NeuroImage **16,** 259–268 (2002)
doi:10.1006/nimg.2002.1061, available online at http://www.idealibrary.com on IDEAL®

Frequency of Concrete Words Modulates Prefrontal Activation during Semantic Judgments

Michael W. L. Chee,* Nicholas H. H. Hon,† David Caplan,‡ Hwee Ling Lee,* and Joshua Goh*

**Cognitive Neuroscience Laboratory, Singapore General Hospital, Singapore 169856; †Department of Social Work and Psychology, National University of Singapore, Singapore 119260; and ‡Neuropsychology Laboratory, Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts 02114*

Received June 7, 2001

The effect of word frequency on semantic processing was characterized by studying two groups of right-handed participants using fMRI. Stimuli were presented in blocks of either high frequency or low frequency word triplets where a sample word appeared above a pair of test words. One group ($n = 8$) made semantic judgments by selecting the word from the test pair that was more closely associated with the sample. Stimulus triplets were designed such that relatedness between sample and "correct" items was obvious. The other group ($n = 8$) read the words silently without making any semantic decision and pressed a button on completing the reading of each triplet. Semantic judgments while no less accurate, were associated with greater left prefrontal BOLD signal change when they involved low frequency words, whereas there was no reliable effect of word frequency in the reading condition. These findings suggest that retrieval effort modulates left prefrontal activity when deliberate access to semantics is required.

Key Words: **fMRI; semantic memory; word frequency; processing.**

INTRODUCTION

How neural activity varies with different cognitive tasks is of considerable interest because it may lend understanding to how mental activity is organized in the brain. Activity, as measured by increased task-related blood flow, was higher in the left prefrontal region when healthy volunteers had to select a word using a specific set of rules compared to when the criterion for selection was less constrained (Thompson-Schill *et al.*, 1997), when they had to discern how pairs of words were related when their degree of relatedness was low (Fletcher *et al.*, 2000; Seger *et al.*, 2000) when item classification involved less typical compared to more typical exemplars of a category (Roskies *et al.*, 2001) and when a greater depth of processing for word meaning (as opposed to perceptual features) was required (Otten *et al.*, 2001). Word categorization that required the use of semantic information resulted in greater left prefrontal activation than a task which took longer to perform but which did not require access to semantics (Demb *et al.*, 1995). What emerges from these observations is that the prefrontal region plays an important role in retrieving verbal semantic information.

Various accounts have been proposed to account for the modulation of left prefrontal activity observed with tasks requiring access to word meaning. One view attributes the modulation to the amount of controlled semantic retrieval necessary to perform the task (Wagner *et al.*, 2001). An alternative view is that the modulation results from the need to select task-relevant information from a number of competing alternatives (Thompson-Schill *et al.*, 1997) and that it is not semantic retrieval per se that drives left prefrontal activation.

We previously observed that when bilingual volunteers performed semantic judgments on concrete words, left prefrontal activity was higher in the volunteers' less proficient language (Chee *et al.*, 2001). Neither associative strength nor selection demands between test items was varied in these experiments and we suggested that lesser familiarity with words in the volunteers' less proficient language could have accounted for the observed effect.

Invasive electrophysiological recordings in the prefrontal cortex have shown that neurons in this region respond differently to familiar compared to unfamiliar items (Rainer and Miller, 2000). Specifically, they fire in a more spatially restricted manner and less frequently, in response to visual representations of familiar (relative to novel) pictures. One might suggest that familiar items require less "neuronal effort" for recognition. We hypothesized that words in the less familiar language may have less well-tuned representations, requiring greater neuronal activity (or "effort") during the retrieval of semantic information. This hypothesis,

1053-8119/02 $35.00

while intuitively appealing, is difficult to assess in a cross-language comparison because processing words in different languages may be subject to differences beyond retrieval effort.

To determine the effect of semantic retrieval effort on brain activation, we manipulated the frequency of concrete words used in an associative semantic judgment task while controlling the semantic relatedness between the test items as the latter can independently modulate brain activation (Fletcher *et al.*, 2000). We chose to manipulate word frequency because it is an index of how much exposure people have had to a particular word. High frequency words, those that occur more frequently in print, are named more quickly than low frequency words and are also more rapidly recognized as words in lexical decision tasks (Balota and Chumbley, 1985; Forster and Chambers, 1973; Frederiksen and Kroll, 1976). Frequency effects have also been observed in tasks that overtly require semantic access (Young and Rugg, 1992). Although the locus of this effect remains controversial, Monsell (1991) proposed that frequency effects reflect the cumulative effect of experience on the facility with which an observer identifies a word and recovers its meaning. It follows, then, that more effort is likely to be necessary when retrieving the meaning of a low frequency word, relative to a high frequency one.

To evaluate how deliberately accessing semantic information contributes to left prefrontal activation, we performed a second experiment where volunteers were required only to silently read words of differing frequency. We expected this experiment to reveal the relative contribution of word frequency to orthographic and phonological processing, acknowledging that some degree of automatic semantic access occurs with reading (Price *et al.*, 1996). If cognitive demands are higher with deliberate as opposed to incidental retrieval of semantic information, we could expect more activation in the semantic judgment experiment, compared to the reading-only experiment.

METHODS

Words used to create the stimulus triplets were obtained from the MRC Psycholinguistic Database (http://www.itd.clrc.ac.uk/Projects/Psych/psych.html). High frequency words had a median frequency of 59 occurrences per million words (Mean = 93.2, SD = 92.2). Low frequency words had a median frequency of 3 occurrences per million words (Kucera and Francis, 1967)(Mean = 2.8, SD = 1.5). Both high and low frequency words were matched on concreteness with respective "concreteness value" means of 579.76 and 579.02.

