

Brain areas engaged during visual judgments by involuntary access to novel semantic information

Thomas W. James^{*}, Isabel Gauthier

Department of Psychology, Vanderbilt University, 111 21st Ave. S., Nashville, TN 37203, USA

Received 23 May 2003; received in revised form 30 September 2003

Abstract

Theories of visual recognition place different emphasis on the role of non-stimulus factors. Previously, we showed that arbitrary semantic associations influenced visual recognition of novel objects. Here, the neural substrate of this effect was investigated. During a visual task, novel objects associated with arbitrary semantic features produced more activation in frontal and parietal cortex than objects associated with names. Because the task required no semantic retrieval, access to semantics appears to be involuntary. The brain regions involved have been implicated in semantic processing, thus recently acquired semantics activate a similar network to semantics learned over a lifetime.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Object recognition; fMRI; Neuroimaging; Implicit; Semantic

1. Introduction

It is widely known that most of the brain is involved in the visual recognition of objects (for review, see Aguirre & Farah, 1998; Grill-Spector, Kourtzi, & Kanwisher, 2001). Theoretically, this widespread activation during object recognition represents a combination of visual (perceptual) and non-visual (semantic) processes that are involved in recognizing common objects. In other words, the pattern of activation produced by viewing an object is a combination of activation from direct visual stimulation and from associations developed during previous experience. This idea is described well by Forde and Humphreys' (Humphreys & Forde, 2001) hierarchical interactive theory (HIT). According to HIT, visual objects are processed in a series of interactive stages, including a structural description stage and a semantic stage. A common practice for investigating brain regions involved in object recognition is to compare viewing of intact images of objects with scrambled images (Grill-Spector et al., 2001; Kanwisher, Chun, McDermott, & Ledden, 1996; Malach et al., 1995). Scrambling the images destroys the structural

information in the image and a comparison with intact images isolates brain regions that are sensitive to object structure. Scrambling the images also disrupts access to semantic knowledge about the object represented by the image. Therefore, not only are brain regions isolated that are sensitive to object structure, but also regions that are activated by these semantic associations. These two types of activation, direct activation from visual stimulation and indirect activation from semantic associations, can be separately isolated.

One method of dissociating activation related to direct visual stimulation and semantic associations is to compare the activation produced by common as opposed to novel objects. With this method, brain regions that are activated by common objects more than by novel objects are assumed to process semantic information associated with the common objects and not available for the novel objects. Two studies that used this method (Martin, 1999; Vuilleumier, Henson, Driver, & Dolan, 2002) revealed a common focus of activation in the inferior prefrontal cortex, suggesting that this region is involved in processing the semantic information as opposed to visual information. One limitation of this method, however, is that novel objects may produce different patterns of activation because they are less familiar visually, and not only due to a lack of semantic associations with them. This problem was

^{*} Corresponding author. Tel.: +1-615-322-6386; fax: +1-615-322-4706.

E-mail address: thomas.james@vanderbilt.edu (T.W. James).

addressed in one study (Leveroni et al., 2000) in which participants were familiarized with novel faces before testing. Activation with these newly learned faces was compared to activation with famous faces. Similar to the other two studies (Martin, 1999; Vuilleumier et al., 2002), Leveroni et al. (2000) found greater activation in the inferior frontal cortex. These findings suggest that the inferior frontal cortex provides a significant contribution to the processing of semantic information associated with common objects.

Another method of dissociating activation related to direct visual stimulation and semantic associations is to compare the effects of perceptual and conceptual priming. In neuroimaging, priming is usually defined as a decrease in activation with a stimulus that is brought about by previous exposure to that same stimulus (repetition priming) or to a similar stimulus (for review, see Schacter & Buckner, 1998; Wiggs & Martin, 1998). Perceptual priming effects arise when the manipulation of the stimulus between study and test is restricted to the perceptual attributes of the stimulus, for instance, a chair seen from two different viewpoints. Conceptual priming takes place between stimuli that are related conceptually. For instance, chairs, couches and stools are perceptually distinct, but are conceptually related because they are all used for sitting. Examining two studies that compared conceptual and perceptual priming (Koutstaal et al., 2001; Vuilleumier et al., 2002) revealed a common site of activation in the inferior frontal cortex that was affected by conceptual priming, but not perceptual priming. One limitation of this method is that conceptually similar items are often perceptually similar (e.g., a stool is more similar to a chair than to a lamp) and consequently, a conceptual priming effect may actually reflect a combination of conceptual and perceptual factors. Despite this limitation, these findings converge with those of the studies comparing common and novel objects (Leveroni et al., 2000; Martin, 1999; Vuilleumier et al., 2002) to suggest that an important neural processing site for semantic associations exists in the inferior frontal cortex. Also apparent from these studies is that many regions in addition to the inferior frontal cortex are recruited for semantic processing for a particular combination of stimuli and tasks. In fact, for a particular combination of stimuli and task, the contributing regions are generally quite widespread throughout the brain.

The purpose of the present experiment was to determine which neural substrates underlie the activation of recent associations of semantic information with visual structural features of objects. Previously, this question has been investigated by comparing common objects with novel objects. This method, however, has at least two limitations: first, in addition to having associated semantic information, common objects are also more perceptually familiar than novel objects; second, the

semantic information associated with common objects tends to be related systematically to the visual features of those objects. In the present experiment, we circumvented these problems by training participants to associate clusters of semantic features (arbitrary concepts) with novel objects. This procedure ensures the independence of visual features and semantic features. In a previous study, we demonstrated that this type of semantic training was sufficient to produce changes in behavioral performance on a visual task (Gauthier, James, Curby, & Tarr, 2003). In that study, participants associated arbitrary concepts, which were simply clusters of three semantic features (e.g., fast, friendly, heavy), with each of four novel objects. Participants were divided into two groups; the dissimilar concept group associated dissimilar arbitrary concepts with each of the four objects (e.g., fast, friendly, heavy versus loud, nervous, flexible); the similar concept group associated similar arbitrary concepts with each of the four objects (e.g., fast, friendly, heavy versus fast, friendly, loud). To be precise, any pair of dissimilar arbitrary concepts shared no common features, whereas any pair of similar arbitrary concepts shared exactly two out of three semantic features. Despite equivalent exposure to the objects during training, participants in the dissimilar group were faster to respond in a sequential matching task with the trained objects than their counterparts in the similar group. Finding this relative difference between the groups suggested that limited semantic training (under one hour) had a reliable effect on visual performance. Furthermore, this effect was found using a visual matching task, a task that requires no explicit retrieval of the learned semantic features, and was also found under conditions of verbal interference, suggesting that access to the semantic associations may be automatic.

