

# Semantic Context and Visual Feature Effects in Object Naming: An fMRI Study using Arterial Spin Labeling

Julia Hocking, Katie L. McMahon, and Greig I. de Zubicaray

## Abstract

■ Previous behavioral studies reported a robust effect of increased naming latencies when objects to be named were blocked within semantic category, compared to items blocked between category. This semantic context effect has been attributed to various mechanisms including inhibition or excitation of lexico-semantic representations and incremental learning of associations between semantic features and names, and is hypothesized to increase demands on verbal self-monitoring during speech production. Objects within categories also share many visual structural features, introducing a potential confound when interpreting the level at which the context effect might occur. Consistent with previous findings, we report a significant increase in response latencies when naming categorically related objects within blocks, an effect associated

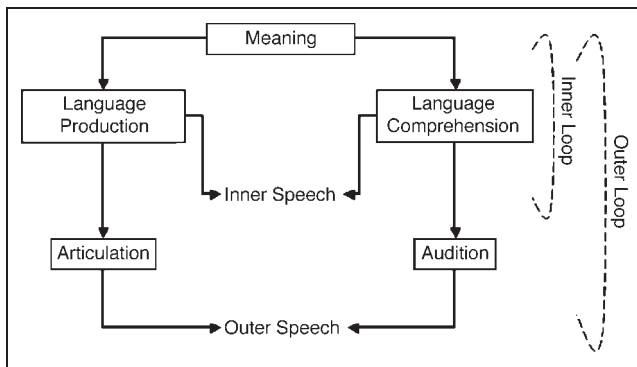
with increased perfusion fMRI signal bilaterally in the hippocampus and in the left middle to posterior superior temporal cortex. No perfusion changes were observed in the middle section of the left middle temporal cortex, a region associated with retrieval of lexical-semantic information in previous object naming studies. Although a manipulation of visual feature similarity did not influence naming latencies, we observed perfusion increases in the perirhinal cortex for naming objects with similar visual features that interacted with the semantic context in which objects were named. These results provide support for the view that the semantic context effect in object naming occurs due to an incremental learning mechanism, and involves increased demands on verbal self-monitoring. ■

## INTRODUCTION

Nearly all modern theories of speech production make the assumption that competition occurs naturally between a target word and its lexico-semantic neighbors prior to its selection for articulation (see Levelt, Roelofs, & Meyer, 1999). Ad hoc evidence for a competitive selection process in normal speech is provided by analyses of occasional errors showing semantic substitutions, in which an intended word (e.g., warm) is spontaneously replaced by a semantically related word (e.g., cold; Harley & MacAndrew, 2001). Self-repairs of these errors, or corrections of them without external prompting, imply the existence of a monitoring process (or processes) operating during speech production (Postma, 2000). Evidence from spontaneous self-correction during speech has shown that this monitoring process can occur at different levels. For example, a word can be articulated in full and then corrected, thus self-monitoring of overt speech provides the speaker an opportunity to correct any speech errors. Alternatively, only the first syllable of a word may be spoken before being corrected. This self-interruption at the onset of articulation suggests an

inner monitor that checks the correctness of the verbal message at a phonological level (i.e., the output of phonological encoding) prior to phonetic encoding and articulation (Levelt et al., 1999; see also Indefrey & Levelt, 2004) (see Figure 1).

In a recent meta-analysis of the relevant neuroimaging literature on word production, Indefrey and Levelt (2004) identified the bilateral superior temporal gyri as the possible neural correlates of verbal self-monitoring. Although the meta-analysis was only able to directly identify the correlates of the outer loop, an “economical assumption” was made attributing the inner loop to the same regions or to a subset of them (p. 125). Self-monitoring has been predominantly investigated through the experimental manipulation of auditory feedback during overt speech. For example, using PET, McGuire, Silbersweig, and Frith (1996) increased demands on speech monitoring by giving normal, distorted, or alien external feedback to subjects reading aloud nouns. This enabled them to determine the brain regions involved in a mismatch between subjects’ intended and perceived verbal output. Increased activation was observed in posterior and middle superior temporal cortices when demands on self-monitoring were increased. Similarly, Fu et al. (2006) manipulated auditory feedback in an overt adjective naming task using fMRI. They reported



**Figure 1.** The perceptual loop theory of speech monitoring (Levelt, 1983, 1989; adapted from Slevc & Ferreira, 2006).

a network of regions engaged in self-monitoring, with greater activation for correct recognition of self-generated speech in the middle and superior temporal gyri.

Both the Fu et al. (2006) and McGuire et al. (1996) studies reported bilateral superior temporal activation, a result supported by additional studies that manipulated different types of verbal feedback (Allen et al., 2005; Hirano et al., 1997). However, manipulations of phonological or semantic characteristics of speech are reported to elicit predominantly left-lateralized temporal activation (Binder et al., 1995; Demonet et al., 1992; Zatorre, Evans, Meyer, & Gjedde, 1992). Are both the right and left superior temporal regions responding to the process of verbal self-monitoring? It is possible that previous methods may have introduced an attentional confound to their experiment due to the qualitative differences between conditions in which the verbal signal is being monitored. For example, when giving alien or distorted verbal feedback to a subject, there is a mismatch between the actual spoken output and the expected verbal feedback, that is, the subject expects to hear their own voice but instead receives either a distorted signal or someone else's voice. The monitoring process being measured in these mismatched speech conditions is thus not specific to verbal self-monitoring, being confounded by increased attention to the unexpected difference between self-generated speech and the experimentally modified verbal feedback signal (e.g., Fu et al., 2006; McGuire et al., 1996). In contrast to the bilateral activation reported in these studies, Allen et al. (2005) found that verbal monitoring of degraded speech increased activation solely in the left superior temporal gyrus, a finding consistent with the role of the left superior temporal gyrus in auditory verbal processing (Scott, Blank, Rosen, & Wise, 2000). Importantly, this effect was increased for self-generated speech but attenuated during alien speech, suggesting that the process of verbal self-monitoring (rather than monitoring distorted or alien speech) more specifically engages the left hemisphere. Taking all this evidence into account, we therefore predict that left (rather than bilat-

eral) superior temporal activation will be observed in a task where self-monitoring of speech is increased while attention and auditory feedback is held constant across conditions.

According to the perceptual loop theory (Levelt, 1989), in addition to monitoring for speech errors, the verbal self-monitoring system additionally monitors for semantic correctness/appropriateness. In order to manipulate the demands on verbal self-monitoring, the present study uses a semantic blocking paradigm to increase the level of competition during speech production (Kroll & Stewart, 1994). This paradigm entails a manipulation of the context in which identical object pictures are named. Behavioral studies have shown that naming latencies increase when objects are grouped or blocked into semantically related or homogeneous categories (e.g., chicken, whale, rabbit, giraffe, lamb) compared with noncategorized or heterogeneous groups or blocks (e.g., skirt, melon, tractor, cat, radish) (Ganushchak & Schiller, 2008; Belke, Meyer, & Damian, 2005; Maess, Friederici, Damian, Meyer, & Levelt, 2002; Vigliocco, Lauer, Damian, & Levelt, 2002; Damian, Vigliocco, & Levelt, 2001; Levelt et al., 1999; Kroll & Stewart, 1994). When naming objects in semantically homogeneous contexts, the self-monitoring system is presumed to be engaged to a greater degree than in heterogeneous contexts (e.g., to check whether the correct alternative has been chosen; see Ganushchak & Schiller, 2008; Maess et al., 2002).

