Behavioral/Systems/Cognitive

Separable Prefrontal Cortex Contributions to Free Recall

Nicole M. Long, Ilke Öztekin, and David Badre

Department of Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, Rhode Island 02912

In everyday life, we often must remember the past in the absence of helpful cues in the environment. In these cases, the brain directs retrieval by relying on internally maintained cues and strategies. Free recall is a widely used behavioral paradigm for studying retrieval with minimal cue support. During free recall, individuals often recall semantically related items consecutively—an effect termed semantic clustering—and previous studies have sought to understand clustering to gain leverage on the basic mechanisms supporting strategic recall. Successful recall and semantic clustering depend on the prefrontal cortex (PFC). However, as a result of methodological limitations, few functional magnetic resonance imaging (fMRI) studies have assessed the neural mechanisms at encoding that support subsequent recall, and none have tested the event-related correlates of recall itself. Thus, it remains open whether one or several frontal control mechanisms operate during encoding and recall. Here, we applied a recently developed method (Öztekin et al., 2010) to assess event-related fMRI signal changes during free recall. During encoding, dorsolateral prefrontal cortex (DLPFC) activation was predictive of subsequent semantic clustering. In contrast, subregions of ventrolateral prefrontal cortex (VLPFC) were predictive of subsequent recall, whether clustered or nonclustered, and were inversely associated with clustering during recall. These results suggest that DLPFC supports relational processes at encoding that are sufficient to produce category clustering effects during recall. Conversely, controlled retrieval mechanisms supported by VLPFC support item-specific search during recall.

Introduction

Often, we must remember the past without helpful cues in the environment. Free recall (FR) is a behavioral paradigm that tests such internally driven retrieval. During free recall, participants overtly report previously studied words in any order without external cues (Moscovitch, 1994; Stuss et al., 1994; Gershberg and Shimamura, 1995; Moscovitch and Winocur, 2002; Becker and Lim, 2003). Participants often recall semantically related items consecutively, an effect termed semantic clustering (Bousfield, 1953). Organizational effects, such as semantic clustering, have been used to study mechanisms giving rise to recall. Differing accounts have attributed such effects to associative (i.e., automatic) versus strategic (i.e., controlled) mechanisms operating at encoding, retrieval, or both (Tulving, 1962; Hunt and Einstein, 1981; Moscovitch, 1994; Gershberg and Shimamura, 1995; Becker and Lim, 2003; Cinan, 2003; Alexander et al., 2009; Polyn et al., 2009). Thus, the source of organizational effects in memory and the neural systems that support them remains an important open question.

Neuropsychological, EEG, and positron emission tomography (PET) studies of recall (Jetter et al., 1986; Janowsky et al., 1989; Stuss et al., 1994; Hildebrandt et al., 1998; Savage et al., 2001; Sederberg et al., 2003, 2007) have broadly implicated pre-frontal cortex (PFC) control mechanisms in clustering and strategic recall. However, methodological limitations constrain our ability to resolve whether one or several frontal control mechanisms operating at encoding and/or retrieval contribute to recall and clustering. The contribution of functional magnetic resonance imaging (fMRI) to these questions has been limited, and event-related fMRI data have come entirely from encoding (Strange et al., 2002; Staresina and Davachi, 2006). To address this gap, we directly compared event-related fMRI activation at encoding and recall.

We sought to contrast frontal contributions at encoding and recall as well as determine the underlying nature of these mechanisms. We focused on three previously implicated control mechanisms and their associated brain regions. (1) Clustering at recall may arise from relational strategies used at encoding (Hunt and Einstein, 1981; Troyer et al., 1998; Cinan, 2003). Regions sensitive to relational processing should demonstrate greater encoding activation for subsequently clustered items over items recalled but not clustered, but no such effect should be evident at recall. We predicted that right dorsolateral prefrontal cortex (DLPFC) would exhibit this pattern given its previous association with working memory mechanisms that give rise to relational encoding (Blumenfeld and Ranganath, 2006; Murray and Ranganath, 2007). (2) Alternatively, clustering could arise from strategic semantic search (Raaijmakers and Shiffrin, 1981), resulting in greater activation of clustered over nonclustered items at recall. We predicted that left anterior (aVLPFC) and mid-ventrolateral (VLPFC) PFC might show this pattern given their association with control and selection processes during semantic retrieval (Badre and Wagner, 2007). (3) Finally, nonclustered items may also engage semantic search during recall, however, at the item rather than category level. This item-specific semantic search

would result in nonclustered greater than clustered activation at recall, potentially in anterior and mid-VLPFC. The current study provides evidence that frontally mediated relational processes at encoding are sufficient to produce clustering effects at recall, whereas PFC supports item-specific semantic search during recall.

Materials and Methods

Participants

Twenty-eight (16 female) right-handed adults (age, 18–28 years; mean, 22 years) enrolled in the study. All had normal or corrected-to-normal vision and were native English speakers. In addition, all participants were screened for use of CNS-affecting drugs, for psychiatric or neurological conditions, and for contraindications for MRI, such as implanted metal. Participants gave written informed consent according to guidelines established and approved by the Human Research Protections Office of Brown University and were compensated for their participation. Four participants were excluded, two for technical issues with the MRI scanner and two for excessive head motion.

Procedure

There were four phases within each of five scan runs: encoding, distractor, free recall, and recognition (Fig. 1a). Participants were given specific instructions about all four phases before entering the magnet. Instructions presented on the screen during scanning cued subjects when to perform the recall and recognition phases.

Encoding. During the encoding phase, participants judged whether the meanings of words presented one at a time in the center of the screen were pleasant or unpleasant. Participants made their response by pressing one of two keys on a button box. On each trial, a single uppercase word was presented for 1000 ms, followed by an additional 1000 ms of fixation (Fig. 1b). Participants could respond during the word presentation or the fixation that followed. If participants responded during the word presentation, the word was replaced with a fixation cross for the remainder of the first 1000 ms. Trials were separated by jittered fixation-null events (0–8 s). The duration and order of the null events were determined by optimizing the efficiency of the design matrix so as to permit estimation of event-related hemodynamic response (Dale, 1999).

The key manipulation during encoding was whether the words came from the same or a distinct semantic category than other words within that block. Half the words came from the same category (CAT words) (e.g., fruits) and half the words did not come from any common category (NCAT words); all were concrete nouns balanced across CAT/NCAT for length and frequency. CAT words were taken from normative data (Battig and Montague, 1969). Different categories were used in each run so that categories did not repeat across runs. CAT and NCAT words were pseudorandomly organized during encoding such that a CAT word was never followed by another CAT word (Stuss et al., 1994). Interleaving CAT and NCAT words ensures that any effects of clustering seen at retrieval are attributable to semantic and not immediate temporal contiguity during encoding. None of the words were repeated at any point in the experiment. Participants were not informed of the semantic category in each study block, although post-task questionnaires revealed that all participants were aware of the semantic categories.