The present experiment made use of a condition that was similar to the dissimilar arbitrary concept condition from the previous study (Gauthier et al., 2003). For this semantic (SEM) condition three different semantic features and a name were associated with each of four visually similar novel objects. The SEM condition was contrasted with a control condition (NAM) for which participants learned to associate names, but no semantic features, with a different set of four novel objects. In this way, the SEM and NAM objects would be equally familiar and would have names associated with them, but only the SEM objects would have significant semantic associations. The names for the SEM and NAM conditions were all proper names, which appear to be distinct from the common names of objects (for review, see Semenza, 1997). This seems particularly true in terms of their semantic associations; common names inherit the associations of their object counterparts, whereas proper names, because of the arbitrary manner in which they are assigned to individuals, are unlikely to maintain a

consistent set of semantic associations. Another control condition was a non-trained set of objects (NON); no training occurred for these objects and therefore the only exposure was during testing phases of the experiment. All three object sets were similar in regard to their visual features and were fully counterbalanced across participants with the three training conditions. Therefore, only the type of training differed between the sets.

Previous neuroimaging studies investigating the brain regions involved in processing semantic information associated with objects have, for good reason, relied primarily on semantic tasks (for review, see Bookheimer, 2002; Cabeza & Nyberg, 2000; Murtha, Chertkow, Beauregard, & Evans, 1999). Findings from Gauthier et al. (2003) indicate that semantic associations can influence visual judgments during tasks that do not require explicit semantic retrieval. Therefore, we hypothesized that brain regions involved in the processing of semantic associations with objects would be engaged automatically during visual processing of those objects. To this end, we used a simultaneous matching task; this task has only a slight memory requirement and certainly does not require explicit retrieval of semantic associations. We also hypothesized, based on our review of the literature, that regions of the inferior frontal cortex would be recruited for processing of the SEM objects more than for the NAM objects. Support for this last hypothesis would suggest that recently created associations activate a similar neural network that is engaged by associations that are created over a lifetime of experience.

2. Materials and methods

2.1. Participants

Participants were graduate students, postdoctoral fellows or research assistants in the psychology department at Vanderbilt University. All participants reported that they were right-handed, had normal or corrected to normal vision, and had no history of neurological disorders. There were five females and seven males, with ages ranging from 22 to 42 years with a mean age of 28.3. The protocol was approved by the Vanderbilt University Institutional Review Board.

2.2. Stimuli and procedures

The experiment was conducted in four phases; a training phase that was preceded and followed by behavioral testing phases, which were all followed by the Neuroimaging phase. Behavioral testing was done before and after training to measure any changes in behavioral performance that may have resulted from the training. Participants were trained and tested with 12 highly similar novel objects (YUFOs; Fig. 1; images

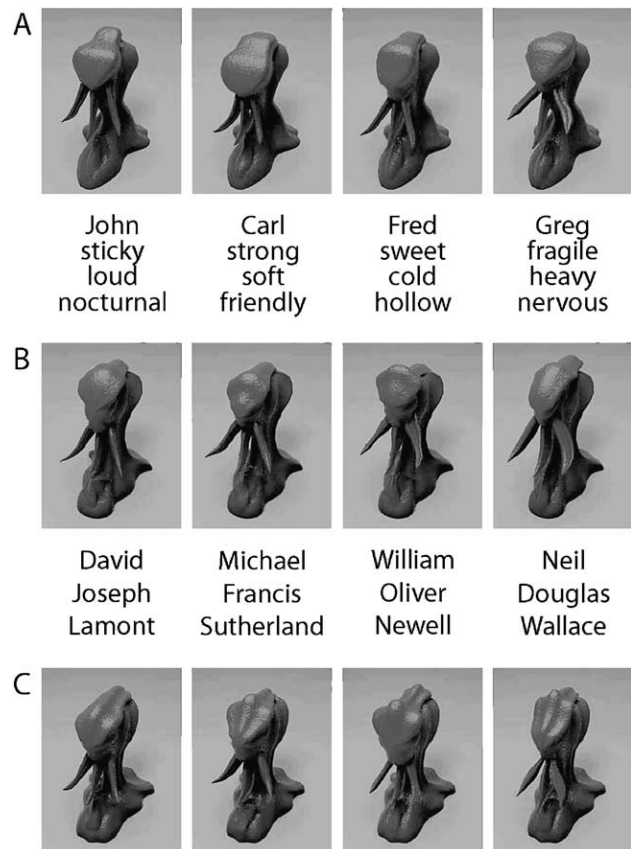


Fig. 1. Novel objects, semantic features and proper names. An example of a possible combination of training conditions and object sets. (A) The semantic condition, in which nicknames and semantic features are learned. (B) The name-only condition, in which proper names are learned. (C) The non-trained condition. All possible combinations of object sets and training conditions were used across participants. Names and semantic features were randomly assigned for each participant.

provided courtesy of Michael J. Tarr, Brown University, Providence, RI). Some changes were made to the stimuli and procedures after half ($N = 6$) of the 12 total participants were tested. For Group 1 ($N = 6$), the three object sets were presented in the same blue color; for Group 2 ($N = 6$), the three object sets were presented in three different colors, red, yellow and cyan, to make the different sets more distinctive. Semantic features were identical for both groups and were chosen to be non-visual in nature and to be unlikely to arouse vivid visual imagery (Fig. 1). For Group 1, names were all four-letter adult men's nicknames (e.g., Mike); for Group 2, full names were used that included first, middle and surname (e.g., Michael Francis Sutherland). Changing the format of the names was done to equate the amount of verbal material that the participants had to learn for the SEM and NAM conditions. The two groups were first analyzed separately, but after finding significant relative overlap between the two groups in their patterns of activation, particularly in hypothesized regions of

interest, the groups were collapsed and analyzed together (see Section 3).