In the only neuroimaging study of the semantic blocking paradigm conducted to date, using MEG, Maess et al. (2002) reported effects for naming in the homogeneous context observed in the left temporal cortex. Firstly, they reported an early effect in a time window of 150–225 msec post picture onset that they attributed to the mechanism responsible for the context effect. Secondly, they reported a later effect also in the left temporal cortex (450–475 msec post picture onset) that they attributed to internal self-monitoring. Unfortunately, as Maess et al. acknowledged, due to the limited source-localization capabilities of MEG, their analysis did not provide the real extent and shape of the activated regions (p. 460). Consequently, they were not able to identify which regions of the temporal cortex were responsible for the different effects.

At least three different mechanisms have been proposed to account for the semantic context effect. Two of these emphasize activation changes occurring at a processing level involving retrieval of lexico-semantic information. For example, increased naming latencies could be due to suppression of activation of a target word's representations following naming of a recently used, semantically related target (an *inhibition* account; Vitkovitch, Rutter, & Read, 2001; McCarthy & Kartsounis, 2000). Alternatively, coactivation of lexical representations of semantically related items may occur, thus increasing competition between a number of possible

responses, as has been proposed for picture–word interference (PWI) and competitor priming effects in naming (an *excitation* account; Maess et al., 2002; Damian et al., 2001). As Damian and Als (2005) noted, one problem with these accounts is that they do not explain how previously named objects become competitors on subsequent trials. Moreover, these accounts predict that spacing trials further apart will reduce the effect, as activation should return to a resting level, which is not the case (Howard, Nickels, Coltheart, & Cole-Virtue, 2006). Damian and Als therefore proposed an alternative account attributing the context effect to an episodic memory mechanism involving incremental learning of the associations between semantic features and names, as each object is named several times across the experiment and within experimental blocks. This explanation has recently been instantiated successfully in a computational model (Oppeheim, Dell, & Schwartz, 2007). The three differing accounts thus provide distinct anatomical and directional hypotheses for the current experiment. If the context effect occurs at a lexico-semantic level, then increased (excitation) or decreased (inhibition) activation would be expected in the left middle temporal gyrus—a region associated with retrieval of lexico-semantic information in PWI and competitor priming studies of naming (de Zubicaray, McMahan, Eastburn, & Pringle, 2006; Indefrey & Levelt, 2004; de Zubicaray, Wilson, McMahan, & Muthiah, 2001). On the other hand, if the context effect is due to incremental learning, one might expect regions involved in associative encoding mechanisms such as the medial temporal lobe, in particular, the hippocampal formation, to be involved (see Cabeza & Nyberg, 2000; Mayes & Roberts, 2000 for reviews).

One further question remains as to whether these potential mechanisms could interact with visual features of the to-be-named objects within homogeneous sets. In their experiment, Kroll and Stewart (1994) controlled for object familiarity across different categories, but they did not control for visual similarity of stimuli. Within-category items, in particular, objects from the living category such as animals or fruits, share many visual features. This visual “confusability” may therefore have driven the increased naming latencies during the categorized blocks. Indeed, lexical selection is more difficult for objects that share perceptual features with other objects from their category (Humphreys, Price, & Riddoch, 1999; Vitkovitch, Humphreys, & Lloyd-Jones, 1993), that is, both visual and semantic variables influence naming latencies. This leaves open the possibility that the observed blocking effect arose from nonlinguistic processes such as visual similarity, which may interact with the naming process.

To investigate this, Damian et al. (2001) utilized the Kroll and Stewart (1994) blocking paradigm, but attempted to control for visual similarity of the stimuli. Participants judged visual similarity on a 1 (*not at all similar*) to

5 (*very similar*) rating scale, independent of conceptual category. They replicated the Kroll and Stewart semantic blocking effect using within-category items that had been rated with an average similarity of 2.45/5 (noncategory items were rated at 1.97/5). This rating reflected categorically related stimuli that were not visually similar. However, this issue needs to be addressed more stringently by utilizing both visually similar and non-similar items in both homogeneous and heterogeneous contexts. In the present study, we therefore explicitly manipulate visual feature similarity in both homogeneous and heterogeneous contexts, using empirically derived ratings for shared and nonshared visual form and surface features based on the corpus of semantic feature norms provided by McRae, Cree, Seidenberg, and McNorgan (2005). Using a comparison between both similar and nonsimilar features provides a more rigorous test of whether visual semantic features influence naming latencies.

Processing complex visual features has been reported to modulate activation in the more anterior portions of the ventral visual stream, in the anterior fusiform gyri and the anteromedial temporal cortex (Moss, Rodd, Stamatakis, Bright, & Tyler, 2005; Tyler et al., 2004). In the nonhuman primate literature, there is compelling evidence that the anteromedial temporal cortex not only is involved in mnemonic processes during object recognition tasks (Buffalo, Reber, & Squire, 1998; Meunier, Bachevalier, Mishkin, & Murray, 1993; Zola-Morgan, Squire, Amaral, & Suzuki, 1989) but is also engaged during perceptual discrimination tasks with no memory component (Bussey, Saksida, & Murray, 2003; Murray & Gaffan, 1994). Increasing evidence from human functional imaging is supporting the involvement of the medial anterior temporal lobes, in particular, the perirhinal cortex, in tasks involving discrimination between complex visual featural combinations (Devlin & Price, 2007; Moss et al., 2005; Tyler et al., 2004; Devlin et al., 2002; Lerner, Hendler, Ben-Bashat, Harel, & Malach, 2001). Correspondingly, patients with selective damage to the medial temporal cortex have more difficulties recognizing living than nonliving things (Gainotti, 2000; Gainotti, Silveri, Daniele, & Giustolisi, 1995): a class of objects that have more shared visual features compared with nonliving items, hence, placing greater demands on the selection of an appropriate living concept from stored representations. Converging evidence from both the human and nonhuman primate literature therefore suggests that the anteromedial temporal cortex is involved in discriminating between visually similar items. Based on this evidence, we expect to observe increased activation in the anteromedial temporal cortex, particularly the perirhinal cortex, when subjects name items in blocks comprising similar visual features relative to dissimilar visual features.

The semantic blocking paradigm requires that an overt verbal response is made; however, the use of

overt speech in conventional fMRI experiments is problematic due to the introduction of susceptibility artifacts (Huang, Francis, & Carr, 2008). For example, speech-related movement causes signal changes in brain regions close to tissue boundaries (Friston, Williams, Howard, Frackowiak, & Turner, 1996; Hajnal et al., 1994). This can lead to false-positive results, where activation due to movement-related artifactual signal changes mimics task-related blood oxygen level-dependent (BOLD) activation. The experimental design requires blocking trials with a short stimulus onset asynchrony, thus precluding the use of a sparse imaging design such as that used with previous PWI and competitor priming paradigms (e.g., de Zubicaray et al., 2001, 2006). This is because the sparse design requires a relatively long stimulus onset asynchrony in order to acquire images coincident with the estimated hemodynamic response to each trial. For this reason, we decided to employ a perfusion-based arterial spin labeling (ASL) fMRI sequence as an alternative to BOLD, as this method does not depend on susceptibility effects for contrast and is thus not influenced by overt speech artifacts (see Troiani et al., 2007; Kemeny, Ye, Birn, & Braun, 2005). In addition, ASL has advantages over BOLD imaging in areas such as the medial temporal lobe that show partial signal loss due to static susceptibility gradients. ASL is a noninvasive technique that uses a radio-frequency pulse to change (label or tag) the magnetization of protons in the arterial water as it flows through the neck into the brain. Tagged images are subtracted from control (non-tagged) scans to produce a direct measure of arterial blood flow (see Methods section). Although the signal-to-noise ratio obtained by ASL is lower than that of BOLD, this can be ameliorated by the use of higher field strengths (4 T in this study) and ASL provides a direct rather than indirect (as for BOLD) measure of cerebral blood flow (for a comparison between BOLD and ASL techniques, see Aguirre, Detre, Zarahn, & Alsop, 2002).