Distractor. After encoding, participants solved math problems during a 2.5 min distractor task. Math problems required the addition and/or subtraction of three two- or three-digit numbers (e.g., 22-87-100 = −165?). Math problems were presented for 6 s during which time participants indi-
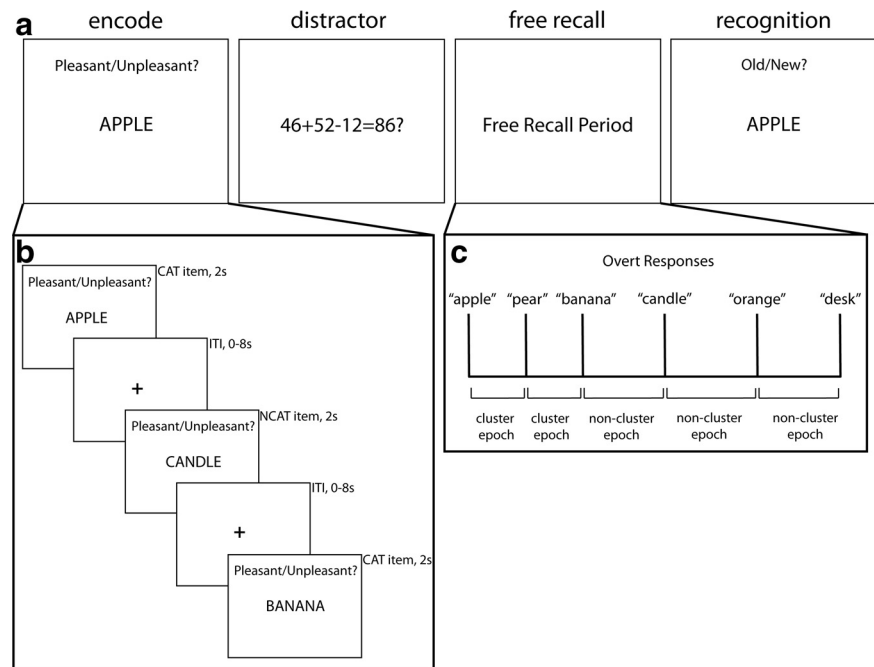
Figure 1. Schematic depiction of experimental events. *a*, During each scan run, participants encountered encoding, distractor, free recall, and recognition phases. *b*, During each encoding trial, participants saw a word for 1000 ms, followed by a green fixation cross for an additional 1000 ms. Participants could respond both during the word presentation and when the green fixation was present. Trials were separated by jittered (0–8 s) intertrial interval (ITI), during which a red fixation cross was presented. *c*, An example of how free recall epochs were modeled for statistical analysis. An epoch was modeled from the beginning of the offset of the previous word (e.g., "apple") to the onset of the following word (e.g., "pear"). Three consecutive same-category words constituted a cluster, and therefore each epoch preceding those words was modeled as a cluster epoch. Noncategory words and category words that were not part of a same-category group of three were preceded by noncluster epochs.

cated whether the presented solution was correct using one of two keys on a button box. Participants could respond at any time in the 6 s during which the problem was on the screen. The null intervals between math problems ranged from 2 to 16 s to make the onset of recall (such as always after a long delay) unpredictable and so to prevent recall preparation.

Free recall. A tone presented to the participants through headphones indicated the beginning of the FR period, at which time participants recalled any words from the most recent word list (and not word lists from other runs) in any order they chose. Verbal responses were recorded using an Avotec SS-3100 Silent Scan Audio System. The FR phase was 60 s long, and participants were encouraged to continue trying to free recall throughout the entire phase. Importantly, previous work from our laboratory, using both empirically and theoretically derived recall latency distributions, has specified the circumstances under which the natural jitter between recall events conveys sufficient design efficiency to estimate the event-related hemodynamic response (Öztekin et al., 2010). As described in the analysis section below, all free recall onset distributions were determined to meet these conditions.

Recognition. In the final recognition phase, participants were presented with all 30 study words from the most recent list (old words) as well as 30 unstudied words (new words). The 30 new words were matched in length and frequency to the 30 old words. In addition, 15 of the 30 new words came from the same category as the 15 studied category words. In each trial, a single uppercase word was presented for 2000 ms, during which time participants had to respond whether the word was old or new by pressing one of two buttons. Trials were separated by jittered fixation-null events (0–8 s), again determined using an optimizing algorithm. This recognition test was used to ensure that better recall of CAT than NCAT words resulted from clustering at FR and not a chance benefit at encoding (Stuss et al., 1994).

Each run followed this four-phase format. Five categories were used across the five runs of the experiment: fruits, tools, animals, sporting activities, and professions. Category order across runs, category repre-

sentativeness of CAT words, key press mappings, and old/new word sets were counterbalanced across participants.

Analysis of free recall behavior

Recordings. A spectral subtraction algorithm, part of Matlab Voicebox toolbox (http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html), was used to subtract scanner noise from the recordings. Each recording underwent two iterations of the algorithm, because additional iterations of the algorithm produced only minimal additional gains. Three independent raters listened to each participant's FR recordings and wrote down the recorded responses. A minimum of two raters had to agree on the identity of any response for it to be scored. In cases in which none of the raters agreed, the word was excluded from correct-trial analyses (see below). At least two raters agreed on 91% of trials.

Cluster scores. We used a modified version of the list based clustering index (LBC) (Stricker et al., 2002) to calculate cluster scores. The LBC is commonly used to score the California Verbal Learning Test (CVLT) (Delis et al., 1987). The LBC calculates an expected cluster value, the number of clusters expected by chance, based on number of words recalled (r), number of words in a category (m), and total number of words studied (NL): $\text{EXP} = [(r - 1) \times (m - 1)]/(NL - 1)$. The LBC score is calculated by subtracting the expected value given in the equation from the observed number of clusters, which is just the total number of consecutive same CAT pairs. In the current experiment, only half the studied words (CAT words) come from a common category and hence can be clustered; therefore, we modified the LBC by dividing the expected value by 2. With the modified LBC equation, the maximum score is 10.62 ($r = 15$, observed = 14), and the minimum score is -7 ($r = 30$, observed = 0).

Onset timing. Tau values were estimated for each individual participant to examine whether free recall latencies were adequately suited to fMRI analysis (Öztekin et al., 2010). The *tau* parameter was extracted from the best fit of an ex-Gaussian to the free recall data (Rohrer and Wixted, 1994). The ex-Gaussian function is used because it best approximates the distribution of response latencies during free recall, with the initiation period preceding recall of the first item corresponding to the Gaussian portion and the words that follow representing the exponential portion (Rohrer and Wixted, 1994). The *tau* parameter, which characterizes the exponential portion of the ex-Gaussian, represents an ongoing memory search process during recall, which decays exponentially. Accordingly, *tau* determines the lags between recall events (i.e., greater *tau* values produce greater variability between lags) and therefore influences fMRI design efficiency. We used an ex-Gaussian model fit to each participant's free recall latencies and estimated *tau* to ensure that our data result from the actual events and not spurious findings. We also estimated *tau* for all participants combined, e.g., a single large distribution of all the participants' free recall responses. The results of this estimation produced a range of *tau* values (13.25–21.5) that fall within the range of *tau* values previously associated with high design efficiency (Öztekin et al., 2010).

fMRI procedure

Whole-brain imaging was performed on a Siemens 3 T TIM Trio MRI system. Functional images were acquired using a gradient-echo echo-planar sequence [repetition time (TR), 2 s; echo time (TE), 30 ms; flip angle, 90°; 33 axial slices; $3 \times 3 \times 3.5$ mm]. After the five functional runs, high-resolution T1-weighted (magnetization-prepared rapid-acquisition gradient echo) anatomical images were collected for visualization (TR, 1900 ms; TE, 2.98 s; flip angle, 9°; 160 sagittal slices; $1 \times 1 \times 1$ mm). Head motion was restricted using firm padding that surrounded the head. Visual stimuli were projected onto a screen and viewed through a mirror attached to a matrix eight-channel head coil. Free recall speech was recorded using a microphone and the program Audacity (http://audacity.sourceforge.net).

fMRI data analysis

Preprocessing and data analysis were performed using SPM2 (http://www.fil.ion.ucl.ac.uk/spm/). Following quality assurance procedures to assess outliers or artifacts in volume and slice-to-slice variance in the global signal, functional images were corrected for differences in slice acquisition timing by resampling all slices in time to match the first slice.