All testing and training was conducted using Macintosh computers and RSVP software (www.cog.brown.edu/~tarr/RSVP). During training outside of the scanner, the object and word stimuli were presented on an iMac computer screen that was positioned approximately 57 cm from the participant. The dimensions of the objects varied slightly, but a representative object was 5.6 cm wide and 7.6 cm tall and thus subtended approximately $5.6^\circ \times 7.6^\circ$ of visual angle. During the imaging phase, the objects were presented on two small LCD screens mounted within a Visuastim XGA goggle system (MRI Devices Inc., <http://www.mrvideo.com>) worn by the participant. The virtual sizes of the screens were 76.2×57.2 cm and they were viewed at a virtual distance of 120 cm. Thus, the virtual images of the objects appeared approximately 24.8×33.7 cm in size and subtended approximately $11.7^\circ \times 15.7^\circ$ of visual angle. For all training and testing, the objects were presented at the orientation shown in Fig. 1.

2.3. Pre- and post-training test phases

Before the training phase, participants performed a matching task with the objects to provide a baseline (non-trained performance). Each trial began with the presentation of a central fixation cross for 750 ms, followed by the simultaneous presentation of two objects, separated by 8.7° of visual angle from center to center, for 1200 ms, followed by a fixation cross. Participants responded with a button press indicating whether the objects were the same or different. Objects were divided into three sets of four objects (Fig. 1) for the purposes of counterbalancing. Pairs of objects were chosen from within a subset of four objects, not between subsets. There were 288 trials in total and they were presented in completely randomized order. Participants were given a break every 48 trials. Trials timed out after 6 s had elapsed. The post-training test phase was identical to the pre-training test phase, but occurred after the training phase.

2.4. Training phase

During the training phase, participants learned to associate some of the objects with semantic features and/or names. The set of 12 objects was divided into three subsets of four objects each (Fig. 1). Each participant was trained with two of these three object sets, leaving the third set as a non-trained control set. With the first trained set, participants learned to associate a nickname and three semantic features with each object (e.g., Fig. 1). With the second trained set, participants learned to associate only a nickname or full name with each object (Fig. 1). The assignment of object sets to the SEM,

Table 1
Names used in the NAM and SEM conditions

Nickname	Full name
John	Jonathan Wesley Abraham
Carl	Carlos Richard Baird
Fred	Frederick Kyle Drummond
Greg	Gregory Childress Forsythe
Bill	William Oliver Newell
Dave	David Joseph Lamont
Neil	Neil Douglas Wallace
Mike	Michael Francis Sutherland

NAM or NON conditions was counterbalanced across participants. Nicknames (Table 1) were associated with the SEM objects by participants in both Groups 1 and 2. For the NAM condition, participants in Group 2 associated full names (first, middle and surname) with the objects, while participants in Group 1 associated nicknames like those used in the SEM condition. Full names were used with Group 2 to help equate the SEM and NAM conditions on the amount of verbal information that was learned, but seemed to make no difference in the results. The training was accomplished using three different types of trials. For “Show” trials, an object was presented with two or three adjectives and/or a name for 5 s. For instance, an object might be presented together with the words, “This one is: MIKE”, “FRIENDLY and LOUD”. The participants were not required to make a response for Show trials. For “Verification” trials, an object was shown together with either one or two adjectives or a name, which remained on the screen until the participants made a response. For instance, an object might be presented with the words, “Is this one: MIKE”, or “Is this one: LOUD and SOFT”. Participants responded to Verification trials by pressing one of two buttons, indicating whether the adjective(s) or name matched the object or not. For “Criterion” trials, three objects were presented together with either three adjectives or a name, which remained on the screen until the participants made a response. For instance, an object might be presented with the words “Which one is: FRIENDLY LOUD and SWEET” or “Which one is: MIKE”. Participants responded during the Criterion trials by pressing one of three buttons, indicating which of the three objects matched the adjectives or name. The buttons were mapped onto the numbers 1, 2 and 3, which were presented below the three objects. For the Show trials and Verification trials, the number of trials was fixed. For the Criterion trials, participants in Group 1 were required to reach a criterion of 22 correct responses per block of 24 trials, while responding to each trial in under 3000 ms. The speed criterion was changed to 4000 ms for Group 2, because one participant in Group 1 required an exceedingly high number of blocks of trials to reach the 3000 ms criterion, which was deemed too conservative. Regardless of their perfor-

mance on the first block of Criterion trials, participants were required to perform the second block, thus the minimum number of Criterion trials that a participant performed was 48. Participants that performed a maximum of 480 Criterion trials were excluded from further testing, however this did not occur for any participants tested. For Practice and Criterion trials, auditory feedback (a beep) was given to participants when an incorrect response was made.

The presentation order of these trials was as follows: Show trials with two features plus a name for each object in the SEM were followed by Verification trials with these same features. Show trials were then performed with three features plus a name, followed by practice trials. Show and Verification trials for the NAM condition were interleaved with the trials for the SEM condition. Even when there was less information to learn in the NAM condition than the SEM condition (nicknames only), the number of trials was kept constant across the SEM and NAM conditions. Finally, Verification trials with the SEM and NAM objects combined were followed by the Criterion trials. Participants performed 20 Show trials and 248 Verification trials for each of the SEM and NAM conditions. There were 220 combined trials, split evenly between SEM and NAM objects. Equal numbers of trials for all trial types ensured that exposure to the SEM and NAM objects was a constant. Use of the Criterion trials demonstrated that, although the NAM associations may have been easier to learn, retrieval of SEM and NAM associations was equated by the end of training. More detailed information about the sequence and number of these trials can be obtained upon request from the corresponding author.