A rating exercise was conducted to ascertain the semantic relatedness between sample items and options. Four raters were chosen based on their good performance in the standardized English examination (UK based GCE "O" and "A" levels) taken as a requirement for tertiary education. The rating list was constructed by dividing the words in a given trial into two pairs, with the sample word paired with both "correct" and "incorrect" options separately. For example, in a trial where the sample word was "vampire" and the response options were "blood" and "leaf," the two pairs made would be *vampire-blood and vampire-leaf*. Raters indicated on a 9 point Likert-type scale (1 for not related and 9 for closely related) how closely associated the words in each pair were. The raters were blind as to the frequency of the words or the purpose of the experiment.

TABLE 1

Mean Relatedness Ratings between Sample Items (e.g., Vampire) and Their Corresponding Related (e.g., Blood) and Unrelated (e.g., Leaf) Test Options for High and Low Frequency Conditions

	High-frequency words	Low-frequency words
Related pairs	8.14	7.80
Unrelated pairs	1.51	1.09

Note. Ratings were based on a 9-point Likert-type scale, where 9 denotes "very closely related" and 1 denotes "not at all related."

There was high reliability across raters ($R = 0.99$). Mean ratings are shown in Table 1. An ANOVA taking option (related vs unrelated) and frequency (high vs low) as factors confirmed that the samples were rated as more closely related to the correct options than the wrong options [$F(1,3) = 2107.85$, $P < 0.001$]. Raters might have been more prone to giving higher ratings to high frequency word pairs compared to low frequency ones but this was only marginally reliable [$F(1,3) = 6.97$, $P < 0.08$]. Critically though, there was no interaction between the correctness of the option and the word frequency [$F < 1$, n.s.].

Experiment 1: Semantic Judgment Task

Eight neurologically normal, right-handed participants, five men and three women aged between 19 and 24 years, gave informed consent for this study. Participants were chosen on the basis of good performance in the standardized English examinations described above. As such, the volunteers can be thought of as having native speaker levels of proficiency in English.

We compared the semantic processing of high frequency and low frequency English words. Stimulus triplets were presented for 3.0 s and followed by 0.5 s of fixation. Participants performed two different matching-to-sample tasks involving visually presented words (Fig. 1). In the semantic task, participants viewed two words and were instructed to choose the word more closely related to the sample stimulus (uppermost item

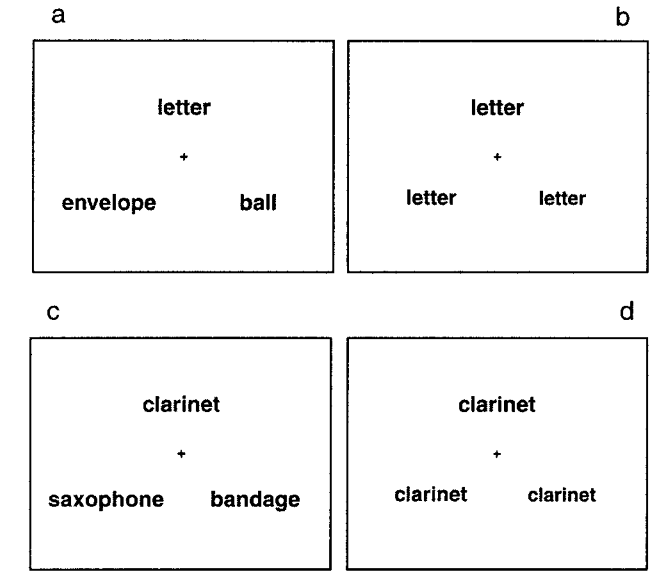


FIG. 1. Exemplars of experimental stimuli used in the semantic judgment and reading tasks (a,c) as well as the size judgment task (b,d). Stimuli involving high frequency test items appear in the upper row and those involving low frequency test items appear in the lower row.

in each panel). This task is known as the Pyramids and Palm Trees (PPT) task (Howard and Patterson, 1992). In the size judgment task, one of the words was 6% smaller (or larger) than the sample word and the other was 12% smaller (or larger). Participants were instructed to choose the item that was closer in size to the sample stimulus. Sample stimuli that appeared in the semantic judgment task were reused in the font-size judgment task.

The stimuli were presented in alternating blocks counter-balanced across runs, comprising a total of 64 trials for each of four conditions: low and high frequency semantic judgment, low and high frequency size judgment. Response time (RT) was collected while participants were scanned using a MR compatible, two-button mouse.

Experiment 2: Reading

A further group of eight, healthy right-handed participants, with the same demographic properties as the first group comprising five men and three women aged between 20 and 24 years gave informed consent for this study. The stimuli used in this experiment were identical to the ones used in Experiment 1. The difference in this experiment was that participants were instructed to silently read the stimulus triplets paying attention to pronunciation and to click the left mouse button once they completed reading the words. The control task was similar to that used in the previous experiment, i.e., size judgment.