Images were then motion corrected across all runs (using sinc interpolation). Functional data were then normalized based on Montreal Neurological Institute stereotaxic space using a 12-parameter affine transformation along with a nonlinear transformation using cosine basis functions. Images were resampled into 2 mm cubic voxels and then spatially smoothed with an 8 mm full-width at half-maximum isotropic Gaussian kernel.

Data analysis was conducted under the assumptions of the general linear model as implemented in SPM2. Four separate statistical models were applied to the data, and each was designed to address separate aspects of the dataset. (1) A general subsequent memory model was intended to mirror the approach taken by previous studies (Staresina and Davachi, 2006). This model was applied only to encoding data and included the subsequently recalled and subsequently non-recalled items, collapsing across clustered conditions. (2) We applied a subsequent clustering model that separated free recalled items into subsequently clustered and nonclustered conditions, collapsing across CAT and NCAT items as there were relatively few CAT nonclustered items across participants (six on average). (3) We applied a subsequent clustering model that separated free recalled items into CAT words subsequently clustered, CAT words subsequently nonclustered, and NCAT words subsequently nonclustered. This model was applied to 10 participants who had at least five CAT noncluster and five NCAT noncluster words and was used to rule out concern that a general category effect might drive cluster-related activation. (4) We applied a free recall and recognition model that included only data from recall and recognition and included regressors for items that were and were not clustered during recall and items that were and were not recognized. All regressors were generated by convolving each epoch with a canonical hemodynamic response function and its temporal derivative. The five runs were concatenated into a single run to have sufficient trials for analysis. The six motion parameters (translation and rotation) were also modeled as nuisance regressors to account for motion during scanning along with other nuisance regressors, included to model out low-frequency signal components, such as those attributable to linear drift.

Subsequent memory model, without clustering. Words studied at encoding could subsequently be recognized, free recalled, and/or clustered. In the first subsequent model, we modeled only words subsequently free recalled or recognized, without delineating across words subsequently clustered. This model was used to test subsequent memory effects observed in other free recall studies (Staresina and Davachi, 2006) without having those effects compromised by the estimation of specific clustering trials. Separate regressors were created for words based on subsequent free recall and recognition, with each encoding trial onset modeled as an epoch lasting 2 s to encompass the duration of the trial.

Subsequent memory model, with clustering. In the second subsequent model, words were modeled according to whether they were subsequently clustered, recalled, and/or recognized. This model was used to test subsequent memory effects specific to clustering. All combinations of the three potential outcomes were modeled as separate regressors, with each encoding trial onset modeled as an epoch lasting 2 s to encompass the duration of the trial.

Subsequent memory model, subset analysis. In the final subsequent model, words were modeled according to whether they were CAT or NCAT, subsequently clustered, recalled, and/or recognized. This model was used to test subsequent memory effects specific to CAT clustering and CAT nonclustering and was run on 10 participants who had at least five CAT noncluster and five NCAT noncluster items. All combinations of the three potential outcomes were modeled as separate regressors, with each encoding trial onset modeled as an epoch lasting 2 s to encompass the duration of the trial.

Free recall and recognition model. For fMRI analysis (in all statistical models), the coding of words during FR as clusters was different than the LBC method. Specifically, when at least three same category words appeared consecutively, all correct words in the group were coded as clusters. We chose to be more conservative in our definition of clusters for the fMRI analysis to ensure that signal associated with clusters was less likely to be attributable to chance recall. Requiring three words of the same category to be recalled contiguously rather than two meant that an average of four less words were coded as clustered than would have been had

we used a less conservative cluster definition. Incorrect CAT words were included in the determination of the clusters but were excluded from the final fMRI analysis. Although incorrect words were modeled separately for fMRI (see below), the inclusion of error trials in the cluster definition is because any CAT word can act as a categorical cue for other words. Therefore, the same semantic information would be available as a cue for correct CAT words, whether preceded by an incorrect or correct CAT word. Across all participants, 0.6% of incorrect CAT words occurred in a cluster of only three words.

Separate regressors were generated for each free recall condition: cluster, noncluster, and incorrectly recalled words, as well as two parametric regressors capturing the output position of clustered and nonclustered words. NCAT and nonclustered CAT words were combined in the noncluster condition because there were too few CAT nonclustered words to be estimable on its own (on average, participants recalled six CAT nonclustered words across the entire experiment). At any point during FR, participants could be verbally responding and/or attempting to recall a word. To model out the verbal response itself, all spoken words were modeled as epochs whose onset and duration matched those of the actual word, without respect to CAT/NCAT membership. Separate cluster, noncluster, and incorrect events were modeled using epochs onsetting at the offset of the previous verbal response and offsetting before the onset of the next verbal response event (Fig. 1c). It is important to note that additional mechanisms supporting recall may be sustained throughout the recall period (Howard and Kahana, 2002; Polyn et al., 2005) or may be temporally misaligned with the recovery of individually recalled items. However, the approach in the present experiment reflects our interest in any processes that differ between cluster and noncluster events. Processes occurring at other timescales across these conditions are interesting, but their estimation would require a hybrid design not implemented in this initial experiment.

Recognition regressors were constructed from the crossing of hits, misses, false alarms (FAs), and correct rejections (CRs) with CAT/NCAT conditions. Recognition events were modeled as 2 s epochs starting at the onset of the target word to encompass the duration of the trial.

Voxelwise contrasts. Statistical effects were estimated using a subject-specific fixed-effects model, with session-specific effects and low-frequency signal components (<0.01 Hz) treated as confounds. Linear contrasts of the whole brain were used to obtain subject-specific estimates for each effect. These estimates were entered into a second-level analysis treating subjects as a random effect, using a one-sample t test against a contrast value of zero at each voxel. Voxel-based group effects were considered reliable to the extent that they consisted of a cluster of at least five contiguous voxels that exceeded an uncorrected threshold of $p < 0.001$.