2.5. Neuroimaging phase

In the Neuroimaging phase, objects from all three object sets, SEM, NAM and NON were presented using a blocked stimulus presentation design. Participants performed a simultaneous match task on pairs of objects selected from the same object set. There were 10 match trials per block, with all pairs in a block selected from the same object set. Therefore, there were SEM blocks, NAM blocks and NON blocks, which could be contrasted. The order of these blocks was counterbalanced across runs and across subjects. The timing parameters of the runs were different for the two groups of six participants. Group 1 performed four runs that each began and ended with fixation. Each stimulus block was 20 s in length and block was followed by a 6 s period of fixation. Each of the 10 trials within a block began with the presentation of a fixation cross for 500 ms, followed by the simultaneous presentation of two objects, side by side, for 1200 ms, followed by a fixation cross for 300 ms. After analyzing the data from Group 1, it was de-

termined that these timing parameters were not optimal. With only a 6 s fixation period after each 20 s block, the BOLD signal did not have time to return to baseline. Therefore, the timing parameters were changed for Group 2. These changes in timing were not expected to influence participants' performance, nor was it expected to produce qualitative changes in the data; the changes were implemented to increase the power of our analyses. Group 2 performed six runs that each began and ended with fixation. Each stimulus block was 12.8 s and was followed by a 7.2 s period of fixation. Each of the eight trials in a block began with the presentation of a fixation cross for 100 ms, followed by the simultaneous presentation of two objects, side by side, for 1200 ms, followed by a fixation cross for 300 ms. Both groups responded on each trial with a button press indicating whether the two objects were the same or different. The ordering of the stimulus presentation blocks was counterbalanced across runs.

2.6. Imaging parameters and analysis

All imaging was done using a 3-T, whole body GE MRI system and a birdcage head coil located at the Vanderbilt University Medical Center (Nashville, USA). The field of view was $24 \times 24 \times 14.0$ cm (or $24 \times 24 \times 12.6$), with an in-plane resolution of 64×64 pixels and 20 (18) contiguous oblique axial scan planes per volume (whole brain), resulting in a voxel size of $3.75 \times 3.75 \times 7.0$ mm. Images were collected using a T2*-weighted EPI acquisition (TE = 25 ms, TR = 2000 ms, flip angle = 70°) for blood oxygen-level dependent (BOLD) based imaging. High-resolution T1-weighted anatomical volumes were also acquired using a 3-D fast spoiled grass (FSPGR) acquisition (TI = 400 ms, TE = 4.18 ms, TR = 10 ms, FA = 20°).

The imaging data were pre-processed using the Brain Voyager™ 3-D analysis tools. The anatomical volumes were transformed into a stereotactic space that was common for all participants (Talairach & Tournoux, 1988). Functional volumes for each subject were aligned to the transformed anatomical volumes, thereby transforming the functional data into a common brain space across participants. Functional data underwent 3-D motion correction, 3-D spatial gaussian filtering (FWHM 6 mm), temporal gaussian filtering (FWHM 2 s), and linear trend removal.

The imaging data were analyzed using the Brain Voyager™ multi-study general linear model (GLM) procedure. This procedure allows the correlation of predictor variables or functions with the recorded activation data (criterion variables) across scanning sessions. Data were analyzed separately for each group ($N = 6$) of participants and on the combined sample ($N = 12$). The predictor functions were a series of gamma functions ($\Delta = 2.5$, $\tau = 1.25$) spaced in time

based on the blocked stimulus presentation paradigm of the particular run being analyzed (Boynton, Engel, Glover, & Heeger, 1996). In other words, the predictors represented the stimulus protocol boxcar functions convolved with the appropriate gamma function.

3. Results

3.1. Training data

All 12 participants passed the criterion test, in which they were required to achieve 22 correct responses in a block of 24 trials, with all trials being completed in under 3000 ms (Group 1) or 4000 ms (Group 2). The median number of blocks of trials required to reach criterion was four (96 trials); the median number of blocks for Group 1 was four and for Group 2 was five (120 trials). At the end of training, there was a small, but significant difference in accuracy during Verification trials between the SEM and NAM conditions, with higher accuracy for the NAM objects than for SEM objects ($t_{(11)} = 2.53, p < 0.05$). This difference was driven mainly by a reliable accuracy difference (6%) between the SEM and NAM conditions ($t_{(5)} = 2.43, p < 0.05$) for participants in Group 1 (which learned only nicknames for the NAM condition). A non-significant effect in the same direction for Group 2 was almost completely driven by a large difference (25%) in only one participant. These data suggest that, as expected, associating full names with the objects was somewhat more difficult than associating nicknames with the objects, and thus helped equate the training difficulty of the NAM condition with the training difficulty of the SEM condition.

The pre- and post-test accuracy and reaction time data were analyzed using a two-way analysis of variance (ANOVA) with test (pre/post) and training condition (SEM/NAM/NON), as predictor variables. One participant's pre-test data were lost due to a computer error and consequently, this participant was excluded from the analysis comparing pre- and post-test performance. For accuracy, there was a significant main effect of test ($F_{(1,9)} = 28.2, p < 0.001, MS_E = 0.00065$), with post-test judgments being made more accurately than pre-test judgments. There was also a significant main effect of test for reaction time ($F_{(1,9)} = 15.6, p < 0.005, MS_E = 10477$), with post-test judgments being made faster than pre-test judgments. These findings are summarized in Fig. 2A, which depicts the performance enhancement between pre- and post-test for both accuracy and reaction time. Although there appears to be a trend for an interaction in accuracy between condition and test, the interaction effect did not reach significance ($F_{(2,11)} = 1.02, n.s.$) and in fact, the lower enhancement for the SEM condition was due mostly to high pre-test accuracy.