Imaging and Image Analysis

Experiments were performed in a 2.0T Bruker Tomikon S200 system (Bruker, Karlsruhe, Germany). A blipped gradient-echo EPI sequence was used with a TR of 2000 ms, a FOV of 23 $\times$ 23 cm and a 128 $\times$ 64 pixel matrix. Fifteen oblique axial slices approximately parallel to the AC–PC line 4 mm thick (2 mm gap) were acquired. High-resolution anatomical reference images were obtained using a three-dimensional spoiled-gradient-recalled-echo sequence. Functional images underwent phase correction prior to further processing that was performed using Brain Voyager 2000 software (Brain Innovation, Maastricht, Holland). Intensity normalization was performed and followed by motion correction. Gaussian filtering was applied in the temporal and spatial domains. In the spatial domain a smoothing kernel of 4 mm FWHM was used for the computation of individual activation maps and 8 mm FWHM for computation of multisubject maps. In the temporal domain, a three time-point FWHM filter was used. Registration of the functional MR data set to the high-resolution anatomical image of the brain was performed by manually registering the stack of T2 images acquired in an identical orientation to the functional MR data set to the 3-D image. The resulting realigned data set was then transformed into Talairach space (Talairach and Tournoux, 1988).

A linear cross correlation map of the size judgment tasks in each experiment was first performed. No significant differences emerged from this comparison and the size judgment task was used as a common baseline condition in both semantic judgment and reading conditions.

Individual subject statistical maps were computed using a general linear model (GLM) using two explanatory variables: low and high frequency test items. The expected BOLD signal change was modeled using a gamma function (tau of 2.5 seconds and a delta of 1.5) synchronized to blocks of cognitive tasks. Statistical maps for individual participants from which ROI based analysis was performed were created using a correlation coefficient threshold of 0.4 that corresponds to a corrected statistical threshold of $P < 0.001$.

For each individual's data, regions of interest (ROI) in the left prefrontal region (corresponding to Brodmann's areas 44, 45, 9, and 47) encompassing the inferior and middle frontal gyri were defined by sampling volumes that were active in both low and high frequency semantic judgment relative to size judgment. We selected ROI's jointly activated in both high and low frequency conditions as this was deemed the least biased comparison of activation between these conditions given that random, as well as unknown systematic effects may contribute to differences in spatial location of activation when high or low frequency items are individually selected as the predictor of interest

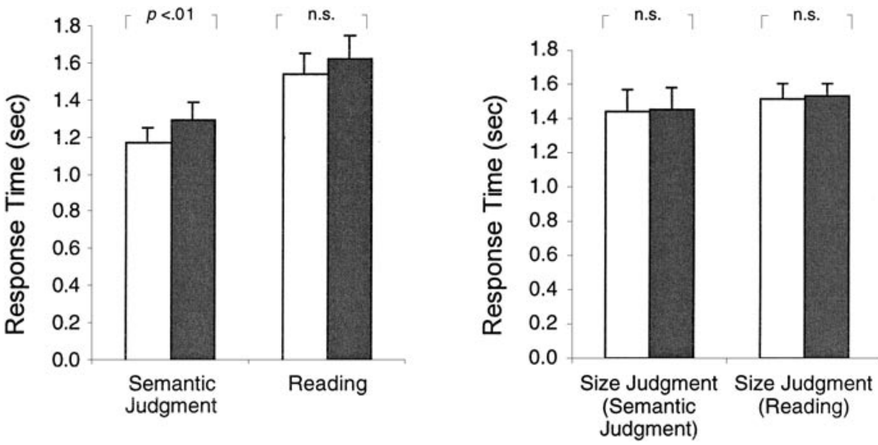


FIG. 2. Response times associated with semantic judgment and reading of high and low frequency items. The statistical significance of differences in pair-wise comparisons between conditions of each experiment is shown. Error bars denote 1 standard error.

(see Fig. 3, Tables 2, 3, and 4). Temporal ROI's lay in the middle and posterior temporal region including the supramarginal, superior temporal gyrus and middle temporal gyrus (corresponding to BA 21 and 22). These prefrontal and temporal areas were chosen for evaluation of percentage signal change information because they yielded the most robust responses across participants. The cluster of voxels, bounded by a $3 \times 3 \times 3$ cm. cube centered on the activation peak defined the ROI (only activated voxels within this bounding cube were counted). Within each individual's ROI, averaged time courses comprising 28 time points (14 task related and 14 baseline points) were calculated in order to show the average BOLD signal change due to the semantic tasks with respect to their size judgment baseline tasks. BOLD signal change was expressed as percentage signal change relative to the baseline (size judgment) task. Mean percentage signal change for each semantic judgment task was calculated from points 5 to 14 located on the plateau of the BOLD response corresponding to the semantic task and from the points in time 20 to 28 corresponding to the size judgment task. In this way, points in the transition phase during the rise and fall of the BOLD signal were omitted.

Percentage signal change values for high and low frequency words associated with each ROI were compared across volunteers using paired t tests. (i.e., signal change for low frequency words in the left prefrontal ROI was compared to signal change for high frequency words in the same ROI). A similar analysis was performed with BOLD signal change data obtained from volunteers performing the reading experiment.

A general linear model (GLM) was used to compute the group level voxel-by-voxel activation maps. For each experiment group, the GLM was computed by setting low and high frequency items as explanatory variables, in addition to coding each subject as an explanatory variable. This generated within and between group comparisons of the effect of frequency and task, respectively. The corrected statistical threshold used for group level analysis was $P < 0.01$.

Talairach transformed group data was displayed on a volume-rendered brain of an individual from the cohort. Activations to a depth of 15 mm were projected to the surface and displayed using a red (min), yellow (max) color scale.