Region of interest analysis. Region of interest (ROI) analyses complemented whole-brain voxelwise contrasts to test predicted effects in a priori defined regions. First, we selected a set of ROIs functionally defined from the current dataset. Foci were functionally defined based on activated voxels in a task effects contrast. All significant voxels within an 8 mm radius of a chosen focus defined an ROI. Specifically, we chose ROIs in right DLPFC (BA 9; $x, y, z = 50, 36, 30$) and left anterior VLPFC (BA 47; $x, y, z = -56, 31, -7$), given their previous association with relational encoding and semantic search mechanisms, respectively. We also chose an ROI in medial temporal lobe (MTL) ($x, y, z = -20, -28, -10$), given its role in subsequent memory and memory strength. Second, to draw a tighter link with the previous literature, we defined a set of ROIs using coordinates from previous fMRI studies that tested analogous mechanisms outside of the context of free recall. Specifically, we defined ROIs in DLPFC from Murray and Ranganath (2007) ($x, y, z = 49, 38, 32$) from a manipulation of relational encoding, left anterior VLPFC (BA 47; $x, y, z = -51, 27, -3$) from a manipulation of controlled semantic retrieval (Badre et al., 2005), and mid-VLPFC (BA 45; $x, y, z = -54, 30, 12$) from a manipulation of post-retrieval selection (Badre et al., 2005).

Selective averaging extending 16 s after stimulus onset allowed assessment of the time-dependent signal change associated with each condition. Integrated percentage signal change (iPSC) was then computed based on the peak plus and minus one TR. The peak was defined neutrally
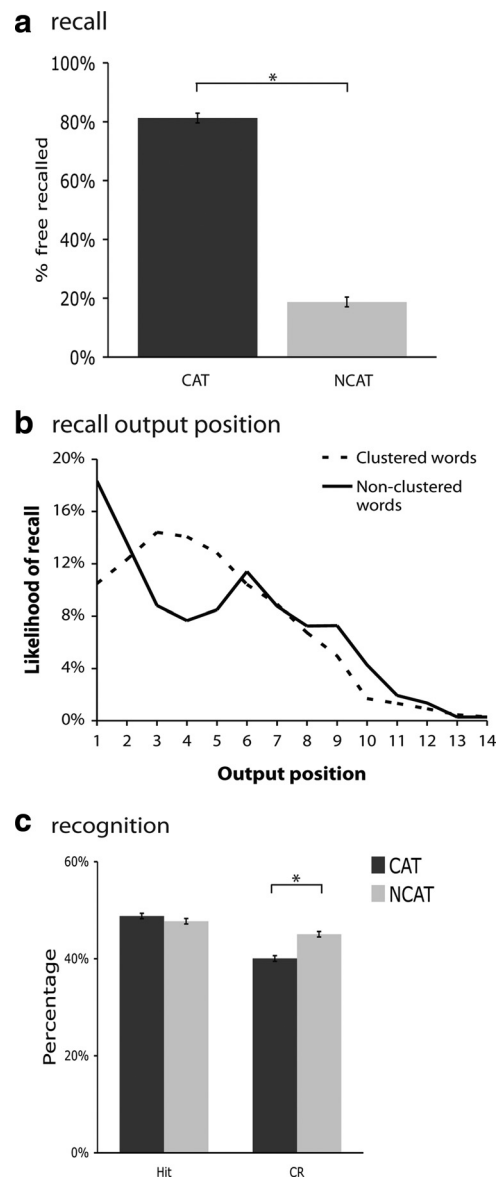

a recall

b recall output position

c recognition

Figure 2. Behavioral free recall results. *a*, The percentage of CAT and NCAT words recalled are plotted. CAT words were remembered better than NCAT words (*$p < 0.05$). *b*, Likelihood of recalling a cluster or noncluster word in a specific output position. *c*, The proportion of CAT and NCAT hits and CRs are plotted. Hit rates were equal across CAT and NCAT words, whereas NCAT words had greater CR rates than CAT words (*$p < 0.05$).

for each ROI based on the average time course. The resultant data were subjected to repeated-measures ANOVA.

Results

Behavioral evidence of semantic clustering

Participants took advantage of the common semantic category and clustered during free recall. On average participants recalled 27% of all words studied. Recall was better for CAT (81%) than NCAT (19%) words ($t_{(23)} = 18.9, p < 0.0001$) (Fig. 2a). Importantly, participants clustered CAT items more than would be expected by chance using both the LBC (mean cluster score, 2.6; one-sample $t_{(23)} = 16.9, p < 0.0001$) and the modified LBC for fMRI (mean cluster score, 2.5; one-sample $t_{(23)} = 13.2, p < 0.0001$). The average number of CAT clustered trials was 26, CAT nonclustered trials was 6, and NCAT nonclustered trials was 8.

The specific categories (e.g., fruits, tools, etc.) did not differ in numbers of clustered words, suggesting no category effects (for additional analysis by category, see supplemental material, available at www.jneurosci.org).

To investigate the potential role of output position on free recall condition (cluster or noncluster), we computed the likelihood of a cluster or noncluster word being recalled (of all words in that condition) in each output position (1 to 14) of the total number of cluster or noncluster words recalled. An omnibus test found a main effect of position ($F_{(13,23)} = 62.9$, $p < 0.0001$) and a significant interaction of output position and cluster type ($F_{(13,23)} = 3.2$, $p < 0.0001$) but no main effect of cluster condition ($F_{(1,23)} = 0.65$). In light of this result, we investigated output position as a parametric regressor for the fMRI analysis of recall.

A post-recall recognition test was used to ensure that better recall of CAT than NCAT words resulted from clustering at FR and not a chance benefit at encoding (Stuss et al., 1994). CAT words were no more likely to be recognized than NCAT words. Specifically, hit rates for CAT and NCAT words (Fig. 2b) were 95 and 96%, respectively, and did not statistically differ ($t_{(23)} = 1.3$). False alarm (FA) rates were higher for CAT (18%) than NCAT (9%) words ($t_{(23)} = 5.9$, $p < 0.0001$). Reaction time for correct rejections was reliably slower than for hits ($F_{(1,23)} = 55.9$, $p < 0.0001$), and there was an interaction such that CAT correct rejections were slower than NCAT correct rejections ($F_{(1,23)} = 28.1$, $p < 0.0001$). The CAT/NCAT differences in FA rates and reaction time for correct rejections are consistent with previous studies showing higher false recognition for words semantically related to a study list (Roediger and McDermott, 1995; Nessler et al., 2001). Although the recognition phase was scanned, these results were not central to the hypotheses under consideration across encoding and free recall. These data are provided in the supplemental material (available at www.jneurosci.org).

CAT words were more likely to be rated pleasant than NCAT words ($F_{(1,23)} = 51.7$, $p < 0.0001$). Importantly, however, this difference did not account for the subsequent clustering of CAT words at recall. Specifically, CAT words rated pleasant at encoding were no more likely to be clustered than CAT words rated unpleasant at encoding ($t_{(23)} = 1.6$).