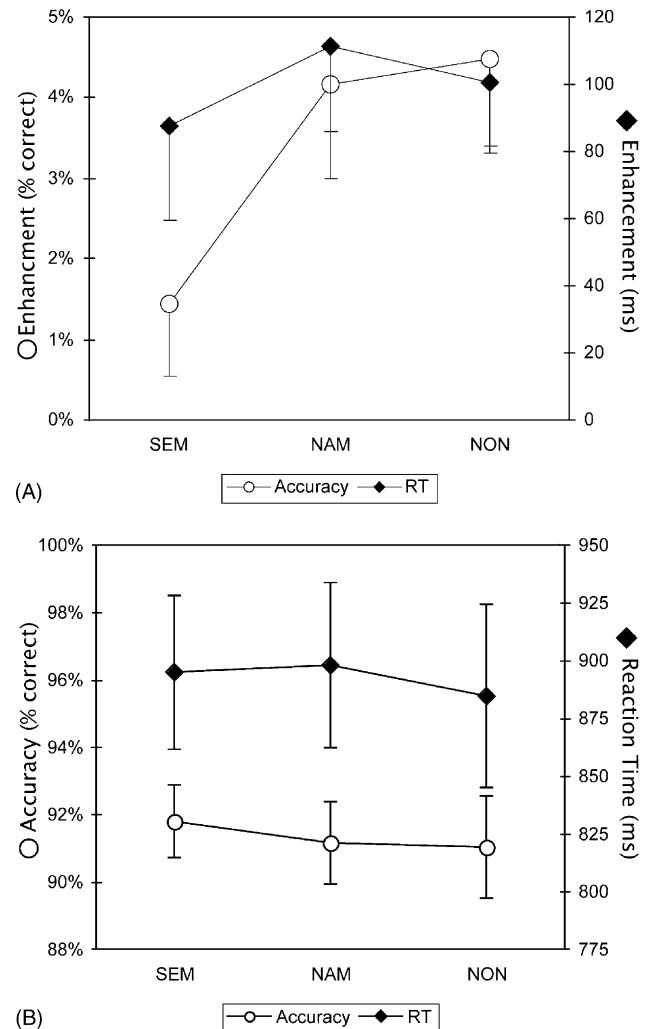


Fig. 2. (A) Performance enhancement between pre- and post-test. Change in accuracy is plotted on the left axis and change in reaction time is plotted on the right axis. For both axes, increasing values represent an increase in performance from pre- to post-test. Error bars represent standard error calculated for a single-sample t -test. (B) Mean accuracies and reaction times collapsed across conditions. Accuracy is plotted on the left axis and reaction time is plotted on the right axis. Error bars are standard error of the mean.

3.2. Imaging data

For the main contrast of interest, SEM > NAM, the overall analysis and both group analyses produced overlapping activation in the left inferior frontal cortex ($t_{(11)} > 3.0, p < 0.012$), a region that was predicted to be involved with processing the SEM objects. The posterior parietal cortex was also activated in all three analyses. The only region that reached significance in the overall analysis but that was only significant in one of the two group analyses was the post-central gyrus, which was highly active only for Group 2. A few other regions were activated in one group or the other, but failed to reach significance in the overall analysis, including activation in the fusiform gyrus for Group 1. Finally, Group 2

produced below threshold ($t_{(11)} = 2.0$, $p < 0.06$) NAM < SEM activation bilaterally at the temporal pole (coordinates $\pm 52, 12, -15$) that was not seen in Group 1. These interactions are likely a result of the slightly different training procedures used with Groups 1 and 2 or even due to individual differences between the participants. The important point is that most of the regions that activated in the overall analysis were activated in both group analyses, even though the power of these analyses was much lower than the overall analysis.

A contrast comparing the SEM condition and the NAM condition from the matching task data for all 12 participants produced the activation map shown in Fig. 3. These two conditions were matched on the amount of experience with each object, but the type of experience differed. There were no significant clusters of voxels that showed greater activation for NAM objects than for SEM objects, even at a relatively low statistical threshold ($t_{(11)} > 2.0$, $p < 0.06$, uncorrected) even though, as described above, Group 2 by itself did produce a NAM > SEM effect at this threshold. There were clusters of voxels that showed significantly greater activation for SEM objects than for NAM objects (indicated in red). These clusters of voxels are detailed in Table 2. Yellow circles (Fig. 3) indicate regions of the left inferior frontal cortex (LIF) which, consistent with our hypothesis, produced significantly more activation ($t_{(11)} = 4.69$, $p < 0.001$) when participants matched SEM objects than when they matched NAM objects.

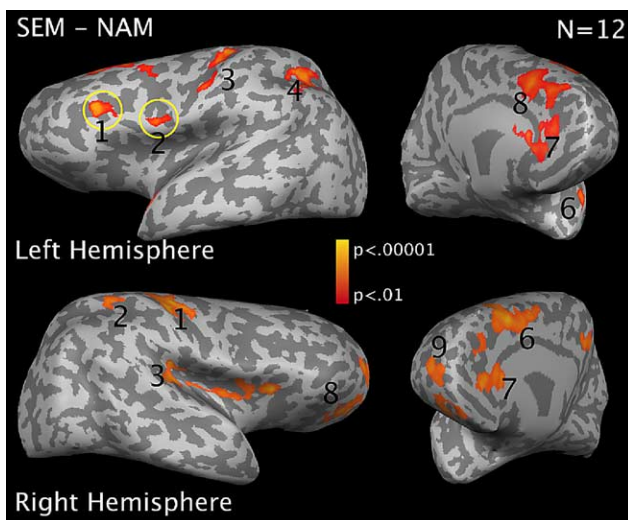


Fig. 3. Brain regions involved in processing artificial concepts. Brains are “inflated” representations of the cortical surface. Light grey areas represent gyri and dark grey areas represent sulci. The two images on the left are lateral views and the two images on the right are anterior medial views. Activation maps represent a contrast between SEM and NAM conditions. All significant activations represent brain regions where SEM > NAM. Descriptions of numbered regions are given in Table 2. Yellow circles indicate regions in the inferior frontal cortex.

The SEM condition was also contrasted with the NON condition. Participants were not given the same visual experience with these objects that was available to them for the SEM condition. Table 3 details the clusters of voxels for this contrast. In short, SEM objects produced more activation than NON objects in inferior frontal and parietal locations ($t_{(11)} > 3.0$, $p < 0.015$). The LIF ROIs for the SEM–NAM and SEM–NON contrasts were non-overlapping, which suggests that the functional anatomy of the LIF is not homogenous. In addition to these anterior ROIs, NON objects produced more activation than the SEM objects in occipital regions, which may have been due to the difference in familiarity between the SEM and the NON objects and may be related to repetition priming effects, which also have an occipital locus (Cabeza & Nyberg, 2000; Schacter & Buckner, 1998; Wiggs & Martin, 1998), or may have been due to increased scrutiny of the NON objects because of their novelty.