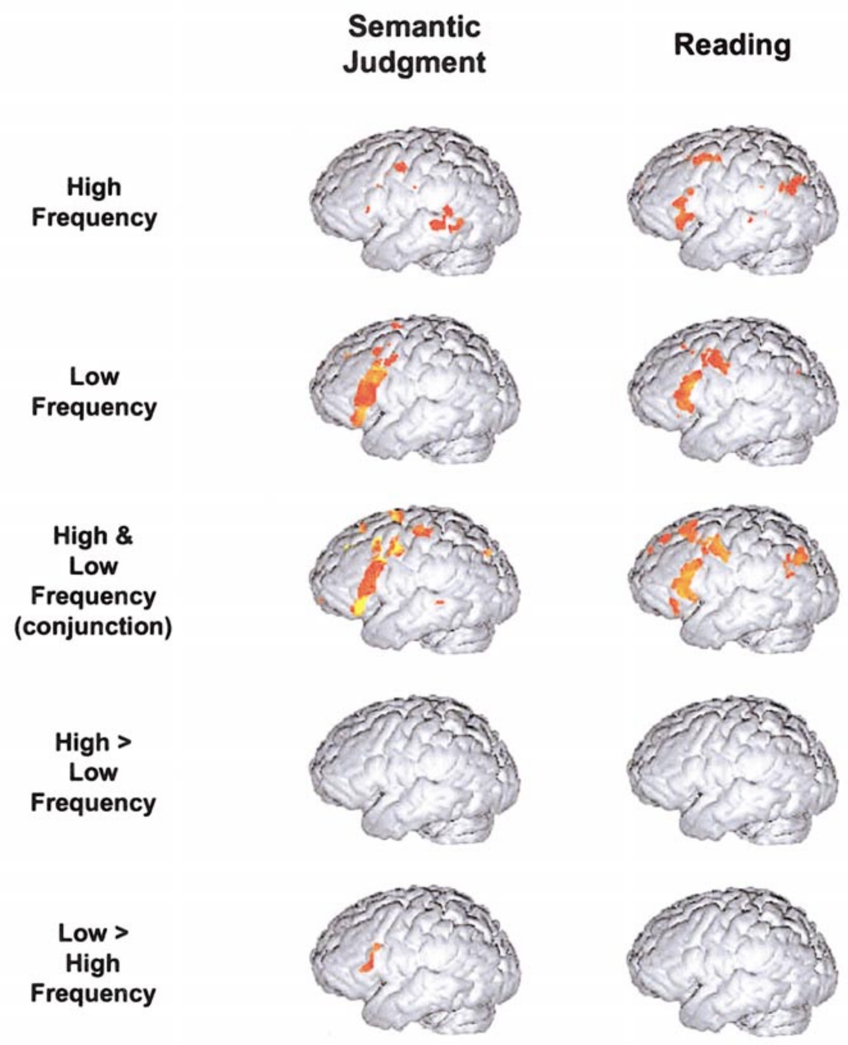
RESULTS

Experiment 1: Semantic Judgment

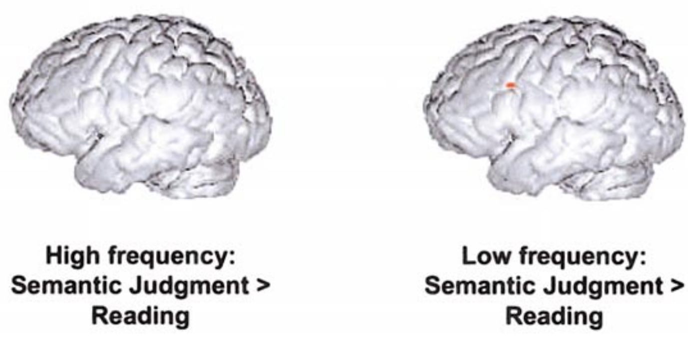
There was no significant difference in accuracy whether participants performed semantic judgments on high frequency (Mean = 98.0%, SD = 1.1) or low frequency (Mean = 95.9%, SD = 0.9) word triplets [$t(7) < 1$, n.s.]. Response times were longer when the triplets were made up of low frequency words [$t(7) = 3.86$, $P < 0.01$] (Fig 2).

We performed additional analyses to check for word or syllable length effects. On the average, high frequency words were shorter than low frequency ones (5.0 vs 5.6 letters per word, respectively). There were no reliable correlations between RT and total number of letters in a trial for either high frequency ($P > 0.6$) or low frequency trials ($P > 0.2$). Low frequency words had more syllables than high frequency ones. Separating the data on the basis of frequency, we found no reliable correlations between RT and total number of syllables in a trial for either high frequency words ($P > 0.4$) or low frequency words ($P > 0.5$). As such, we determined that word and syllable length did not significantly influence response times in this experiment.

Semantic judgments involving both low and high frequency words relative to size judgment activated a network of areas that included the left prefrontal (BA



3



4

FIG. 3. Regions activated during semantic judgment or reading with high or low frequency words compared to word size judgment. The two lowermost rows show the direct contrast between activation associated with low and high frequency items in each experiment.

FIG. 4. Areas active in the interaction between experimental task and word frequency.

9, 44), middle and posterior temporal (BA 21, 22), and parietal (BA 7) regions (Fig. 3). A strong left hemisphere predominance of activation was observed and subsequent discussion relates only to left hemisphere activation. All eight individuals showed left prefrontal activation. Five of eight showed left posterior temporal activation above threshold (Table 2).