Finally, temporal or serial clustering occurs when participants show a greater than chance likelihood of recalling words in the order in which those words were studied (Howard and Kahana, 2002). Although not the focus of the current study, the mechanisms of temporal clustering are important and may share features with those supporting semantic clustering (Polyn et al., 2009). Thus, we calculated serial clustering scores for each participant using the list-based serial clustering index (Stricker et al., 2002). From this index, the average serial clustering score was −0.05, which did not differ significantly from zero ($t_{(23)} = 0.1$). It might be the case that semantic clustering somehow interferes with serial clustering and that NCAT words were in fact being clustered. To evaluate this possibility, we calculated separate serial cluster scores for NCAT words and CAT words alone. The average serial clustering score for NCAT items (cluster score, −0.005) was not significantly different from zero ($t_{(23)} = 0.37$). The average serial clustering score for CAT items (cluster score, 0.42) was significantly different from zero ($t_{(23)} = 5.3$, $p < 0.0001$), which suggests an interaction between semantic relatedness and temporal contiguity, potentially consistent with previous models of free recall (Polyn et al., 2009).
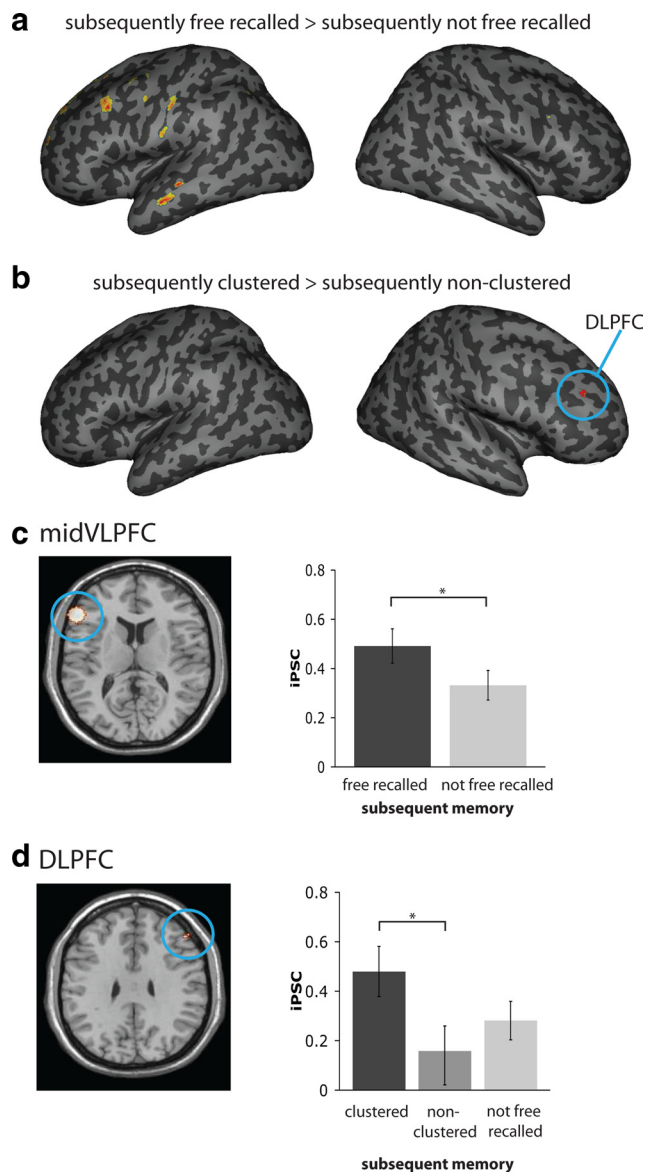

Figure 3. Voxelwise and ROI results at encoding. *a*, The contrast of subsequently free recalled greater than subsequently not free recalled words revealed activation in left middle temporal gyrus, left superior temporal sulcus, left caudate, left superior frontal sulcus, and left DLPFC. *b*, The contrast of subsequently clustered greater than subsequently nonclustered words yielded a distinct pattern of activation from that observed for subsequent recall. Activation was observed in right temporal pole and right DLPFC. *c*, ROI analysis of mid-VLPFC (*x*, *y*, *z* = 54, 30, 12) (Badre et al., 2005) showed greater activation at encoding for words subsequently recalled than words subsequently not recalled. *d*, ROI analysis of DLPFC demonstrated activation for subsequently clustered words greater than subsequently nonclustered words (*$p < 0.05$).

fMRI analysis of subsequent recall and subsequent clustering

To assess which regions were associated with subsequent recall and/or subsequent clustering at recall, events at encoding were coded in the fMRI analysis depending on their outcome at recall. A voxelwise contrast of all words subsequently free recalled compared with those not free recalled, regardless of whether they were members of a cluster, revealed activation in left middle temporal gyrus (−58, −10, −18), left superior temporal sulcus (−58, −20, −4), left caudate (−6, −8, 14), left superior frontal sulcus (−20, 56, 20), and left DLPFC (−48, 20, 38) (Fig. 3a, Table 1). Analysis of the mid-VLPFC ROI (BA 45; *x*, *y*, *z* = −54, 30, 12) revealed a significant subsequent free recall effect ($t_{(23)} = 2.2$, $p < 0.05$) (Fig. 3c). There was no subsequent free recall effect

Table 1. Task activations

Coordinates (x, y, z)	Location	BA	Peak Z
Contrast: subsequently free recalled > subsequently not free recalled			
−58, −10, −18	Middle temporal gyrus	21	3.05
−58, −20, −4	Superior temporal sulcus	20	3.17
−6, −8, 14	Caudate		3.72
−20, 56, 20	Superior frontal sulcus	10	3.31
−48, 20, 38	DLPFC	9	3.11
Contrast: subsequently clustered > subsequently not clustered			
42, 48, 32	DLPFC	9	4.05
42, 2, −12	Temporal pole	28/35	3.88
Contrast: noncluster > cluster			
16, 22, 4	Caudate		3.43
40, −38, 10	Superior temporal gyrus	22	4.31
60, −26, 48	Postcentral gyrus	5/40	4.38

(t values < 0.2) in either anterior ventrolateral PFC ROI (anterior VLPFC; BA 47).

Locating activation in mid-VLPFC predictive of subsequent recall might be initial evidence in favor of the semantic search hypothesis. However, it remains important to establish whether this subsequent memory effect is related to category-level search giving rise to clustering or some other form of semantic elaboration at encoding (such as item-specific processing). Thus, we tested a second subsequent model that included subsequently clustered and subsequently recalled but not clustered regressors. The voxelwise contrast of subsequently clustered over nonclustered items revealed a distinct set of regions from those associated with general subsequent free recall (Fig. 3b, Table 1). In particular, activation was greater in right DLPFC (42, 48, 32) and right temporal pole (42, 2, −12) (Fig. 3b) for subsequently clustered than subsequently nonclustered words. Consistent with the whole-brain analysis, ROI analysis of DLPFC using our functionally defined ROI (BA 9; $x, y, z = 50, 36, 30$) and the DLPFC ROI defined from Murray and Ranganath (2007) ($x, y, z = 49, 38, 32$) showed reliably greater activation for subsequently clustered over nonclustered items (t values >2.2, p values <0.05) (Fig. 3d). However, neither right DLPFC ROI showed a general subsequent recall effect across clustered and nonclustered conditions (t values <1.8). In contrast to DLPFC, neither mid-VLPFC nor anterior VLPFC differed between subsequently clustered and subsequently nonclustered words (t values <1.5).