In summary, the present study investigates different mechanisms potentially responsible for the semantic context effect in object naming. We also examine whether naming objects in semantically related (homogeneous) blocks increases demands on self-monitoring of speech. Finally, the semantic blocking paradigm also enables us to manipulate the demands on discriminating between visually similar and dissimilar items.

## METHODS

### Subjects

Eighteen people (9 women, mean age = 25.4 years, range = 19–36 years) gave written informed consent to take part. All had normal or corrected-to-normal vision, were right-handed, and were free of any history of neurological problems. The study was approved by The University of Queensland's Medical Research Ethics Committee, and all subjects were paid a gratuity of AU\$30 to cover their expenses.

### Experimental Design and Stimuli

#### Stimuli

In this experiment, subjects were required to overtly name visually presented black and white photographs depicting real objects. Fifty items were selected from five object categories: animals, clothing, fruits, vegetables, and vehicles (see Appendix for a complete list). Photographs were obtained from either the Hemara Photo Objects CD collection or downloaded from the Internet. Visual feature ratings were based on the empirically derived visual form and surface feature ratings provided by McRae et al. (2005). Although derived from words (as opposed to the pictorial exemplars used here), these visual feature ratings are based on a “knowledge type” taxonomy that is presumed to reflect the brain regions responsible for processing that type of knowledge (i.e., visual form and surface; see Cree & McRae, 2003). During naming, we therefore expected greater demands to be placed on regions involved in stored knowledge of visual form and surface features because it is through access to stored abstract representations that raters list these features. Selection of stimuli was based on equating variables at visual perceptual, semantic, and lexical levels: (1) luminance, (2) mean pixelwise Euclidean distance between pictures, (3) number of shared visual features within a category, (4) overall familiarity rating, (5) number of syllables, (6) number of phonemes, (7) CELEX noun frequency, and (8) phonological-neighborhood density (Balota et al., 2007; Cree & McRae, 2003; Baayen, Piepenbrock, & Gulikers, 1995; see Table 1).

In order to segregate the stimuli into groups with similar visual features (V+) versus dissimilar features (V–), five objects from each category were selected on the basis that they had been rated as having (1) no visual

**Table 1.** Matching Variables for Stimuli Used in the Experiment

| <i>Stimuli</i>      | <i>Luminance</i> | <i>Familiarity</i> | <i>Syllables</i> | <i>Phonemes</i> | <i>CELEX (Log) Frequency</i> | <i>Phonological Neighbors</i> |
|---------------------|------------------|--------------------|------------------|-----------------|------------------------------|-------------------------------|
| Visually similar    | 237.38 (4.86)    | 6.03 (1.64)        | 1.76 (0.66)      | 5.08 (1.58)     | 1.04 (0.64)                  | 6.42 (7.99)                   |
| Visually dissimilar | 238.91 (4.86)    | 5.87 (1.63)        | 2.08 (1.04)      | 5.2 (2)         | 0.92 (0.69)                  | 10.5 (12.37)                  |

Mean value, with standard deviation in parentheses, for variables used to match 50 stimuli (25 visually similar, 25 visually dissimilar). Luminance ( $\text{cd/m}^2$ ) was matched across all stimuli using Adobe Photoshop v.7.0. Rating data for phonological neighborhood density was not available for two items (earmuffs, unicycle).

features in common with each other, or (2) shared a minimum of 3/4 features with other stimuli within the group. For example, in the homogeneous visually similar (HomV+) animal category, the object “cat” shared the following features with the other exemplars in the block: has\_4\_legs (3 others), has\_fur (2 others), has\_a\_tail (all others), has\_legs (3 others). In contrast, for the homogeneous visually dissimilar (HomV-) animal category, the object “chicken” was rated with four visual form and surface features: has\_a\_beak, has\_feathers, has\_legs, has\_wings, none of which were listed as features for the other four animals in that block.<sup>1</sup> In order to ensure that any differences that might be detected between visually similar and dissimilar items were not due to low-level visual perceptual differences, we measured the pixelwise Euclidean distance between mean images, adapted from Grill-Spector et al. (1999). The measure gives an index of the physical difference between pixels in two images, where physically identical images have an index of 0 and the greatest difference between images has an index of 0.8684 (based on the 400 × 400 pixel size used here). We obtained an index of 0.0254 for the mean Euclidean distance between a mean image of the 25 visually similar pictures and mean image of the 25 visually dissimilar pictures (see Table 1 for data on additional matching variables).

### Design

Prior to scanning, subjects were familiarized with all 50 experimental stimuli by viewing them on a computer screen, in random order, and overtly naming all stimuli

three times. On the first viewing, the appropriate label was provided on the screen with the picture. On the subsequent two viewings, the name was removed and erroneous responses were corrected. All subjects were instructed to name the presented objects as quickly and as accurately as possible during scanning. Vocal responses were recorded via a microphone attached to the head coil in order to measure naming accuracy. Stimuli were presented on a front-projected screen viewed via a mirror mounted on the head coil.

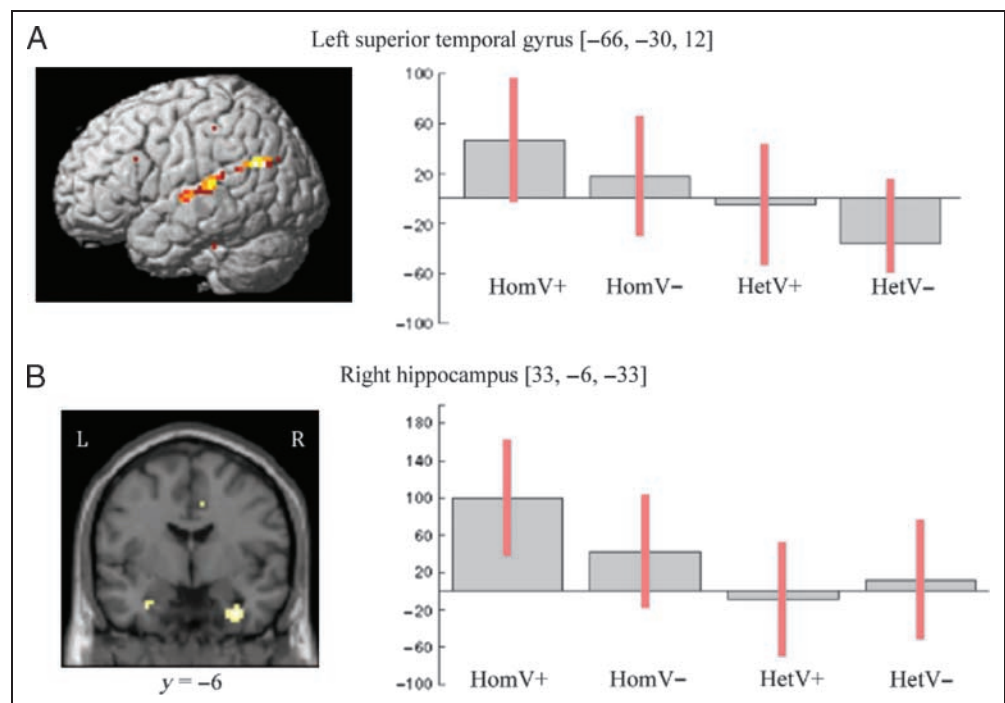
The manipulation of semantic category and visual features resulted in a total of four conditions.

1. Homogeneous (i.e., within category) blocks comprised of visually similar objects (HomV+)
2. Homogeneous blocks comprised of visually dissimilar objects (HomV-)
3. Heterogeneous (i.e., between category) blocks comprised of visually similar objects from Condition 1 (HetV+)
4. Heterogeneous blocks, comprised of visually dissimilar objects from Condition 2 (HetV-)

The same 50 objects were used across homogeneous and heterogeneous blocks, thus heterogeneous blocks acted as a control for the homogeneous blocks by using the same stimuli but presenting blocks of items between category, with objects that were either visually similar (HetV+) or visually dissimilar (HetV-).