Within the left prefrontal cortex, a region at the superior end of the inferior frontal gyrus corresponding to BA 44 showed significantly higher BOLD signal in the low frequency condition (Fig. 3; Table 3). This area was also revealed in the contrast between low and high frequency conditions (Fig. 3). The effect of frequency was also observed in the ROI based analysis (Fig. 5)

TABLE 2

Activation Peak Loci and Correlation Coefficient (cc) Values in Individual Left Prefrontal and Posterior Temporal ROI That Showed Conjoined Activation in High and Low Frequency Conditions during Semantic Judgment and Reading

	Semantic judgment							
	Prefrontal				Temporal			
	x	y	z	*cc*	x	y	z	*cc*
S1	−47	15	24	0.64	−49	−51	6	0.78
S2	−47	24	19	0.66	−59	−30	5	0.52
S3	−36	15	31	0.78	—	—	—	—
S4	−52	22	27	0.66	−60	−26	−3	0.82
					−60	−47	−3	0.86
S5	−46	10	37	0.58	—	—	—	—
S6	−27	8	42	0.68	−45	−57	4	0.74
S7	−49	27	27	0.66	−52	−51	−3	0.76
S8	−45	18	24	0.48	—	—	—	—
	Reading							
	Prefrontal				Prefrontal			
	x	y	z	*cc*	x	y	z	*cc*
S9	−46	13	26	0.66	−49	−29	0	0.72
S10	−46	7	18	0.80	−58	−22	3	0.78
S11	−38	16	30	0.46	−48	−36	8	0.64
S12	−37	10	29	0.68	−55	−35	30	0.74
	−38	15	22	0.58				
S13	−49	19	22	0.83	−61	−51	12	0.60
S14	—	—	—	—	−50	−54	9	0.46
S15	−33	16	26	0.62	—	—	—	—
S16	−49	16	25	0.46	−59	−23	0	0.42

Note. Regions with a correlation coefficient of less than 0.4 were not included in the ROI analysis.

[$t(7) = 8.20$, $P < 0.0001$] and was consistently present in all eight volunteers.

As a result of the relatively lesser spatial extent and greater variability of temporal activation as well as the modest spatial smoothing used to process the data, the activated area depicted in the voxel-by-voxel analysis was less conspicuous than that seen in individual volunteers (Fig. 3; Tables 2 and 4). The apparently greater activation in the left midtemporal region (BA 21, 22) with the high frequency condition was not statistically significant in the direct contrast between high and low frequency conditions.

Experiment 2: Reading

There was no effect of word frequency on response times [$t(7) = 1.47$, $P > 0.1$] (Fig. 2).

The left prefrontal regions activated by reading were broadly similar to those activated during semantic judgment when comparing the respective conjunctions of activation in relation to high and low frequency items. This was also true of the comparison of activation associated with low frequency items (Fig. 3). As the activation peaks in the high frequency condition were in relative proximity to those observed for the low frequency condition but at lower levels of signal (Tables 3 and 4), we speculate that the differences in spatial location of activation between experiments in the high frequency conditions were contributed by just-below-threshold activation.

Activation of the left angular gyrus that has been demonstrated in some studies involving reading (see (Price, 2000) for a review) was evident in the present data but the difference between reading and semantic judgment was not significant at the group level (Fig. 4).

The group level voxel-by-voxel analysis did not show a significant effect frequency on left prefrontal activation. This was a result of the individual variability in magnitude of activation with reading (Fig. 5). Of the eight individuals, four showed more activation for low frequency words, one showed no activation above threshold in the ROI and three activated more for high frequency words [$t(6) < 1$, n.s.] (Fig. 4).

Comparison between Semantic Judgment and Reading

A mixed design ANOVA with word frequency (High vs Low) as a within subject variable and task (Semantic Judgment vs Reading) as a between subject vari-

TABLE 3

Activation Peak Loci and Correlation Coefficient (cc) Values in Individual Left Prefrontal ROI (BA 44, 9) during the High and Low Frequency Conditions in Semantic Judgment and Reading Experiments

	Semantic judgment							
	High frequency				Low frequency			
	x	y	z	*cc*	x	y	z	*cc*
S1	−47	15	24	0.46	−49	16	24	0.46
S2	−46	22	16	0.42	−47	25	17	0.54
S3	−34	14	32	0.44	−46	13	33	0.68
S4	−51	21	26	0.40	−51	22	28	0.50
S5	−47	10	36	0.34	−50	16	25	0.58
S6	−37	4	33	0.42	−37	4	33	0.60
S7	−52	16	15	0.26	−43	22	17	0.46
S8	−35	3	33	0.52	−43	11	39	0.62
	Reading							
	High frequency				Low frequency			
	x	y	z	*cc*	x	y	z	*cc*
S9	−55	16	24	0.42	−46	13	26	0.56
S10	−49	10	19	0.62	−46	7	18	0.54
S11	−40	18	30	0.38	−43	16	27	0.34
S12	−37	10	29	0.44	−37	10	29	0.50
	−37	25	21	0.50	−46	28	16	0.54
S13	−52	21	17	0.52	−49	18	25	0.66
S14	—	—	—	—	—	—	—	—
S15	−33	16	26	0.36	−33	16	26	0.52
S16	−49	14	27	0.24	−49	16	25	0.44

able, showed main effects of frequency [$F(1,14) = 9.99$, $P < 0.05$] and task [$F(1,14) = 5.369$, $P < 0.01$] on response times but no interaction [$F(1,14) < 1$, n.s.]. Overall, participants took longer for reading than for semantic judgment despite taking a comparable time for the perceptual judgment task (all $F < 1$). This was possibly due to the delay imposed by the self-monitoring required to indicate the completion of silent reading task.