To ensure that the difference in DLPFC activation between clustered and nonclustered words was not driven by a general category effect, we further broke down subsequently nonclustered into CAT and NCAT words that were subsequently nonclustered (all clustered words were CAT). Across participants, there were relatively few CAT noncluster items preventing separation of these conditions in the full group. However, we selected the subset of 10 participants who had at least five CAT noncluster items and five NCAT noncluster items for additional analysis. Importantly, the voxelwise contrast of subsequently clustered CAT items versus subsequent nonclustered CAT items in this subset revealed right DLPFC activation ($x, y, z = 52, 42, 24$). This result indicates that the difference observed between clustered and nonclustered items relates to clustering and not a general category effect, per se. Moreover, the voxelwise contrast of subsequently clustered versus subsequently nonclustered items (the original contrast) in these 10 participants showed

similar DLPFC activation ($x, y, z = 42, 48, 32$) as shown with the full group of 24 participants, suggesting that this subset is not qualitatively different from the group in their clustering effects.

To summarize the results from encoding, right DLPFC showed a subsequent clustering effect but not a general subsequent recall effect. Right DLPFC also showed greater activation for subsequent CAT cluster items than subsequent CAT noncluster items in a subset analysis. In comparison, anterior VLPFC showed neither a subsequent clustering nor subsequent recall effect. In contrast, mid-VLPFC showed a subsequent recall effect but not a subsequent clustering effect. The difference between DLPFC and mid-VLPFC was confirmed by a reliable region by subsequent clustering effect interaction ($F_{(1,23)} = 4.5, p < 0.05$).

fMRI analysis of free recall

Analysis of event-related data at recall indicated a frontal contribution to item-specific search at recall rather than category-level search supporting clustering effects. The voxelwise contrast of clustered greater than nonclustered items at recall did not yield any reliable differences. In contrast, the voxelwise contrast clustered versus nonclustered items at recall revealed activation in right caudate (16, 22, 4) (Fig. 4a), right superior temporal gyrus (40, −38, 10), and right post central gyrus (60, −26, 48) (Table 1). Notably, neither contrast yielded reliable activation in the PFC. However, our hypotheses about aVLPFC (see Introduction) and DLPFC, as well as the association of DLPFC and mid-VLPFC at encoding with subsequent clustering and recall, prompted a specific test of these DLPFC and VLPFC ROIs during FR.

ROI analysis of DLPFC (Fig. 4b) revealed no significant difference between clustered and nonclustered words at recall in either DLPFC ROI (t values < 0.6). Similarly, there was no effect of clustering in mid-VLPFC ($t_{(23)} = 1.0$). In contrast, activation in aVLPFC (Fig. 4c) was reliably greater for nonclustered than clustered words in both the functionally defined ROI and the ROI from Badre et al. (2005) (t values > 2.1, p values < 0.05). This difference between aVLPFC and DLPFC at recall was supported by a reliable region by effect (clustered/nonclustered) interaction ($F_{(1,23)} = 4.6, p < 0.05$). A comparable region × effect (clustered/nonclustered) interaction was evident between aVLPFC and mid-VLPFC ($F_{(1,23)} = 7.5, p < 0.05$). We note that these effects emerged while controlling for output position in the model.

Motivated by the behavioral results, however, we investigated the effect of output position during recall. The voxelwise contrast of output position greater than baseline revealed activation in bilateral orbital frontal cortex (−40, 40, −7; 43, 44, −7), caudate (−7, 15, 7), and putamen (26, −3, 14). We discuss the potential implications of activation of this striatofrontal network below.

To summarize the results from recall, DLPFC showed no difference between clustered and nonclustered words, whereas aVLPFC showed greater activation for nonclustered words than clustered words. This is notably different from encoding, during which DLPFC showed greater activation for subsequently clustered than subsequently nonclustered words, whereas aVLPFC showed no difference in activation. These regional differences between encoding and recall were supported by a phase (encoding/recall) × effect (clustered/nonclustered) interaction ($F_{(1,23)} = 10.5, p < 0.01$) and a region × effect (clustered/nonclustered) interaction ($F_{(1,23)} = 5.5, p < 0.05$).

fMRI analysis of MTL activation

Although the voxelwise contrasts reported above did not reveal reliable activation in the MTLs, we sought to directly test MTL
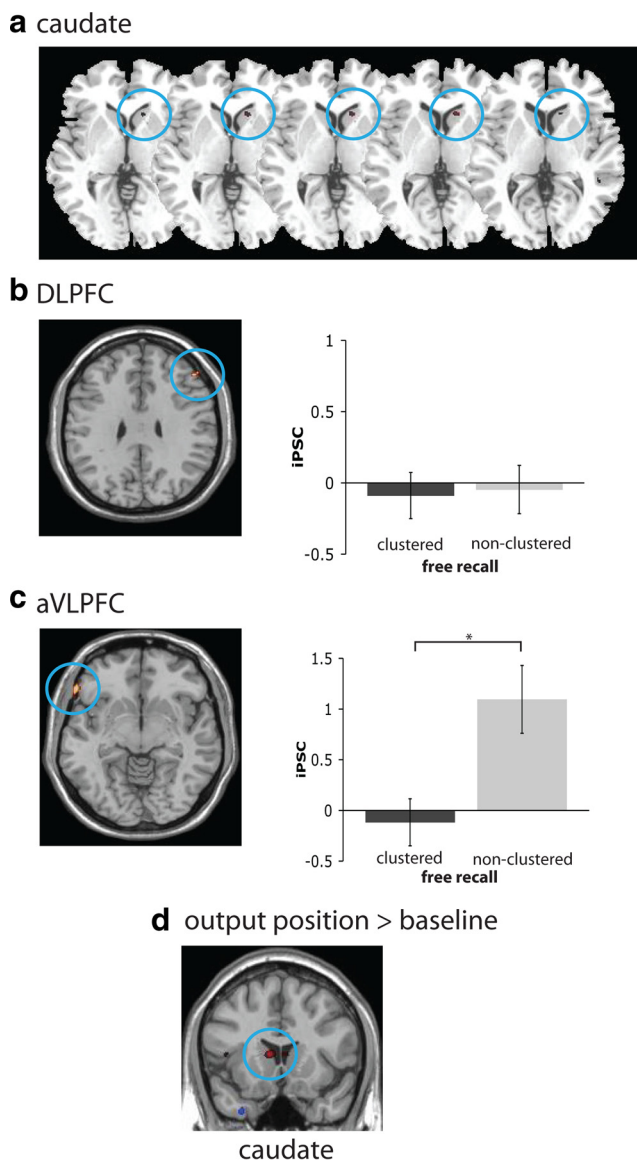
a caudate

b DLPFC

c aVLPFC

d output position > baseline

caudate

Figure 4. Voxelwise and ROI analysis during free recall. *a*, The whole-brain contrast of free recall nonclustered greater than free recall clustered produced activation in right caudate. *b*, No difference in activation at free recall for clustered and nonclustered words was evident in DLPFC. *c*, Anterior VLPFC ($x, y, z = -56, 31, -7$) showed greater activation during free recall for nonclustered words than for clustered words (*$p < 0.05$). *d*, Voxelwise contrast of output position greater than baseline, highlighting activation in caudate ($x, y, z = -7, 15, 7$).

contributions to recall given its role in subsequent memory, specifically subsequent recall (Strange et al., 2002; Eldridge et al., 2005; Staresina and Davachi, 2006). Hence, we tested an ROI in left MTL ($x, y, z = -20, -28, -10$) (Fig. 5a) defined from the contrast of all encoding conditions versus baseline.