Finally, the NAM condition was contrasted with the NON condition. Although the participants had far greater experience with the NAM objects than the NON objects, there were no brain regions that produced more activation to NAM objects than NON objects, even at a relatively low statistical threshold ($t_{(11)} > 2.0$, $p < 0.06$, uncorrected). There were, however, several regions in the medial temporal lobe that produced more activation for NON objects than NAM objects, which may have been due to the novelty of the NON objects.

4. Discussion

Novel objects that were associated with semantic information (SEM) produced more activation in the left inferior frontal cortex (LIF) than either non-trained (NON) novel objects or novel objects that were associated with people’s names (NAM). Furthermore, this effect was found when participants engaged in a simultaneous visual-matching task, a task that requires no explicit recall of the associated knowledge. Many previous neuroimaging studies of semantic encoding, retrieval and generation have found semantic-related activation in the LIF (for review, see Bookheimer, 2002; Cabeza & Nyberg, 2000; Murtha et al., 1999). More specifically, several studies of object recognition have determined that the LIF may contribute to the processing of the meaning associated with objects (Koutstaal et al., 2001; Leveroni et al., 2000; Martin, 1999; Vuilleumier et al., 2002). Our findings provide converging evidence that the LIF is involved in processing semantic features associated with objects. More specifically, although the LIF has previously been found to process semantic associations learned over a lifetime and which are consistent with one’s entire body of semantic knowledge, we have shown that the LIF is also engaged

Table 2
Significant clusters for SEM–NAM comparison

ID	mm ³	X	Y	Z	Name	BA	<i>t</i> ₍₁₁₎	<i>p</i> <
<i>Left hemisphere</i>								
1	1171	−37	14	28	L inferior/middle frontal	9	4.69	0.0007
2	623	−52	−5	15	L inferior frontal	44	3.73	0.003
3	1346	−28	−33	39	L post-central	2	7.16	0.00002
4	2048	−26	−56	39	L precuneus	19	7.13	0.00002
5	1060	−40	−27	0	L superior temporal	22	6.94	0.00002
6	185	−46	5	−6	L superior temporal	22	4.04	0.002
7	5560	−19	24	25	L anterior cingulate	24	4.88	0.0005
8	1373	−4	14	42	L cingulate	32	4.69	0.0007
<i>Right hemisphere</i>								
1	1880	33	−25	43	R post-central	2	5.66	0.0001
2	274	22	−54	45	R precuneus	7	5.42	0.0002
3	570	39	−31	13	R superior temporal	42	4.81	0.0005
4	846	22	−16	−1	R para-/hippocampus		5.05	0.0004
5	688	26	−69	7	R middle occipital	19	4.31	0.001
6	2613	12	17	37	R cingulate	32	6.45	0.00005
7	3002	16	41	12	R medial frontal	10	5.74	0.0001
8	751	39	44	−10	R middle frontal	10	7.30	0.00002
9	688	19	60	−8	R superior frontal	10	4.59	0.0008

Note: *p*-values are uncorrected for multiple comparisons; region 5 in the left hemisphere is absent from Fig. 3 because it was too deep in the sulcus, region 5 in the right hemisphere is absent because it appeared to be almost completely in the white matter, and region 4 in the right hemisphere was absent because this part of the hippocampus is masked as a subcortical structure.

Table 3
Significant clusters for SEM–NON comparison

ID	mm ³	X	Y	Z	Name	BA	<i>t</i> ₍₁₁₎	<i>p</i> <
<i>Left hemisphere</i>								
SEM > NON								
a	301	−39	35	13	L inferior frontal	46	4.29	0.001
b	754	−47	5	31	L inferior frontal	9	3.36	0.006
c	573	−53	−22	35	L post-central	2	4.25	0.001
NON > SEM								
d	514	−35	−73	0	L middle occipital	19	−5.84	0.0001
e	552	−35	−73	0	L inferior temporal	37	−5.85	0.0001
f	364	−10	−91	2	L lingual	17	−4.80	0.0006
<i>Right hemisphere</i>								
SEM > NON								
a	306	17	63	−3	R medial frontal	10	4.83	0.001
b	325	8	14	39	R cingulate	32	5.34	0.0002
NON > SEM								
c	364	10	−91	2	R lingual	17	−4.80	0.001
d	758	53	−22	35	R post-central	2	−5.75	0.0001
e	555	46	5	−6	R superior temporal	22	−3.75	0.003

by recent and arbitrary associations. Moreover, by manipulating the semantic information associated with each object set independent of their visual features, we find more direct support for the processing of semantic associations in the LIF. This does not imply, however, that the LIF would respond to all arbitrary associations. The semantic features in our study were familiar and associating novel semantic features (e.g., <is glorpable>) with objects may have produced different results.

The use of a visual matching task implies that semantic associations stored in the LIF may be accessed somewhat automatically during visual processing. That is, whereas our results do not guarantee that semantic associations would be obligatorily retrieved under any test condition, they reveal that they are retrieved under conditions that are typical of a visual task used to study early stages of visual perception (Biederman & Gerhardstein, 1993; Ellis & Allport, 1986; Hayward & Williams, 2000; Lawson &

Humphreys, 1996; Tarr, Hayward, Williams, & Gauthier, 1998). This reinforces behavioral work from our lab that found differences in performance on a visual task between groups that were trained under different semantic conditions (Gauthier et al., 2003). Together, these results imply that the processes responsible for producing our perceptions may not be completely encapsulated from other cognitive processes (Pylyshyn, 1999). Thus, caution should be used when distinguishing between low-level perceptual abilities and high-level cognitive abilities, a common distinction in neuropsychology and cognitive neuroscience (Farah, 1990; Humphreys & Riddoch, 1987; Kolb & Whishaw, 1996).