In the left prefrontal region, the task by frequency analysis showed a main effect of frequency that was largely driven by the semantic judgment task (Figs. 3–5). There was no main effect of task but there was greater activation at the superior end of the left inferior frontal gyrus (Talairach coordinate: −34, 11, 32; BA44) for low frequency semantic judgments compared to the reading of low frequency words (Fig.5).

ROI based analysis corroborated these findings showing a main effect of frequency [$F(1,13) = 13.07$, $P < 0.005$], no effect of task [$F(1,13) < 1$, n.s.] and an interaction between frequency and task [$F(1,13) = 6.66$, $P < 0.05$].

DISCUSSION

We found higher left prefrontal BOLD signal change when volunteers performed semantic judgments on low frequency words. This effect was not evident when volunteers had only to read these words. The contrast between the semantic judgment experiment and the reading experiment suggests that the locus of this effect is likely to be semantic, as the effect of frequency on orthographic, phonological and lexical processing would likely have been accounted for in the reading experiment. These findings suggest that retrieval effort modulates prefrontal activity when deliberate access to semantics is required.

A similar idea has recently been advanced by Wagner (Wagner et al., 2001), who proposed that left prefrontal activity is modulated by the degree of controlled retrieval necessary to perform the task. Greater left prefrontal activation occurred when associative judgments made on loosely associated words were contrasted with those made on strongly associated ones arguably because retrieving information that links weakly associated items is more effortful. In our semantic judgment task, the correct response was always strongly associated to the sample word for both high and low frequency conditions. Further, we did not specify the dimension along which an associative decision was to be made, allowing for all available information about the test items to be used. As such, the effort required for semantic retrieval would be determined

TABLE 4

Activation Peak Loci and Correlation Coefficient (cc) Values in Individual Left Posterior-Lateral Temporal ROI (BA 21, 22) during the High and Low Frequency Conditions in Semantic Judgment and Reading Experiments

	Semantic judgment							
	High frequency				Low frequency			
	x	y	z	*cc*	x	y	z	*cc*
S1	−49	−52	6	0.58	−46	−59	6	0.58
S2	−61	−29	3	0.36	−58	−30	6	0.42
S3	−50	−56	−6	0.32	−49	−54	−12	0.34
S4	−49	−29	−4	0.60	−61	−26	−3	0.60
	−58	−47	−3	0.64	−61	−47	−3	0.56
S5	−58	−56	9	0.38	−58	−56	9	0.28
S6	−49	−61	0	0.48	−42	−56	3	0.60
S7	−52	−49	−3	0.52	−52	−49	−3	0.58
S8	−52	−49	0	0.26	—	—	—	—
	Reading							
	High frequency				Low frequency			
	x	y	z	*cc*	x	y	z	*cc*
S9	−49	−27	−2	0.50	−49	−27	−2	0.52
S10	−58	−23	3	0.56	−58	−23	3	0.54
S11	−47	−41	9	0.50	−49	−35	9	0.40
S12	−57	−41	3	0.60	−56	−36	3	0.44
S13	−50	−29	6	0.36	−61	−51	12	0.40
S14	−50	−55	9	0.32	−50	−53	9	0.32
S15	−50	−69	9	0.26	−53	−71	9	0.30
S16	−44	−47	9	0.40	−47	−49	9	0.40

primarily by the extent of prior experience with the stimuli.

Thompson-Schill's (Thompson-Schill *et al.*, 1997) findings from a previous investigation are also relevant to the retrieval effort account we propose. In that study, greater left prefrontal activation was observed when volunteers had to match words along a specific dimension (e.g., color) than when any available information could be used. This result was interpreted as showing the effect of increased selection demands on prefrontal activity. However, an alternative interpretation of this finding is that increased activation occurs whenever responses require access to less common semantic features and that this (access) requires more effort (see also Roskies *et al.*, 2001). For example, in the "match-to-color" ("high selection demand") version of that experiment, "bone" was the correct match for "tooth" instead of "mouth." Given that color may be a less commonly invoked semantic feature of these items than anatomical location, this information could require more effort to retrieve.

Frequency and Effort

High-frequency words are encountered more often in print and appear in more contexts than low frequency words (Nelson and McEvoy, 2000). As a result of fewer encounters, low-frequency items may either be represented more sparsely in fewer neuronal arrays, requiring more effort to access (in a multiple memory trace model) (Hintzman, 1988) or accessed through a less well-specified or tuned manner as suggested by data from prefrontal microelectrode recordings (Rainer and Miller, 2000). In either case, retrieval of sufficient information to make a decision will require more "effort" for low frequency words even though the specific information required for decision-making is likely to be the same as that for high frequency words.

The relatively greater left prefrontal activation associated with retrieval of semantic information for low frequency items might correlate to their enhanced recognition in future probes. In a recent study, volunteers studied a list of words prior to undergoing scanning. While being scanned, they were given a mixture of novel and previously studied words and were told to make old/new judgments. Postscan, a surprise test was administered to evaluate recognition of the novel words shown during the scanning phase of the experiment and fMRI responses were sorted according to whether the novel words were successfully recognized or not (Buckner *et al.*, 2001). The ability of volunteers to make this second round of old/novel judgments showed that encoding occurred even when the explicit