MTL demonstrated greater activation for subsequently recalled versus not recalled items ($t_{(23)} = 2.2, p < 0.05$) (Fig. 5a). However, there was not a reliable effect of subsequent clustering in left MTL ($t_{(23)} = 1.1$). It is notable that the quantitative pattern of data in the MTL was similar to that observed in mid-VLPFC, which also showed a subsequent recall but not a subsequent clustering effect. Interestingly, from ROI analysis, both regions showed an apparently graded activation pattern across conditions such that subsequently clustered items were greater than subsequently recalled but nonclustered items, which were in turn greater than unrecalled items (Fig. 5a). Thus, to investigate

whether this graded effect was consistent between mid-VLPFC and MTL, we computed the slope of the increase in activation across the three conditions ("not recalled" to "recalled/nonclustered" to "recalled/clustered") in each region for each participant. Across participants, the slopes in mid-VLPFC were positively correlated with those in MTL ($r_{(23)} = 0.6, p < 0.01$) (Fig. 5b), indicating that the degree of increased activation across these three conditions was related in the two regions.

During the free recall phase, whole-brain analysis did not reveal differential activation in MTL relative to baseline. However, inspection of the time course suggested that this null result might be attributable to a baseline shift. In light of noted concerns about estimates of baseline in the MTL in absence of a neutral task (Stark and Squire, 2001), we used the first and last time points from the time course of signal change in MTL to estimate its baseline. We compared this estimate of baseline with the integrated peak (2–4) and found significantly greater activation at the integrated peak during recall ($t_{(23)} = 2.8, p < 0.01$). However, beyond this general positive change in activation in the MTL during free recall, there were no effects of clustering in this ROI ($t = 0.03$).

Discussion

The goal of the present study was to advance our understanding of the neural mechanisms supporting free recall. Four findings are of central importance. (1) Right DLPFC was associated with subsequent clustering at encoding but not subsequent recall generally or clustering during recall. (2) Mid-VLPFC activation at encoding was associated with subsequent recall but not subsequent clustering. (3) Mid-VLPFC and MTL showed correlated patterns of activation across subsequent recall conditions that decreased across clustered, nonclustered, and unrecalled conditions. (4) Anterior VLPFC was insensitive to subsequent recall or clustering effects at encoding but was more active for nonclustered than clustered items at recall. Together, our results indicate that DLPFC-supported relational encoding mechanisms are sufficient to produce clustering effects during recall, and there is no evidence of category-level semantic search during recall itself under the basic recall conditions used in the current task. Rather, we demonstrate that, during recall, anterior VLPFC shows increased activity for nonclustered over clustered items, an effect potentially consistent with item-based retrieval processing. We now consider the nature of these distinct mechanisms in more detail.

The only region specifically associated with successful clustering was right DLPFC, consistent with previous PET evidence (Savage et al., 2001). Data from studies of nonhuman primates (Petrides, 1994) and fMRI studies in humans have long associated DLPFC with monitoring and/or manipulating the contents of working memory (D'Esposito et al., 1999; Petrides, 2000; Postle et al., 2000; Wagner et al., 2001; Curtis and D'Esposito, 2003; Wager and Smith, 2003). Moreover, previous research has shown DLPFC involvement in deciding whether two items share a common feature (Holyoak and Thagard, 1995; Bunge, 2004; Christoff and Keramatian, 2007). Activation in DLPFC attributable to working memory manipulation has been specifically tied to the formation of subsequent "relational" memories, wherein the association between items must be retrieved rather than the item itself (Blumenfeld and Ranganath, 2006, 2007; Murray and Ranganath, 2007). Hence, the subsequent clustering effect in DLPFC may reflect active attention to the categorical relationship among items during encoding, relating each to the previously encountered items. Creating an elaborated relational trace, this

process increases the likelihood of recalling items sharing the trace (Hunt and Einstein, 1981; Cinan, 2003).

A related hypothesis regarding DLPFC follows from computational models of free recall, specifically the temporal context model and the context maintenance and retrieval model (Howard and Kahana, 2002; Polyn et al., 2009). These models propose that, during encoding, an item and its current context are integrated into a contextual representation that is stored in long-term memory. A retrieved context can cue recall of a specific study item or related item contexts. For multiple items to have overlapping contexts, the context representation must be maintained and slowly updated, a function assumed to be supported by DLPFC (Polyn et al., 2009). Our results are consistent with this view in that DLPFC may be maintaining and integrating a semantic contextual representation specifically during encoding. During retrieval, semantic contexts may be strong enough to permit clustered retrieval to proceed through more automatic, associative mechanisms.

In contrast to DLPFC, mid-VLPFC activation at encoding was predictive of subsequent recall for both clustered and nonclustered items. This effect is consistent with the well established association of mid-VLPFC with subsequent memory during recognition (Wagner et al., 1998) and recall tasks (Brassen et al., 2006; Staresina and Davachi, 2006). Contrasted with DLPFC, mid-VLPFC appears to impact memory strength generally rather than relational processes useful for clustering. As has been suggested previously, mid-VLPFC may affect subsequent memory by focusing attention on details potentially useful for distinguishing that item during subsequent retrieval.

Perhaps consistent with this view, the mid-VLPFC pattern of activation across conditions was graded such that activation to subsequently clustered items appeared greater than subsequently nonclustered items, which were both greater than unrecalled items. This pattern could reflect memory strength or ease of access at retrieval, because items clustered at encoding receive the most elaboration and would therefore have greater memory strength. Notably, a similar graded pattern of activation across subsequent memory conditions was observed in MTL, a region observed previously to exhibit memory strength effects (Strange et al., 2002; Eldridge et al., 2005; Gonsalves et al., 2005; Staresina and Davachi, 2006; Shrager et al., 2008). Importantly, this graded pattern in mid-VLPFC correlated with that in MTL, further supporting the idea that mid-VLPFC activity is related to subsequent memory strength.

