & Marslen-Wilson, 2002; Borowsky & Masson, 1996), and semantic richness (Pexman, Hargreaves, Siakaluk, Bodner, & Pope, 2008).

### Interpretation of Deactivations Relative to Baseline

Implicit in the direct contrast between words and nonwords is the idea that direct task comparisons are interpretable

even if both conditions deactivate relative to a baseline such as, in this case, silent fixation (as in the PC activations in Figures 2E and 3E). This raises the question of what is happening during “baseline.” Prominent suggestions include semantic processing (Binder & Desai, 2011; Binder et al., 1999); recall of other information such as phonology, syntax, or episodic memory (Humphreys & Lambon Ralph, 2015; Cabeza, Ciaramelli, Olson, & Moscovitch, 2008); or a general idling relative to the task-positive state that involves internal metalizing or largely random associations of thought (Buckner et al., 2008; Andreasen et al., 1995). We propose a dynamic alternative in which, for example, the PC activates for semantic processing under conditions of minimal task difficulty (or at least does not activate for nonwords more than high-frequency words) but can be deactivated even for meaningful compared with non-meaningful stimuli when stimulus characteristics make the task difficult.

## Conclusion

The distinction between words and well-formed, pronounceable nonwords is thought to primarily arise from the fact that the word condition consists of meaningful strings and the nonword condition does not. This distinction has been a major factor in mapping brain areas thought to be related to semantic processing. Our results suggest that the neural basis of this fundamental distinction between words and nonwords can be altered by stimulus difficulty effects. Specifically, the finding that areas such as the AG and PC that feature prominently in the putative semantic network can be more active for nonwords than words calls into question the role of these areas in semantic processing. Overall, our results suggest that information-processing accounts may only partly explain neuroimaging data, whereas effects of task difficulty exert a greater influence than previously appreciated.

## Acknowledgments

This work was supported by a grant from the National Institutes of Health, Eunice Kennedy Shriver National Institute of Child Health and Human Development (grant number K99/R00 HD065839) to W. W. G. The authors thank Jeffrey R. Binder, M.D., for help with conceptualizing Experiment 1 and two anonymous reviewers for their extremely constructive feedback.

Reprint requests should be sent to William W. Graves, Department of Psychology, Rutgers University, 101 Warren St., Newark, NJ 07102, or via e-mail: william.graves@rutgers.edu.

## REFERENCES

- Adelman, J. S., Brown, G. D. A., & Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science, 17*, 814–823.
- Andreasen, N. C., O’Leary, D. S., Cizaldo, T., Arndt, S., Rezai, K., Watkins, G. L., et al. (1995). Remembering the past: Two facets of episodic memory explored with positron emission tomography. *American Journal of Psychiatry, 152*, 1576–1585.
- Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language, 55*, 290–313.
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database* (2.5 ed.). Linguistic Data Consortium, University of Pennsylvania, Philadelphia.
- Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 340–357.
- Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133*, 283–316.
- Bedny, M., & Thompson-Schill, S. L. (2006). Neuroanatomically separable effects of imageability and grammatical class during single-word comprehension. *Brain and Language, 98*, 127–139.
- Binder, J. R. (2007). Effects of word imageability on semantic access: Neuroimaging studies. In J. H. Hart Jr. & M. A. Kraut (Eds.), *Neural basis of semantic memory* (pp. 149–181). New York: Cambridge University Press.
- Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences, 15*, 527–536.
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19*, 2767–2796.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Rao, S. M., & Cox, R. W. (1999). Conceptual processing during the conscious resting state: A functional MRI study. *Journal of Cognitive Neuroscience, 11*, 80–93.
- Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience, 17*, 905–917.
- Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers, 33*, 73–79.
- Borowsky, R., & Masson, M. E. J. (1996). Semantic ambiguity effects in word identification. *Journal of Experimental Psychology: Learning Memory and Cognition, 22*, 63–85.
- Boukhrina, O., & Graves, W. W. (2013). Neural networks underlying contributions from semantics in reading aloud. *Frontiers in Human Neuroscience, 7*, 518.
- Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychologica Belgica, 51*, 5–13.
- Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain’s default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences, 1124*, 1–38.
- Cabeza, R., Ciaramelli, E., Olson, I. R., & Moscovitch, M. (2008). The parietal cortex and episodic memory: An attentional account. *Nature Reviews Neuroscience, 9*, 613–625.
- Cattinelli, I., Borghese, N. A., Gallucci, M., & Paulesu, E. (2013). Reading the reading brain: A new meta-analysis of functional imaging data on reading. *Journal of Neurolinguistics, 26*, 214–238.