Each of the four conditions listed above contained five trials per block, with a trial duration of 3000 msec. The visual stimulus was presented on a screen with a white background for 800 msec, replaced immediately by a

**Figure 2.** Main effect for naming in homogeneous versus heterogeneous blocks. (A, left) Left hemisphere activation for homogeneous relative to heterogeneous blocks rendered on a standard MNI surface template at  $p < .001$ , uncorrected, with a 5-voxel minimum cluster size. Plot (right) shows the mean centered effect size at the peak voxel  $[-66, -30, 12]$ . (B) Hippocampal activation, rendered on an averaged T1 coronal section of the standardized brain, and plot (right) of the mean centered effect size at the peak voxel  $[33, -6, -33]$ . Hom = homogeneous; Het = heterogeneous; V+ = visually similar; V- = visually dissimilar.



centrally positioned black fixation cross for 2200 msec. This resulted in a total block time of 15 sec. Between blocks, the screen was blank (white) for 4000 msec, followed by the presentation of a fixation cross 1000 msec prior to the onset of the next block. Over the course of the experiment, every subject was presented with 20 blocks from each of the four conditions. This gave a total of 80 blocks (400 trials), split into two successive scanning runs (40 blocks per run). The order of conditions was counterbalanced within and between subjects. Although we were not interested in the effect of object category, this was fully counterbalanced across conditions, runs, and subjects.

### Behavioral Study

We recorded subject naming responses during the fMRI experiment. However, even after filtering the scanner noise utilizing both the Cusack, Cumming, Bor, Norris, and Lyzenga (2005) and Nelles et al. (2003) noise cancellation software packages, enough distortion remained to prohibit accurate measurement of response latencies with the voice key software. Therefore, an additional behavioral study was run with a separate group of nine subjects to investigate differences in naming latencies for each condition. These subjects received the same experimental design parameters as the scanned subjects, presented on a laptop PC using COGENT presentation software ([www.vislab.ucl.ac.uk/cogent](http://www.vislab.ucl.ac.uk/cogent)) on a Matlab v7.1 platform (Mathworks, Sherborne, MA, USA). Response latencies were recorded via a microphone with a voice activated relay system from stimulus to response onset and were analyzed with a 2 Context (Hom vs. Het)  $\times$  2 Visual (V+ vs. V-)  $\times$  2 Run (first vs. second) repeated measures ANOVA.

### Data Acquisition

MR images were acquired with a 4-T Bruker Medspec, using a transverse electromagnetic (TEM) transmit and receive head coil (Vaughan et al., 2002). Head movement was controlled with foam padding within the head coil. Pulsed ASL was used to collect the functional time series, using a modified version of FAIR (Kim, 1995). A saturation slab was applied inferior to the imaging slice, at a time  $TI_1$  after either global or slice-selective inversion (control and tagging phases; Wang et al., 2002, 2003). The saturation slab allowed for a more defined tagging bolus, with  $TI_1$  set to 800 msec (Wong, Buxton, & Frank, 1998b). The area of the slice-selective inversion was 2 cm thicker than the imaging slab, with a 1-cm margin at each edge, to ensure optimum inversion. A delay between saturation and excitation, greater than the arterial transit time, was used to allow all labeled blood to flow into the imaging slices by the time images were acquired ( $TI_2 = 1000$  msec) (Wong, Buxton, &

Frank, 1998a; Alsop & Detre, 1996). Control and tagging phases were interleaved throughout the time series, and acquisition of the imaging slices was done using gradient-echo EPI. The imaging parameters were as follows: matrix size  $64 \times 64$ ; TR/TE = 2500/11 msec; and FOV  $230 \times 230$ . Twelve slices, 6 mm thick with a 1.5-mm gap, were acquired in ascending order, and angled to optimize coverage of the temporal lobes. Two imaging runs were acquired, each with 327 brain volumes. A 3-D T1-weighted high-resolution image was acquired within the same scanning run, using an MP-RAGE sequence (TI = 700 msec, TR = 1500 msec, TE = 3.35 msec,  $256 \times 256 \times 256$  matrix, and  $0.9 \text{ mm}^3$  isotropic voxels).

### Data Analysis

Following reconstruction of the raw images, the first four scans in the time series of each scanning run were removed to allow for T1 equilibration effects. Rigid-body motion correction was carried out on all images using INRIalign (Freire, Roche, & Mangin, 2002) and a mean realigned image was created for each subject. Acquisition of the perfusion image time series was carried out by implementing a pairwise simple subtraction between temporally adjacent labeled (tagged) and control (non-tagged) acquisitions, resulting in perfusion image volumes with an effective TR of 5 sec (Aguirre et al., 2002; Wong, Buxton, & Frank, 1997). The mean image for each subject was coregistered to the corresponding structural (T1) image, using SPM5 (Wellcome Trust Centre for Neuroimaging, London). Each individual's T1 image was then normalized to the SPM5 MNI T1 template. The resulting spatial normalization parameters were applied to the CBF time series (across both runs), and resliced to  $3 \times 3 \times 3$  mm voxels. Following normalization of all images in the time series, images were spatially smoothed with a full-width half-maximum Gaussian kernel of 10 mm. Each subject's coregistered T1 structural scan was segmented to create a gray matter image. This image was subsequently used as an explicit mask when estimating effects at the first-level, single-subject analysis.

Stimulus onset times corresponding to the small number of erroneous responses were modeled as effects of no interest (200 responses, or 2.8% of scans). Incorrect responses were due to semantic substitutions (41%), hesitations (e.g., saying "er...") followed by a correct response (34%), no response at all (17.5%), hesitation followed by no response (3.5%), unrelated or nonword response (3%), or hesitation followed by semantic substitution (1%). First-level statistical analyses (single subject and fixed effects) modeled each trial type independently by convolving the onset times with a first-order gamma response function. The data were high-pass filtered using a set of discrete cosine basis functions with an infinite cutoff period. Parameter estimates (for correct responses only) were calculated for all voxels

**Table 2.** Mean Naming Latencies

| <i>Condition</i> | <i>Mean (msec)</i> |
|------------------|--------------------|
| Hom V+           | 774 (72)           |
| Hom V-           | 778 (69)           |
| Het V+           | 739 (56)           |
| Het V-           | 743 (60)           |

Mean naming latencies collapsed across runs for each condition with standard deviation in parentheses. Hom = homogeneous context; HET = heterogeneous context; V+ = visually similar; V- = visually dissimilar.

using the general linear model by computing a contrast image for each condition relative to fixation. These parameter estimates were then fed into a second-level  $2 \times 2 \times 2$  ANOVA, modeling context (Hom vs. Het), visual similarity (V+ vs. V-), and run (first vs. second), resulting in four different conditions (HomV+, HomV-, HetV+, HetV-) repeated across the two scanning runs. In the preliminary analysis that modeled the effects of each run separately, the effects of context and visual features did not interact with run. The functional imaging results reported here are therefore based on an analysis that summed over the effect of the two runs, analyzed in a  $2$  (context)  $\times$   $2$  (visual features) ANOVA.

### Regions of Interest

Based on the evidence reviewed in the Introduction, we adopted the following a priori regions of interest (ROIs): the entire left superior temporal gyrus, the entire left middle temporal gyrus, the bilateral hippocampus, and the perirhinal cortex. This entailed constructing binary image masks for the first three ROIs using the LONI probabilistic atlas for gray matter regions at a 50% probability level (Shattuck et al., 2008). We used the probabilistic atlas of the perirhinal cortex provided by Devlin and Price (2007). Left and right hemisphere masks were generated separately for both the hippocampal and perirhinal ROIs. In addition to a whole-brain exploratory threshold of  $p < .001$ , we report only those effects significant at  $p < .05$ , family-wise error corrected for the search volume for these three ROIs.