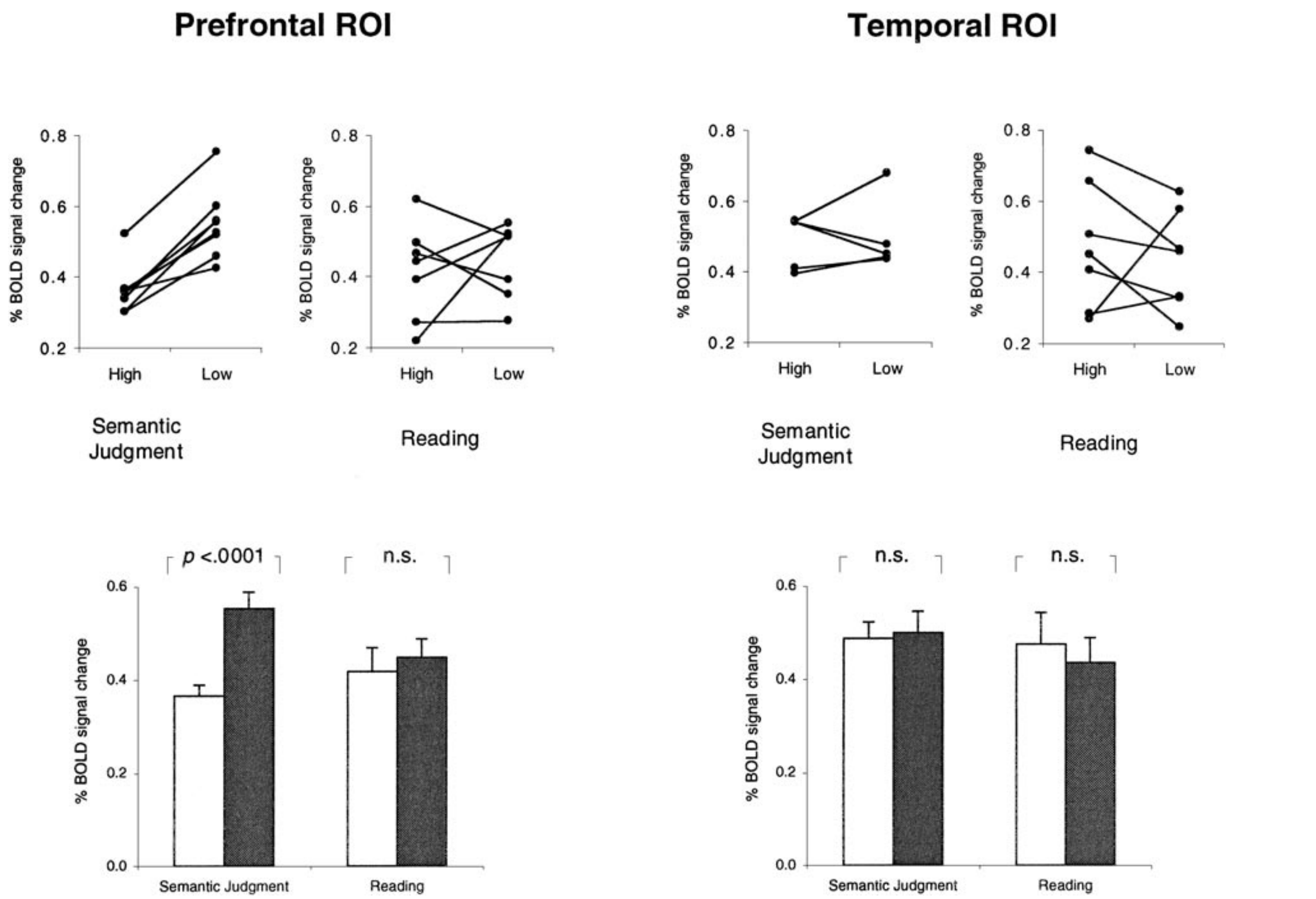


FIG. 5. Line plots showing individual BOLD signal change (in percent) for high and low frequency items when volunteers performed semantic judgment or reading tasks. Bar charts show the mean BOLD signal in each experimental condition. Error bars denote 1 standard error.

instruction was to discriminate between old and new items. Although this study examined episodic as opposed to semantic memory, it is noteworthy that the area shown to have a higher BOLD signal change in the contrast between successfully recognized items and forgotten items (demonstrating incidental encoding) was a subset of the areas activated during the in-scanner discrimination of old and new items (which was ostensibly a retrieval task). In relation to the present study, the higher BOLD signal associated with semantic retrieval of low frequency words may also relate to concurrent encoding processes that facilitate recognition when memory for these words is subsequently probed. Brief exposure to words in the context of incidental encoding may influence long-term memory (Davachi *et al.*, 2001). Stronger incidental encoding following exposure to novel or infrequently encountered words (for example, those in a second language) could in this manner, facilitate neural changes that ultimately result in lower neural activation when semantic retrieval involving identical words is tested after attaining a higher level of language proficiency (Chee *et al.*, 2001). While appealing, this point needs to properly characterized in a longitudinal study.

Neural Correlates of the Word Frequency Effect

While we have demonstrated a strong effect of word frequency on brain activation using a semantic task, this and a previous study (Fiez *et al.*, 1999), did not demonstrate it using a reading task. (Fiez *et al.*, did show effects when frequency and consistency were combined). We propose that deliberate semantic access may be a critical factor in demonstrating the frequency effect. A study that demonstrated frequency-based modulation of neural activity used a sentence verification task that also required deliberate semantic access (Keller *et al.*, 2001). Additionally, event related potential studies have demonstrated significant postlexical effects of frequency when semantic decision tasks are involved (Young and Rugg, 1992) and that the most salient locus of frequency effects may be postlexical. The present findings of greater contrast in activation between high and low frequency words in a semantic judgment task, compared to reading, support this notion.