- Clark, J. M., & Paivio, A. (2004). Extensions of the Paivio, Yuille, and Madigan (1968) norms. *Behavior Research Methods, Instruments, & Computers*, *36*, 371–383.
- Colombo, L., Pasini, M., & Balota, D. A. (2006). Dissociating the influence of familiarity and meaningfulness from word frequency in naming and lexical decision performance. *Memory and Cognition*, *34*, 1312–1324.
- Coltheart, M. (1981). MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, *33*, 497–505.
- Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, *36*, 384–387.
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*, 162–173.
- Cox, R. W., & Jesmanowicz, A. (1999). Real-time 3D image registration of functional MRI. *Magnetic Resonance in Medicine*, *42*, 1014–1018.
- Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences*, *9*, 335–341.
- Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., et al. (2010). How learning to read changes the cortical networks for vision and language. *Science*, *330*, 1359–1364.
- Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: Mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, *14*, 172–179.
- Evans, G. A. L., Lambon Ralph, M. A., & Woollams, A. M. (2012). What's in a word? A parametric study of semantic influences on visual word recognition. *Psychonomic Bulletin & Review*, *19*, 325–331.
- Fiez, J. A., & Petersen, S. E. (1998). Neuroimaging studies of word reading. *Proceedings of the National Academy of Sciences, U.S.A.*, *95*, 914–921.
- Fox, M. D., & Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience*, *8*, 700–711.
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences, U.S.A.*, *102*, 9673–9678.
- Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *Neuroimage*, *19*, 1273–1302.
- Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, *12*, 395–427.
- Graves, W. W., Desai, R., Humphries, C., Seidenberg, M. S., & Binder, J. R. (2010). Neural systems for reading aloud: A multiparametric approach. *Cerebral Cortex*, *20*, 1799–1815.
- Gusnard, D. A., & Raichle, M. E. (2001). Searching for a baseline: Functional imaging and the resting brain. *Nature Reviews Neuroscience*, *2*, 685–694.
- Hampson, M., Driesen, N. R., Skudlarski, P., Gore, J. C., & Constable, R. T. (2006). Brain connectivity related to working memory performance. *Journal of Neuroscience*, *25*, 13338–13343.
- Hauk, O., Davis, M. H., Kherif, F., & Pulvermüller, F. (2008). Imagery or meaning? Evidence for a semantic origin of category-specific brain activity in metabolic imaging. *European Journal of Neuroscience*, *27*, 1856–1866.
- Hoffman, P., Lambon Ralph, M. A., & Woollams, A. M. (2015). Triangulation of the neurocomputational architecture underpinning reading aloud. *Proceedings of the National Academy of Sciences, U.S.A.*, *112*, E3719–E3728.
- Humphreys, G. F., Hoffman, P., Visser, M., Binney, R. J., & Lambon Ralph, M. A. (2015). Establishing task- and modality-dependent dissociations between the semantic and default mode networks. *Proceedings of the National Academy of Sciences, U.S.A.*, *112*, 7857–7862.
- Humphreys, G. F., & Lambon Ralph, M. A. (2015). Fusion and fission of cognitive functions in the human parietal cortex. *Cerebral Cortex*, *25*, 3547–3560.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). Regression models for quantitative and qualitative predictors. In *Applied linear statistical models* (pp. 294–342). Boston: McGraw-Hill.
- Lambon Ralph, M. A. (2014). Neurocognitive insights on conceptual knowledge and its breakdown. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *369*, 20120392.
- 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.
- McNorgan, C., Chabal, S., O'Young, D., Lukic, S., & Booth, J. R. (2015). Task dependent lexicality effects support interactive models of reading: A meta-analytic neuroimaging review. *Neuropsychologia*, *67*, 148–158.
- Monsell, S. (1991). The nature and locus of word frequency effects in reading. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 148–197). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General*, *118*, 43–71.
- Nelson, S. M., Cohen, A. L., Power, J. D., Wig, G. S., Miezin, F. M., Wheeler, M. E., et al. (2010). A parcellation scheme for human left lateral parietal cortex. *Neuron*, *67*, 156–170.
- Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*, 255–287.
- Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement*, *76*, 1–25.
- Perani, D., Cappa, S. F., Schnur, T., Tettamanti, M., Collina, S., Rosa, M. M., et al. (1999). The neural correlates of verb and noun processing: A PET study. *Brain*, *122*, 2337–2344.
- Pexman, P. M., Hargreaves, I. S., Siakaluk, P. D., Bodner, G. E., & Pope, J. (2008). There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin & Review*, *15*, 161–167.
- Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology*, *10*, 377–500.
- Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage*, *10*, 15–35.
- Ramsey, J. D., Hanson, S. J., & Glymour, C. (2011). Multi-subject search correctly identifies causal connections and most causal directions in the DCM models of the Smith et al. simulation study. *Neuroimage*, *58*, 838–848.
- Ramsey, J. D., Hanson, S. J., Hanson, C., Halchenko, Y. O., Poldrack, R. A., & Glymour, C. (2010). Six problems for causal inference from fMRI. *Neuroimage*, *49*, 1545–1558.
- Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, *46*, 245–266.
- Rowe, J. B., Toni, I., Josephs, O., Frackowiak, R. S. J., & Passingham, R. E. (2000). The prefrontal cortex: Response selection or maintenance within working memory? *Science*, *288*, 1656–1660.