### Behavioral Results

Latencies shorter than 300 msec and longer than 1500 msec were excluded from the analysis. Along with errors (using the same criteria for responses during scanning), this resulted in 12% of responses being excluded from the analysis. The  $2 \times 2 \times 2$  ANOVA revealed significant main effects for context (homogeneous vs. heterogeneous) and run (first vs. second), but no effect for visual features (similar vs. dissimilar) nor interactions between any factors. The main effect of run was consis-

tent with previous studies investigating the semantic blocking effect. However, as there was no interaction between scanning run and context, we collapsed the data across runs and analyzed the data as a  $2$  (context)  $\times$   $2$  (visual features) ANOVA. This is because we were specifically interested in the effect of context and the manipulation of visual features. This analysis revealed a significant effect of context (Hom vs. Het) [ $F(1, 8) = 7.182, p < .05$ ], with mean subject responses 69 msec faster in the heterogeneous compared with homogeneous conditions. There was no significant main effect of visual features (V+ vs. V-) ( $p > .05$ ) and no interaction ( $p > .05$ ). Mean response latencies with standard deviation, collapsed across runs, are shown in Table 2.

### Perfusion fMRI Results

At an uncorrected threshold of  $p < .001$  across the whole brain, activation in the bilateral medial temporal lobe (predominantly hippocampus) and the left middle to posterior superior temporal gyrus was observed for the main effect of context (see Table 3 for all regions activated at this threshold). The left superior temporal gyrus and the hippocampus bilaterally reached a corrected level of significance within our a priori defined ROIs (see Figure 2). No significant activation was observed in the left middle temporal gyrus ROI. Nor was significant activation observed for the reverse contrast of heterogeneous  $>$  homogeneous context across the whole brain, or in any of our ROIs.

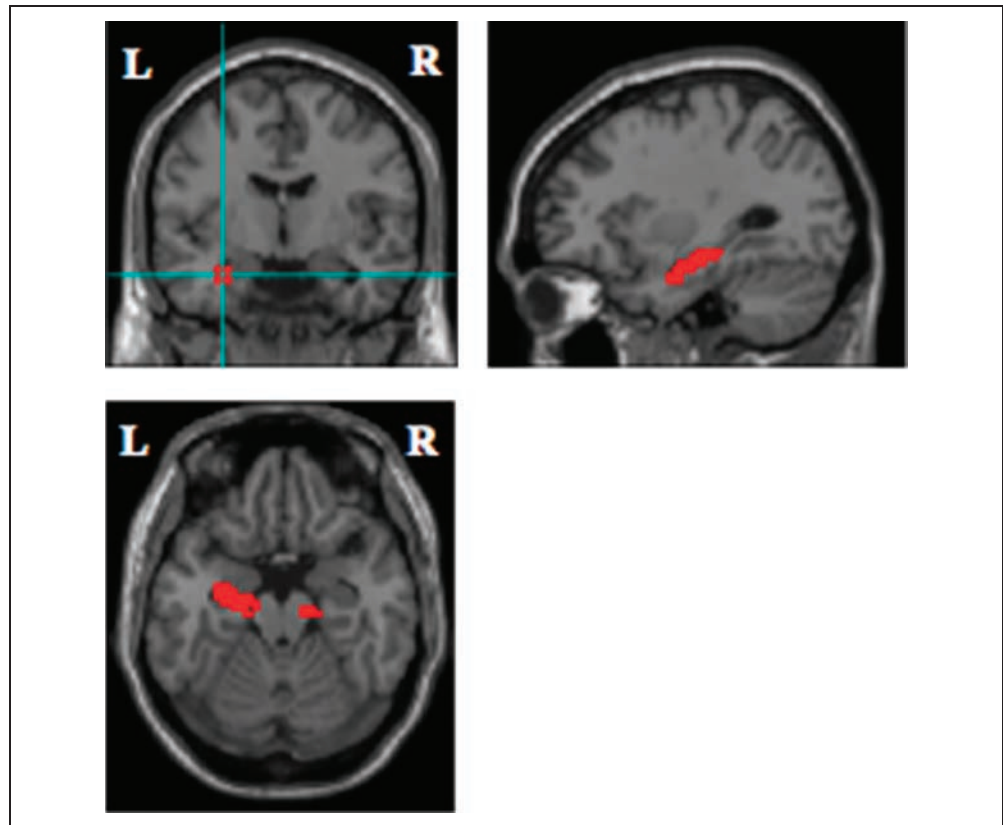
The main effect of visual features independent of context (homogeneous vs. heterogeneous) revealed no

**Table 3.** Regions Engaged for Naming in Homogeneous vs. Heterogeneous Blocks

| <i>Anatomical Region</i>         | <i>x</i>   | <i>y</i>   | <i>z</i>   | <i>Z-score</i> |
|----------------------------------|------------|------------|------------|----------------|
| <b>L Superior temporal gyrus</b> | <b>-66</b> | <b>-30</b> | <b>12</b>  | <b>3.5</b>     |
|                                  | <b>-66</b> | <b>-15</b> | <b>3</b>   | <b>3.4</b>     |
|                                  | <b>-54</b> | <b>-36</b> | <b>15</b>  | <b>3.2</b>     |
| L Superior temporal sulcus       | -42        | -63        | 27         | 3.4            |
| <b>L Hippocampus</b>             | <b>-30</b> | <b>-3</b>  | <b>-30</b> | <b>3.3</b>     |
| L Anterior cerebellum            | -15        | -36        | -27        | 3.4            |
| L Anterior cerebellum            | -9         | -30        | -27        | 3.4            |
| <b>R Hippocampus</b>             | <b>33</b>  | <b>-6</b>  | <b>-33</b> | <b>3.7</b>     |
| R Postcentral gyrus              | 42         | -30        | 45         | 3.7            |
| R Middle cingulate cortex        | 9          | -9         | 45         | 3.5            |
| R Rolandic operculum             | 66         | -12        | 18         | 3.3            |

Table gives coordinates of activation at a threshold of  $p < .001$  (uncorrected, with a minimum cluster of 5 voxels) for naming in homogeneous relative to heterogeneous blocks. Peak coordinates for ROIs in the hippocampus and the left superior temporal gyrus are shown in **bold**. R = right, L = left.

**Figure 3.** Interaction between context and visual features in the left perirhinal cortex. Figure shows activation for the interaction term  $[(\text{HomV+} + \text{HetV-}) > (\text{HetV+} + \text{HomV-})]$ , rendered on an averaged T1 template at  $p < .001$ , uncorrected, with a 5-voxel minimum cluster size. Peak of activation in the perirhinal cortex ROI marked with crosshairs at  $[-30, -6, -24]$  on coronal slice (top left).



significant activation in either the whole brain or ROI analyses. However, we did observe a significant interaction between context and visual features  $[(\text{HomV+} + \text{HetV-}) \times (\text{HomV-} + \text{HetV+})]$  in the anterior medial temporal lobe bilaterally, including the left perirhinal cortex ROI (see Figure 3 and Table 4). To investigate this interaction further, the simple main effects were computed for the effect of context. As shown in Table 4, there was a trend for this interaction to be driven more by the difference between homogeneous and heterogeneous trials

involving objects with similar visual features. No significant effects were found for the opposite interaction.