Additionally, the reading task is probably inappropriate for isolating the neural correlates of the word frequency effect because it does not adequately constrain the type of processing that occurs. In the present study, individual data from the reading task (Fig.5) showed considerable variation in the relative levels of activation, with some people showing greater activation for high frequency words and others showing greater activity for low frequency words. As such, frequency effects at the group level, even if they were present, may have been masked by variances due to heterogeneous processing strategies.

In sum, the present results demonstrate that left prefrontal activation during deliberate semantic retrieval is modulated by word frequency even after controlling for strength of association and selection demands. This adds to the existing evidence pointing to the greater recruitment of this region when retrieval difficulty increases (Buckner and Wheeler, 2001).

ACKNOWLEDGMENTS

This work was supported by NMRC Grants 98/00270, 2000/0477, and the Shaw Foundation. David Caplan received support from NINCD Grant DC02146.

REFERENCES

Balota, D. A., and Chumbley, J. I. 1985. The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *J. Mem. Lang.* **24:** 89–106.

Buckner, R. L., and Wheeler, M. E. 2001. The cognitive neuroscience of remembering. *Nat. Rev. Neurosci.* **2:** 624–634.

Buckner, R. L., Wheeler, M. E., and Sheridan, M. A. 2001. Encoding processes during retrieval tasks. *J. Cogn. Neurosci.* **13:** 406–415.

Chee, M. W., Hon, N., Lee, H. L., and Soon, C. S. 2001. Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. *Neuroimage* **13:** 1155–1163.

Davachi, L., Maril, A., and Wagner, A. D. 2001. When keeping in mind supports later bringing to mind: Neural markers of phonological rehearsal predict subsequent remembering. *J. Cogn. Neurosci.* **13:** 1059–1070.

Demb, J. B., Desmond, J. E., Wagner, A. D., Vaidya, C. J., Glover, G. H., and Gabrieli, J. D. 1995. Semantic encoding and retrieval in the left inferior prefrontal cortex: A functional MRI study of task difficulty and process specificity. *J. Neurosci.* **15:** 5870–5878.

Fiez, J. A., Balota, D. A., Raichle, M. E., and Petersen, S. E. 1999. Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron* **24:** 205–218.

Fletcher, P. C., Shallice, T., and Dolan, R. J. 2000. “Sculpting the response space”—An account of left prefrontal activation at encoding. *Neuroimage* **12:** 404–417.

Forster, K. I., and Chambers, I. M. 1973. Lexical access and naming time. *J. Verb. Learn. Verb. Behav.* **12:** 627–635.

Frederiksen, J. R., and Kroll, J. F. 1976. Spelling and sound: Approaches to the internal lexicon. *J. Exp. Psychol. Hum. Percept. Perform.* **2:** 361–379.

Hintzman, D. L. 1988. Judgments of frequency and recognition memory in a multiple-trace memory model. *Psych. Rev.* **95:** 528–551.

Howard, D., and Patterson, K. 1992. *The Pyramid and Palm Trees Test: A Test of Semantic Access from Words and Pictures,* Thames Valley Test Co., Bury St. Edmunds.

Keller, T. A., Carpenter, P. A., and Just, M. A. 2001. The neural bases of sentence comprehension: A fMRI examination of syntactic and lexical processing. *Cereb. Cortex.* **11:** 223–237.

Kucera, H., and Francis, W. N. 1967. *Computational Analysis of Present Day American English,* Brown Univ. Press, Providence, RI.

Monsell, S. 1991. The nature and locus of word frequency effects in reading. In *Basic Processes in Reading: Visual Word Recognition* (D. Besner and G. W. Humphreys, Eds.), pp. 148–197. Lawrence Erlbaum, Hillsdale.

Nelson, D. L., and McEvoy, C. L. 2000. What is this thing called frequency? *Mem. Cogn.* **28:** 509–522.

Otten, L. J., Henson, R. N., and Rugg, M. D. 2001. Depth of processing effects on neural correlates of memory encoding: Relationship between findings from across- and within-task comparisons. *Brain* **124:** 399–412.

Price, C. J. 2000. The anatomy of language: contributions from functional neuroimaging. *J. Anat.* **197 Pt 3:** 335–359.

Price, C. J., Wise, R. J. S., and Frackowiak, R. S. J. 1996. Demonstrating the implicit processing of visually presented words and pseudowords. *Cereb. Cortex.* **6:** 62–70.

Rainer, G., and Miller, E. K. 2000. Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron* **27:** 179–189.

Roskies, A. L., Fiez, J. A., Balota, D. A., Raichle, M. E., and Petersen, S. E. 2001. Task-dependent modulation of regions in the left inferior frontal cortex during semantic processing. *J. Cogn. Neurosci.* **13:** 829–843.

Seger, C. A., Desmond, J. E., Glover, G. H., and Gabrieli, J. D. 2000. Functional magnetic resonance imaging evidence for right-hemisphere involvement in processing unusual semantic relationships. *Neuropsychology* **14:** 361–369.

Talairach, J., and Tournoux, P. 1988. *Coplanar Stereotaxic Atlas of the Human Brain,* Thieme Medical Publishers, New York.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., and Farah, M. J. 1997. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proc. Natl. Acad. Sci. USA* **94:** 14792–14797.

Wagner, A. D., Pare-Blagoev, E. J., Clark, J., and Poldrack, R. A. 2001. Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron* **31:** 329–338.

Young, M. P., and Rugg, M. D. 1992. Word frequency and multiple repetition as determinants of the modulation of event-related potentials in a semantic classification task. *Psychophysiology* **29:** 664–676.