- Rubenstein, H., Garfield, L., & Millikan, J. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior*, *9*, 487–494.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, *23*(Suppl. 1), 208–219.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic axis of the human brain*. Stuttgart: Thieme.
- 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.
- Taylor, J. S. H., Rastle, K., & Davis, M. H. (2014). Interpreting response time effects in functional imaging studies. *Neuroimage*, *99*, 419–433.
- Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Townsend, J. T., & Ashby, F. G. (1978). Methods for modeling capacity in simple processing systems. In J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 199–239). New York: Psychology Press.
- Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *Neuroimage*, *16*, 765–780.
- Ullsperger, M., & von Cramon, D. Y. (2004). Neuroimaging of performance monitoring: Error detection and beyond. *Cortex*, *40*, 593–604.
- Vatanserver, D., Menon, D. K., Manktelow, A. E., Sahakian, B. J., & Stamatakis, E. A. (2015). Default mode dynamics for global functional integration. *Journal of Neuroscience*, *35*, 15254–15262.
- Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., et al. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage*, *30*, 1414–1432.
- Visser, M., Embleton, K. V., Jefferies, E., Parker, G. J., & Lambon Ralph, M. A. (2010). The inferior, anterior temporal lobes and semantic memory clarified: Novel evidence from distortion-corrected fMRI. *Neuropsychologia*, *48*, 1689–1696.
- Vogel, A. C., Miezin, F. M., Petersen, S. E., & Schlaggar, B. L. (2012). The putative visual word form area is functionally connected to the dorsal attention network. *Cerebral Cortex*, *22*, 537–549.
- Wager, T. D., Sylverster, C.-Y. C., Lacey, S. C., Nee, D. E., Franklin, M., & Jonides, J. (2005). Common and unique components of response inhibition. *Neuroimage*, *27*, 323–340.
- Wise, R. J. S., Howard, D., Mummery, C. J., Fletcher, P., Leff, A., Buchel, C., et al. (2000). Noun imageability and the temporal lobes. *Neuropsychologia*, *38*, 985–994.
- Yoncheva, Y. N., Zevin, J. D., Maurer, U., & McCandliss, B. D. (2010). Auditory selective attention to speech modulates activity in the visual word form area. *Cerebral Cortex*, *20*, 622–632.