## DISCUSSION

Consistent with previous behavioral research, blocking items to be named according to semantic category resulted in increased naming latencies compared with blocking items across categories. This semantic context effect was associated with significantly increased perfusion fMRI signal

**Table 4.** Activation for the Interaction between Context and Visual Features and the Corresponding Simple Main Effect of Each Relevant Contrast

| Anatomical Region          | Interaction   |           |            |            | Simple Main Effects           |                               |
|----------------------------|---|-----------|------------|------------|-------------------------------|-------------------------------|
|                            | $(\text{HomV+} + \text{HetV-}) > (\text{HomV-} + \text{HetV+})$ |           |            |            | $\text{HomV+} > \text{HetV+}$ | $\text{HetV-} > \text{HomV-}$ |
|                            | <i>x</i>  | <i>y</i>  | <i>z</i>   | Z-score    | Z-score                       | Z-score                       |
| L Anterior medial temporal | -27   | -21       | -15        | 3.8        | 3.3                           | 2.6                           |
|                            | -18   | -15       | -21        | 3.8        | 3.2                           | 2.7                           |
|                            | <b>-30</b>  | <b>-6</b> | <b>-24</b> | <b>3.6</b> | <b>4.2</b>                    | <b>1.2</b>                    |
| R Anterior medial temporal | 15  | -18       | -18        | 3.8        | 3.0                           | 2.9                           |
| R Cerebellum               | 33  | -66       | -39        | 3.7        | 3.3                           | 2.5                           |
| R Thalamus                 | 24  | -30       | 15         | 3.6        | 2.7                           | 2.9                           |

Table gives coordinates of activation at a threshold of  $p < .001$  (uncorrected, with a minimum cluster of 5 voxels) for the interaction term. Coordinates for activation that reached a corrected level of significance in the left perirhinal cortex ROI are highlighted in **bold**. R = right, L = left.



in the left middle to posterior superior temporal gyrus and bilaterally in the hippocampus. Although naming latencies were not modulated by the number of shared visual features within a block, there was an interaction with the context in which items were presented at the anatomical level, manifesting as perfusion increases in the anterior medial temporal lobe including the left perirhinal cortex.

### **Self-monitoring in Speech Production and the Source of the Context Effect**

We replicated the robust finding of the semantic context effect. Critically, this effect was independent of whether blocks of within category items shared similar visual features. This demonstrates that the increased naming latencies were not the result of processes occurring at a visual–semantic feature level of representation, corroborating and extending earlier work (Damian et al., 2001; Kroll & Stewart, 1994). The effect of blocking items according to category was associated with increased perfusion bilaterally in the hippocampus and in the left middle to posterior superior temporal gyrus.

Only a small number of studies have specifically investigated verbal self-monitoring using functional imaging. These studies have used either EEG/MEG (Ganushchak & Schiller, 2006, 2008; Maess et al., 2002) to investigate the time course and/or signal amplitude associated with verbal self-monitoring, or fMRI and PET through the use of distorted, masked, or delayed verbal feedback (Christoffels, Formisano, & Schiller, 2007; Fu et al., 2006; Hashimoto & Sakai, 2003; McGuire et al., 1996). Several EEG and MEG studies have pointed to a role for the left temporal cortex in self-monitoring, however, they cannot provide specificity in terms of anatomical regions. Previous functional imaging studies, in contrast, have reported activation in a number of more specific brain regions, but have looked at differences between the monitoring of different types of verbal inputs. In contrast, this experiment held object naming constant, in the context of manipulating demands on naming the same stimuli in homogeneous compared with heterogeneous conditions, without manipulating verbal feedback. This allows us to conclude that increased activation in the left middle to posterior superior temporal gyrus is in response to the increased demands on verbal self-monitoring during naming.

A second, related point is that previous studies have reported bilateral superior temporal gyrus activation when manipulating different types of verbal feedback (e.g., Allen et al., 2005; Hirano et al., 1997; McGuire et al., 1996). We did not observe activation in this study in the right superior temporal gyrus. However, as discussed in the Introduction, the manipulation of verbal feedback may introduce attentional confounds due to the qualitative differences between the actual and the distorted, or alien, acoustic verbal signal that is being monitored by the speaker. Here, we controlled attention by en-

suring that verbal output was identical in the homogeneous and heterogeneous conditions, and that verbal monitoring was specific to self-generated overt speech. Under this context, we provide novel evidence that the left superior temporal gyrus is selectively engaged by verbal self-monitoring in overt speech.

A question which the present study leaves open is whether the increased signal observed in the left superior temporal gyrus was driven by an inner or an outer monitoring loop (Levelt, 1983). Indefrey and Levelt (2004) identified the bilateral superior temporal gyri as the possible neural correlates of self-monitoring, with the inner loop residing within the same, or a subset, of regions engaged by the outer monitoring loop. When naming objects, our subjects demonstrated a clear behavioral effect in terms of response latencies induced by the blocking paradigm, yet naming errors were very low (2.8%). This suggests that the mechanism prolonging naming latencies arose prior to and/or during articulation, implicating the inner monitoring loop. In terms of neural operations, the left-lateralized effect observed here is thus more consistent with the prediction that an inner loop engages a subset of (left hemisphere) regions involved in prearticulatory speech monitoring. Bilateral effects reported in previous studies (e.g., Fu et al., 2006; Allen et al., 2005; Hirano et al., 1997; McGuire et al., 1996) might possibly reflect the operation of an outer monitoring loop, which we suggest is engaged after speech (self- or other-generated) is articulated.

Previous behavioral studies have assumed that the semantic context effect is due to either inhibition or excitation of lexico-semantic representations (e.g., Maess et al., 2002; Damian et al., 2001; McCarthy & Kartsounis, 2000). If that were the case, then we would have expected to observe perfusion changes in the middle section of the left middle temporal gyrus, consistent with previous imaging studies investigating PWI and competitor priming effects in object naming (e.g., de Zubicaray et al., 2001, 2006). Instead, the bilateral hippocampal activation is suggestive of a different functional locus for the semantic context effect. The critical involvement of the hippocampal formation in associative encoding of information into memory is well established (for reviews, see Cabeza & Nyberg, 2000; Mayes & Roberts, 2000). The hippocampal activation observed here is therefore consistent with the idea that an episodic memory mechanism underlies the semantic context effect. It supports recent proposals that the effect is due to incremental learning of the associations between semantic features and names (see Oppenheim et al., 2007; Damian & Als, 2005). As Damian and Als (2005) noted, attributing the context effect to an incremental learning mechanism does not necessarily invalidate Maess et al.'s (2002) MEG assessment of the locus and time course of the effect. As Maess et al. acknowledged, their principal component analysis was not able to localize the early and late responses associated

with the context effect to any specific regions of the left temporal cortex. However, visual inspection of the relevant spatio-temporal map indicates a surprisingly basal “hot spot” in addition to a more posterior and superior one in the temporal cortex (see their Figure 4, SF 6).

### Visual Features in Object Naming

Although there was no effect for naming pictures with similar relative to dissimilar visual features at a neuronal level or in the behavioral study, the visual feature manipulation interacted with context. This manifested as increased perfusion signal in the left perirhinal cortex ROI. Convergent evidence from both the human and nonhuman primate literature suggests involvement of the anteromedial temporal cortex in discrimination between visually similar items. We extend these findings by showing that the context in which items are named modulates this activation. Of particular interest is that the interaction was driven by the difference between naming visually similar items in homogeneous versus heterogeneous blocks (see Table 2). This is intriguing when one considers the fact that the pictures in these two types of blocks were identical, hence, subject to the same level of visual discrimination. We therefore propose that the differential activation between homogeneous and heterogeneous blocks of items with similar shared visual features is driven by the differential task demands brought about by manipulation of semantic context. This is consistent with previous studies showing an effect of task on naming in the anteromedial temporal lobes. For example, Tyler et al. (2004) reported increased activation in the left perirhinal cortex, extend-

ing into the fusiform gyrus, the amygdala, and the hippocampus, for a task comparing naming pictures at a basic level (e.g., dog) relative to naming the same pictures at a domain level (e.g., animal). This was interpreted as being due to the finer-grained discrimination required to differentiate between conceptually similar objects. This result was extended by Moss et al. (2005), who observed an increase in the same region for living things—concepts that share a greater proportion of highly correlated shared properties—relative to more distinctive nonliving things. Interestingly, the anatomical effect that they reported was not reflected by a difference in naming latencies, consistent with the results from the behavioral study reported here for similar versus dissimilar visual features, although as a note of caution, this must be interpreted in the context of an absence of naming latency data from the present imaging study.