The pattern of activation evoked by semantic associations was not restricted to regions of the inferior frontal cortex. As described earlier, activation patterns during tasks involving semantic generation and retrieval also tend to activate widespread regions of the cortex. These regions appear to be quite variable across experiments, with the exception of the inferior frontal cortex. In fact, we found some differences even between the two groups of participants tested in this study. There are a number of reasons why semantic processing may be distributed across the entire cortex. One account is that semantic information is represented in a modality-specific manner and that semantic information of different modalities is processed in different parts of the cortex (Barsalou, 2003; Farah & McClelland, 1991; Martin, Ungerleider, & Haxby, 2000). Related to this is the idea that there are brain regions that process the social meaningfulness of visual stimuli (for review, see Adolphs, 2001). Another theory is that semantic features that are learned through direct sensory stimulation may be stored differently than features that are learned only verbally (Noppeney & Price, 2003). All of the semantic features in the present experiments were learned verbally, not by direct sensory experience, and one benefit of using novel objects trained with new semantic associations is that the types of features that were used can be analyzed. According to the semantic feature type nomenclature of Cree and McRae (2003), the features shown in Fig. 1 were drawn from five different types: encyclopedic (strong, friendly, fragile, nervous, nocturnal), tactile (cold, soft, sticky, heavy), visual, sound and taste. Most of the features belonged to the encyclopedic and tactile types. The present study was not designed to investigate different theories of semantic memory, however, some speculations may still be made. First, the LIF activation may reflect processing of encyclopedic features somewhat like a supramodal convergence zone (Barsalou, 2003); the post-central (somatosensory cortex) activation may reflect the processing of tactile features. In fact, there is evidence from our laboratory that processing of some semantic information may occur in sensory-specific brain regions (James & Gauthier, 2003). Second, the pattern of activation in the left hemisphere

is strikingly similar to that observed for semantic information that is learned verbally (Noppeney & Price, 2003) as compared with semantic information that is learned through direct sensory experience. Third, activation of bilateral somatosensory, anterior cingulate and medial frontal cortices suggest the recruitment of brain regions involved in social cognition (Adolphs, 2001), possibly because the objects were visually biological in nature. These alternative explanations can certainly be tested in future studies.

In the behavioral results shown in Fig. 2, there was no significant interaction between training condition and pre- versus post-test performance. It is possible that the present study lacked sufficient power to show this effect, because the sample size ($N = 12$) was smaller than in the behavioral demonstration, where 32 subjects were tested (Gauthier et al., 2003). Therefore, the trend shown in Fig. 2A is worth discussing. Namely, SEM objects showed less enhancement in performance between the pre- and post-tests than the NAM or NON objects. This suggests that associating adjectives with the novel objects during training actually interfered with any visual learning that should have occurred during the training process. In contrast to this result, we had anticipated that learning artificial concepts with non-overlapping features would actually aid the most in discriminating between the novel objects. This expectation was based on work from our own lab showing that learning dissimilar arbitrary concepts benefits you relative to learning similar artificial concepts (Gauthier et al., 2003) and patient work suggesting that the severity of visual naming deficits are lessened when novel objects are associated with dissimilar concepts (Dixon, Bub, & Arguin, 1997, 1998). A possible explanation concerns the random assignment of features to each artificial concept and the arbitrary mapping of features to a particular object. This design was important in our fMRI study for ensuring independence of visual and other semantic features, but semantic information may only aid visual discrimination when semantic features are associated with objects that display appropriate visual features. For instance, learning that the widest YUFO was <heavy> could help discriminate it better, whereas learning that the slimmest YUFO was <heavy> may have actually hindered discrimination. Associated people names did not apparently lead to similar interference, even though the associated names were arbitrarily assigned. Presumably, names carry little associated semantic information and therefore can be assigned arbitrarily with no repercussion. These speculations are only based on trends in our behavioral data and it is clear that they could be directly investigated in future studies where the mapping between visual and semantic information would be systematically manipulated.

There are three alternative accounts for our inferior frontal cortex activation that need to be addressed.

First, the inferior frontal cortex has been implicated in memory retrieval of labels. Therefore, the pattern of results may imply that memory retrieval was more demanding for the SEM than the NAM condition. The SEM and NAM conditions were equated in terms of name labels; both conditions required the learning of proper names. It is possible, however, that the semantic features were also treated as labels. The fact that reaction times were under 1 s suggests that a labourious search for labels was not occurring. Furthermore, data from our previous study (Gauthier et al., 2003) showed that behavioral changes were resistant to verbal working memory manipulations, suggesting again that retrieval of labels was not used. Second, the difference in activation between the SEM and NAM conditions could have been due to overlearning of the NAM objects. This could have been the case had we only run Group 1, however, the inclusion of Group 2 showed that making the NAM training more difficult had no effect on the pattern of activation. Third, the different patterns of activation between SEM and NAM objects may not have been due to the creation of associations between verbal semantic information and visual features of objects. Alternatively, the pairing of semantic features with the objects during training may have directed attention to different visual features than the pairing of names with the objects. Differences in the allocation of attention during training may have given rise to different visual representations. Differences in representation may have been reflected in the different patterns of activation. There is a good reason to believe, however, that this was not the case; there was little difference in the pattern of activation in ventral temporal and occipital cortices. These regions are thought to be involved in the visual representation of objects (Grill-Spector et al., 2001; James, Humphrey, Gati, Menon, & Goodale, 2000; Kanwisher et al., 1996; Malach et al., 1995), whereas there is little evidence for the involvement of the inferior frontal cortex.

In conclusion, we have shown that visual processing of novel objects associated with arbitrary semantic information leads to activity in brain regions that have been implicated in the processing of semantic knowledge. Specifically, the inferior frontal cortex produced more activation for objects trained with semantics than those trained with proper names. Access to these semantic associations appeared to be involuntary, because recall of the associations was not necessary to perform the visual-matching task that we used. Semantic information about common objects is usually built up over a lifetime of experiences. In our experiment, semantic associations were developed in less than two hours in sessions spanning only a few days, yet these associations appeared to be of sufficient strength to activate a network of semantic processing regions that is normally activated by common objects. Finally, associating arbitrary

semantic information with novel objects promises to provide new insights into the neural substrates of semantic memory without the natural confounds associated with using common objects.