In summary, we report here three key findings: Firstly, we observed, as predicted, that the left middle to posterior superior temporal gyrus is engaged by verbal self-monitoring when objects to be named are blocked according to semantic category. Second, the hippocampal activation supports a recent proposal that the context effect is due to an incremental learning mechanism. Third, we observed an anatomical effect for shared visual features in the perirhinal cortex, reflecting the increased demands placed on differentiating between conceptually similar objects. Interestingly, this was in the absence of a corresponding behavioral effect, showing that there is no cost involved in terms of naming latencies despite the increased neuronal activity. Taken together, these results also demonstrate the importance of ASL as a successful technique for fMRI of overt speech tasks.

## APPENDIX

### Homogeneous Blocks in Rows and Heterogeneous Blocks in Columns

|                                      | <i>Heterogeneous Blocks</i> |          |            |          |            |
|--------------------------------------|-----------------------------|----------|------------|----------|------------|
| <i>a. Similar Visual Features</i>    |                             |          |            |          |            |
| Homogeneous blocks                   | nightgown                   | bra      | cape       | skirt    | dress      |
|                                      | hamster                     | elephant | dog        | cat      | donkey     |
|                                      | plum                        | lemon    | strawberry | melon    | grapefruit |
|                                      | zucchini                    | rhubarb  | pumpkin    | radish   | cabbage    |
|                                      | airplane                    | van      | rocket     | tractor  | train      |
| <i>b. Dissimilar Visual Features</i> |                             |          |            |          |            |
| Homogeneous blocks                   | cap                         | slippers | tie        | swimsuit | earmuffs   |
|                                      | rabbit                      | lamb     | whale      | giraffe  | chicken    |
|                                      | pear                        | raisin   | pineapple  | banana   | avocado    |
|                                      | spinach                     | turnip   | asparagus  | beans    | eggplant   |
|                                      | helicopter                  | balloon  | boat       | unicycle | car        |

## Acknowledgments

We thank all the subjects who participated in these studies. We would also like to thank Matt Meredith for his assistance acquiring the data.

Reprint requests should be sent to Julia Hocking, Centre for Magnetic Resonance, The University of Queensland, St Lucia, Queensland, Australia, 4072, or via e-mail: julia.hocking@cmr.uq.edu.au.

## Note

1. It should be noted that although other items in that block also have legs (i.e., giraffe, lamb, and rabbit; see Appendix for stimulus blocks), none of the other items have beaks, feathers, or wings. Indeed, McRae et al. (2005) reported that raters tended to focus on the more salient, or distinctive, visual properties of objects, a point that emphasizes how alike the blocks of visually similar items are. McRae et al. include a discussion of the ambiguities inherent in the “has\_legs” feature (e.g., a table also “has\_legs”; p. 551).

## REFERENCES

- Aguirre, G. K., Detre, J. A., Zarahn, E., & Alsop, D. C. (2002). Experimental design and the relative sensitivity of BOLD and perfusion fMRI. *Neuroimage*, *15*, 488–500.
- Allen, P. P., Amaro, E., Fu, C. H., Williams, S. C., Brammer, M., Johns, L. C., et al. (2005). Neural correlates of the misattribution of self-generated speech. *Human Brain Mapping*, *26*, 44–53.
- Alsop, D. C., & Detre, J. A. (1996). Reduced transit-time sensitivity in noninvasive magnetic resonance imaging of human cerebral blood flow. *Journal of Cerebral Blood Flow and Metabolism*, *16*, 1236–1249.
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database* [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.
- Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., et al. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*, 445–459.
- Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *Quarterly Journal of Experimental Psychology A*, *58*, 667–692.
- Binder, J. R., Rao, S. M., Hammeke, T. A., Frost, J. A., Bandettini, P. A., Jesmanowicz, A., et al. (1995). Lateralized human brain language systems demonstrated by task subtraction functional magnetic resonance imaging. *Archives of Neurology*, *52*, 593–601.
- Buffalo, E. A., Reber, P. J., & Squire, L. R. (1998). The human perirhinal cortex and recognition memory. *Hippocampus*, *8*, 330–339.
- Bussey, T. J., Saksida, L. M., & Murray, E. A. (2003). Impairments in visual discrimination after perirhinal cortex lesions: Testing “declarative” vs. “perceptual-mnemonic” views of perirhinal cortex function. *European Journal of Neuroscience*, *17*, 649–660.
- Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, *12*, 1–47.
- Christoffels, I. K., Formisano, E., & Schiller, N. O. (2007). Neural correlates of verbal feedback processing: An fMRI study employing overt speech. *Human Brain Mapping*, *28*, 868–879.
- Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, *132*, 163–201.
- Cusack, R., Cumming, N., Bor, D., Norris, D., & Lyzenga, J. (2005). Automated post-hoc noise cancellation tool for audio recordings acquired in an MRI scanner. *Human Brain Mapping*, *24*, 299–304.
- Damian, M. F., & Als, L. C. (2005). Long-lasting semantic context effects in the spoken production of object names. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1372–1384.
- Damian, M. F., Vigliocco, G., & Levelt, W. J. (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, *81*, B77–B86.
- de Zubicaray, G. I., McMahon, K. L., Eastburn, M., & Pringle, A. (2006). Top-down influences on lexical selection during spoken word production: A 4T fMRI investigation of refractory effects in picture naming. *Human Brain Mapping*, *27*, 864–873.
- de Zubicaray, G. I., Wilson, S. J., McMahon, K. L., & Muthiah, S. (2001). The semantic interference effect in the picture-word paradigm: An event-related fMRI study employing overt responses. *Human Brain Mapping*, *14*, 218–227.
- Demonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., et al. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, *115*, 1753–1768.
- Devlin, J. T., Moore, C. J., Mummery, C. J., Gorno-Tempini, M. L., Phillips, J. A., Noppeney, U., et al. (2002). Anatomic constraints on cognitive theories of category specificity. *Neuroimage*, *15*, 675–685.
- Devlin, J. T., & Price, C. J. (2007). Perirhinal contributions to human visual perception. *Current Biology*, *17*, 1484–1488.
- Freire, L., Roche, A., & Mangin, J. F. (2002). What is the best similarity measure for motion correction in fMRI time series? *IEEE Transactions on Medical Imaging*, *21*, 470–484.
- Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S., & Turner, R. (1996). Movement-related effects in fMRI time-series. *Magnetic Resonance in Medicine*, *35*, 346–355.
- Fu, C. H., Vythelingum, G. N., Brammer, M. J., Williams, S. C., Amaro, E., Jr., Andrew, C. M., et al. (2006). An fMRI study of verbal self-monitoring: Neural correlates of auditory verbal feedback. *Cerebral Cortex*, *16*, 969–977.
- Gainotti, G. (2000). What the locus of brain lesion tells us about the nature of the cognitive defect underlying category-specific disorders: A review. *Cortex*, *36*, 539–559.
- Gainotti, G., Silveri, M. C., Daniele, A., & Giustolisi, L. (1995). Neuroanatomical correlates of category-specific semantic disorders: A critical survey. *Memory*, *3*, 247–264.
- Ganushchak, L. Y., & Schiller, N. O. (2006). Effects of time pressure on verbal self-monitoring: An ERP study. *Brain Research*, *1125*, 104–115.
- Ganushchak, L. Y., & Schiller, N. O. (2008). Motivation and semantic context affect brain error-monitoring activity: An event-related brain potentials study. *Neuroimage*, *39*, 395–405.
- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, *24*, 187–203.

- Hajnal, J. V., Myers, R., Oatridge, A., Schwieso, J. E., Young, I. R., & Bydder, G. M. (1994). Artifacts due to stimulus correlated motion in functional imaging of the brain. *Magnetic Resonance in Medicine*, *31*, 283–291.
- Harley, T. A., & MacAndrew, S. B. (2001). Constraints upon word substitution speech errors. *Journal of Psycholinguistic Research*, *30*, 395–418.
- Hashimoto, Y., & Sakai, K. L. (2003). Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: An fMRI study. *Human Brain Mapping*, *20*, 22–28.
- Hirano, S., Kojima, H., Naito, Y., Honjo, I., Kamoto, Y., Okazawa, H., et al. (1997). Cortical processing mechanism for vocalization with auditory verbal feedback. *NeuroReport*, *8*, 2379–2382.
- Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, *100*, 464–482.
- Huang, J., Francis, A. P., & Carr, T. H. (2008). Studying overt word reading and speech production with event-related fMRI: A method for detecting, assessing, and correcting articulation-induced signal changes and for measuring onset time and duration of articulation. *Brain and Language*, *104*, 10–23.
- Humphreys, G. W., Price, C. J., & Riddoch, M. J. (1999). From objects to names: A cognitive neuroscience approach. *Psychological Research*, *62*, 118–130.
- Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*, 101–144.
- Kemeny, S., Ye, F. Q., Birn, R., & Braun, A. R. (2005). Comparison of continuous overt speech fMRI using BOLD and arterial spin labeling. *Human Brain Mapping*, *24*, 173–183.
- Kim, S. G. (1995). Quantification of relative cerebral blood flow change by flow-sensitive alternating inversion recovery (FAIR) technique: Application to functional mapping. *Magnetic Resonance in Medicine*, *34*, 293–301.
- Kroll, J., & Stewart, F. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–174.
- Lerner, Y., Hendler, T., Ben-Bashat, D., Harel, M., & Malach, R. (2001). A hierarchical axis of object processing stages in the human visual cortex. *Cerebral Cortex*, *11*, 287–297.
- Levelt, W. J. (1983). Monitoring and self-repair in speech. *Cognition*, *14*, 41–104.
- Levelt, W. J. (1989). *Speaking: From intention to articulation*. Cambridge: MIT Press.
- Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38; discussion 38–75.
- Maess, B., Friederici, A. D., Damian, M., Meyer, A. S., & Levelt, W. J. (2002). Semantic category interference in overt picture naming: Sharpening current density localization by PCA. *Journal of Cognitive Neuroscience*, *14*, 455–462.
- Mayes, A. R., & Roberts, N. (2000). Theories of episodic memory. *Philosophical Transactions of the Royal Society of London*, *B356*, 1395–1408.
- McCarthy, R. A., & Kartsounis, L. D. (2000). Wobbly words: Refractory anomia with preserved semantics. *Neurocase*, *6*, 487–497.
- McGuire, P. K., Silbersweig, D. A., & Frith, C. D. (1996). Functional neuroanatomy of verbal self-monitoring. *Brain*, *119*, 907–917.
- McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*, 547–559.
- Meunier, M., Bachevalier, J., Mishkin, M., & Murray, E. A. (1993). Effects on visual recognition of combined and separate ablations of the entorhinal and perirhinal cortex in rhesus monkeys. *Journal of Neuroscience*, *13*, 5418–5432.
- Moss, H. E., Rodd, J. M., Stamatakis, E. A., Bright, P., & Tyler, L. K. (2005). Anteromedial temporal cortex supports fine-grained differentiation among objects. *Cerebral Cortex*, *15*, 616–627.
- Murray, E. A., & Gaffan, D. (1994). Removal of the amygdala plus subjacent cortex disrupts the retention of both intramodal and crossmodal associative memories in monkeys. *Behavioural Neuroscience*, *108*, 494–500.
- Nelles, J. L., Lugar, H. M., Coalson, R. S., Miezin, F. M., Petersen, S. E., & Schlagar, B. L. (2003). Automated method for extracting response latencies of subject vocalizations in event-related fMRI experiments. *Neuroimage*, *20*, 1865–1871.
- Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2007). Cumulative semantic interference as learning. *Brain and Language*, *103*, 175–176.
- Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition*, *77*, 97–132.
- Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*, 2400–2406.
- Shattuck, D. W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K. L., et al. (2008). Construction of a 3D probabilistic atlas of human cortical structures. *Neuroimage*, *39*, 1064–1080.
- Slevc, L. R., & Ferreira, V. S. (2006). Halting in single word production: A test of the perceptual loop theory of speech monitoring. *Journal of Memory and Language*, *54*, 515–540.
- Troiani, V., Fernandez-Seara, M. A., Wang, Z., Detre, J. A., Ash, S., & Grossman, M. (2007). Narrative speech production: An fMRI study using continuous arterial spin labeling. *Neuroimage*, doi: 10.1016/j.neuroimage.2007.12.002.
- Tyler, L. K., Stamatakis, E. A., Bright, P., Acres, K., Abdallah, S., Rodd, J. M., et al. (2004). Processing objects at different levels of specificity. *Journal of Cognitive Neuroscience*, *16*, 351–362.
- Vaughan, J. T., Adriany, G., Garwood, M., Yacoub, E., Duong, T., DelaBarre, L., et al. (2002). Detunable transverse electromagnetic (TEM) volume coil for high-field NMR. *Magnetic Resonance in Medicine*, *47*, 990–1000.
- Vigliocco, G., Lauer, M., Damian, M. F., & Levelt, W. J. (2002). Semantic and syntactic forces in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 46–58.
- Vitkovitch, M., Humphreys, G. W., & Lloyd-Jones, T. (1993). On naming a giraffe a zebra: Picture naming errors across different object categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 243–259.
- Vitkovitch, M., Rutter, C., & Read, A. (2001). Inhibitory effects during object name retrieval: The effect of interval between prime and target on picture naming responses. *British Journal of Psychology*, *92*, 483–506.
- Wang, J., Alsop, D. C., Li, L., Listerud, J., Gonzalez-At, J. B., Schnall, M. D., et al. (2002). Comparison of quantitative perfusion imaging using arterial spin labeling at 1.5 and 4.0 Tesla. *Magnetic Resonance in Medicine*, *48*, 242–254.

- Wang, J., Licht, D. J., Jahng, G. H., Liu, C. S., Rubin, J. T., Haselgrove, J., et al. (2003). Pediatric perfusion imaging using pulsed arterial spin labeling. *Journal of Magnetic Resonance Imaging*, *18*, 404–413.
- Wong, E. C., Buxton, R. B., & Frank, L. R. (1997). Implementation of quantitative perfusion imaging techniques for functional brain mapping using pulsed arterial spin labeling. *NMR in Biomedicine*, *10*, 237–249.
- Wong, E. C., Buxton, R. B., & Frank, L. R. (1998a). Quantitative imaging of perfusion using a single subtraction (QUIPSS and QUIPSS II). *Magnetic Resonance in Medicine*, *39*, 702–708.
- Wong, E. C., Buxton, R. B., & Frank, L. R. (1998b). A theoretical and experimental comparison of continuous and pulsed arterial spin labeling techniques for quantitative perfusion imaging. *Magnetic Resonance in Medicine*, *40*, 348–355.
- Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, *256*, 846–849.
- Zola-Morgan, S., Squire, L. R., Amaral, D. G., & Suzuki, W. A. (1989). Lesions of perirhinal and parahippocampal cortex that spare the amygdala and hippocampal formation produce severe memory impairment. *Journal of Neuroscience*, *9*, 4355–4370.