Acknowledgements

This work was supported by the James S. McDonnell Foundation (IG), the National Science Foundation under grant no. 0091752 (IG), the National Institute of Health under grant no. EY13441 (IG), and a postdoctoral fellowship from the Canadian Institutes of Health Research (TWJ). Thanks to James D. Stefansic for his assistance with the neuroimaging protocols and to Karin H. James, other members of the Perceptual Expertise Network (PEN) and three anonymous reviewers for their comments and suggestions on this work.

References

- Adolphs, R. (2001). The neurobiology of social cognition. *Current Opinion in Neurobiology*, *11*, 231–239.
- Aguirre, G. K., & Farah, M. J. (1998). Human visual object recognition: What have we learned from neuroimaging? *Psychobiology*, *26*(4), 322–332.
- Barsalou, L. W. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, *7*(2), 84–91.
- Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 1162–1182.
- Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience*, *25*, 151–188.
- Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *16*(13), 4207–4221.
- Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, *12*(1), 1–47.
- Cree, G., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and chello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, *132*, 163–201.
- Dixon, M., Bub, D. N., & Arguin, M. (1997). The interaction of object form and object meaning in the identification performance of a patient with category-specific visual agnosia. *Cognitive Neuropsychology*, *14*(8), 1085–1130.
- Dixon, M., Bub, D. N., & Arguin, M. (1998). Semantic and visual determinants of face recognition in a prosopagnosic patient. *Journal of Cognitive Neuroscience*, *10*, 362–376.
- Ellis, R., & Allport, D. A. (1986). Multiple levels of representation for visual objects: A behavioural study. In A. G. Cohn & J. R. Thomas (Eds.), *Artificial intelligence and its applications* (pp. 245–247). New York: Wiley.
- Farah, M. J. (1990). *Visual agnosia: Disorders of object recognition and what they tell us about normal vision*. Cambridge: MIT Press.
- Farah, M. J., & McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General*, *120*, 339–357.

- Gauthier, I., James, T. W., Curby, K. M., & Tarr, M. J. (2003). The influence of conceptual knowledge on visual discrimination. *Cognitive Neuropsychology*, 20(2/4–6), 507–523.
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, 41(10–11), 1409–1422.
- Hayward, W. G., & Williams, P. (2000). Viewpoint dependence and object discriminability. *Psychological Science*, 11, 7–12.
- Humphreys, G. W., & Forde, E. M. E. (2001). Hierarchies, similarity, and interactivity in object recognition: “Category-specific” neuropsychological deficits. *Behavioral and Brain Sciences*, 24(3), 453–509.
- Humphreys, G. W., & Riddoch, M. J. (1987). The fractionation of visual agnosia. In G. W. Humphreys & M. J. Riddoch (Eds.), *Visual object processing: A cognitive neuropsychological approach*. London: Lawrence Erlbaum.
- James, T. W., & Gauthier, I. (2003). Auditory and action semantic features activate sensory-specific perceptual brain regions. *Current Biology*, 13, 1792–1796.
- James, T. W., Humphrey, G. K., Gati, J. S., Menon, R. S., & Goodale, M. A. (2000). The effects of visual object priming on brain activation before and after recognition. *Current Biology*, 10, 1017–1024.
- Kanwisher, N., Chun, M. M., McDermott, J., & Ledden, P. J. (1996). Functional imaging of human visual recognition. *Brain Research. Cognitive Brain Research*, 5(1–2), 55–67.
- Kolb, B., & Whishaw, I. Q. (1996). *Fundamentals of human neuropsychology*. New York: W.H. Freeman.
- Koutstaal, W., Wagner, A. D., Rotte, M., Maril, A., Buckner, R. L., & Schacter, D. L. (2001). Perceptual specificity in visual object priming: Functional magnetic resonance imaging evidence for a laterality difference in fusiform cortex. *Neuropsychologia*, 39, 184–199.
- Lawson, R., & Humphreys, G. W. (1996). View specificity in object processing: Evidence from picture matching. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 395–416.
- Leveroni, C. L., Seidenberg, M., Mayer, A. R., Mead, L. A., Binder, J. R., & Rao, S. M. (2000). Neural systems underlying the recognition of familiar and newly learned faces. *The Journal of Neuroscience*, 20(2), 878–886.
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., & Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 92(18), 8135–8139.
- Martin, A. (1999). Automatic activation of the medial temporal lobe during encoding: Lateralized influences of meaning and novelty. *Hippocampus*, 9, 62–70.
- Martin, A., Ungerleider, L. G., & Haxby, J. V. (2000). Category specificity and the brain: The sensory/motor model of semantic representations of objects. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences*. Cambridge: MIT Press.
- Murtha, S., Chertkow, H., Beauregard, M., & Evans, A. (1999). The neural substrate of picture naming. *Journal of Cognitive Neuroscience*, 11(4), 399–423.
- Noppeney, U., & Price, C. J. (2003). Functional imaging of the semantic system: Retrieval of sensory-experienced and verbally learned knowledge. *Brain and Language*, 84, 120–133.
- Pylyshyn, Z. (1999). Is vision continuous with cognition? The case of impenetrability of visual perception. *Behavioral and Brain Sciences*, 22, 341–423.
- Schacter, D. L., & Buckner, R. L. (1998). Priming and the brain. *Neuron*, 20(2), 185–195.
- Semenza, C. (1997). Proper-name-specific-aphasias. In H. Goodglass & A. Wingfield (Eds.), *Anomia neuroanatomical and cognitive correlates* (pp. 115–133). San Diego: Academic Press.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme Medical Publishers.
- Tarr, M. J., Hayward, W. G., Williams, P., & Gauthier, I. (1998). Three-dimensional object recognition is viewpoint dependent. *Nature Neuroscience*, 1, 275–277.
- Vuilleumier, P., Henson, R. N., Driver, J., & Dolan, R. J. (2002). Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nature Neuroscience*, 5(5), 491–499.
- Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology*, 8, 227–233.