In contrast to both mid-VLPFC and DLPFC, anterior VLPFC activation did not correlate with subsequent recall or subsequent clustering. Given previous evidence associating anterior VLPFC with controlled retrieval of goal-relevant semantic knowledge (Badre et al., 2005; Dobbins and Wagner, 2005; Gold et al., 2006), we initially hypothesized that anterior VLPFC might be associated with clustering at encoding and/or retrieval by supporting semantic search (Schneider and Shiffrin, 1977; Raaijmakers and Shiffrin, 1981). For example, at retrieval, participants might strategically search for items from a specific semantic category and
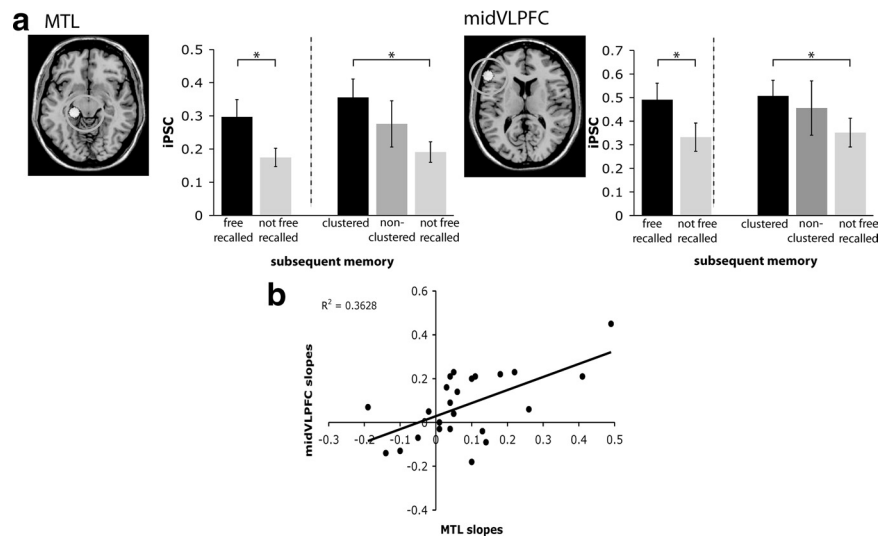

Figure 5. Results from ROI analysis of MTL. **a**, The iPSC in MTL (x, y, $z = -20$, -28, -10) is plotted in the left graph. MTL showed greater activation at encoding for words subsequently free recalled than those not free recalled (left of dashed line). However, MTL showed the graded activation for subsequently clustered, subsequently nonclustered, and subsequently not free recalled words (right of dashed line). For comparison, mid-VLPFC is plotted on the right and showed a similar graded pattern to MTL. **b**, The slope of the activation increase in MTL from not free recalled to free recalled nonclustered to free recalled and clustered (x-axis) is plotted against the slope in activation across these conditions in mid-VLPFC (y-axis). There was a positive and reliable correlation between these slopes across participants ($R = 0.6$) (*$p < 0.05$).

assess the familiarity of each. Such a category fluency strategy would likely require anterior and/or mid-VLPFC (Baldo and Shimamura, 1998; Troyer et al., 1998; Baldo et al., 2001). However, contrary to this hypothesis, anterior VLPFC showed greater activation during recall for nonclustered than clustered items, even with output position controlled.

Anterior VLPFC is engaged when retrieval cannot proceed automatically, such as when associations between cues and target knowledge are weak (Wagner et al., 2001; Badre et al., 2005). Clustered items might be recalled more automatically because of their depth of encoding, whereas weaker nonclustered items are more difficult to retrieve. Consequently, controlled retrieval, focused on item-specific features, is required to recall these weaker items.

We did not find clear effects of frontal laterality between encoding and retrieval. Encoding effects were located in both the left (left mid-VLPFC, left DLPFC, and left hippocampus) and right (right DLPFC) hemispheres. Recall effects were located on the left (left anterior VLPFC). The leftward bias of these activations may be attributable, in part, to our use of verbal materials. However, we particularly implicate right DLPFC in subsequent clustering. Although previous neuropsychological evidence of laterality has been limited, it is notable that at least one neuropsychological investigation highlighted damage to left rather than right DLPFC (BA 9) in association with clustering deficits during CVLT (Alexander et al., 2009). In the present study, we did locate activation associated with subsequent recall (although not clustering) in left DLPFC in the voxelwise analysis (Table 1). However, a neutrally defined ROI in left DLPFC revealed no main effects for either subsequent recall or subsequent clustering (t values <1.4). We do caution, however, that this null result likely reflects the lack of overlap of this ROI with the cluster of activation located in the voxelwise analysis. However, as discussed below, there are components of free recall captured by the CVLT that were not manipulated in the current study and for which this left DLPFC focus might be critical.

The current results indicate that the extent of clustering during recall is likely determined by the degree of relational processing at encoding, as supported by right DLPFC. However, there are other important aspects of organizational effects in memory, untested in the present design. Frontal patients can have deficits not only in the degree to which they retrieve semantically related items consecutively, which in the present study arises entirely from contextual elaboration at encoding, but also in the number of categories along which they cluster (Jetter et al., 1986; Hildebrandt et al., 1998). This latter deficit might reflect a reduced ability to evaluate and/or shift strategies (Becker and Lim, 2003). Unlike the classical CVLT, the present study did not use multiple categories and so could not test the mechanisms involved in this type of strategic switching during recall. Such a mechanism may be frontally mediated, and at least one fMRI study strongly suggests that mid-VLPFC may be particularly critical for this function during fluency tasks analogous to clustered recall (Hirshorn and Thompson-Schill, 2006).

In this context, one potentially intriguing result was the correlation of output position during recall with activation in the striatum and orbital PFC. Considerable evidence has associated frontostriatal circuits in evaluating outcomes associated with control strategies via prediction error signals. Indeed, at least one model of free recall (Becker and Lim, 2003) has suggested that prediction errors during recall may signal when to shift retrieval strategies. As output position increases, latencies also increase between responses. This effect has been assumed in some models (Raaijmakers and Shiffrin, 1981) to reflect repeated retrieval of items that are rejected because they have either already been recovered or were not on the previous list. Hence, later output positions should be associated with greater prediction error, and this might be reflected in the association of striatofrontal activation with output position. However, future research would be necessary to test the relationship of this signal with strategy shifts.

Notwithstanding these potential additional mechanisms, the present study provides evidence for multiple frontal mechanisms acting during encoding and retrieval that support clustering and recall.

References

Alexander MP, Stuss D, Gillingham S (2009) Impaired list learning is not a general property of frontal lesions. J Cogn Neurosci 21:1422–1434.

Badre D, Wagner AD (2007) Left ventrolateral prefrontal cortex and the cognitive control of memory. Neuropsychologia 45:2883–2901.

Badre D, Poldrack RA, Paré-Blagoev EJ, Insler RZ, Wagner AD (2005) Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. Neuron 47:907–918.

Baldo JV, Shimamura AP (1998) Letter and category fluency in patients with frontal lobe lesions. Neuropsychology 12:259–267.

Baldo JV, Shimamura AP, Delis DC, Kramer J, Kaplan E (2001) Verbal and design fluency in patients with frontal lobe lesions. J Int Neuropsychol Soc 7:586–596.

Battig WF, Montague WE (1969) Category norms for verbal items in 56 categories: a replication and extension of the Connecticut category norms. J Exp Psychol Monogr 80:1–46.

Becker S, Lim J (2003) A computational model of prefrontal control in free recall: strategic memory use in the California verbal learning task. J Cogn Neurosci 15:821–832.

Blumenfeld RS, Ranganath C (2006) Dorsolateral prefrontal cortex promotes long-term memory formation through its role in working memory organization. J Neurosci 26:916–925.

Blumenfeld RS, Ranganath C (2007) Prefrontal cortex and long-term memory encoding: an integrative review of findings from neuropsychology and neuroimaging. Neuroscientist 13:280–291.

Bousfield WA (1953) The occurrence of clustering in the recall of randomly arranged associates. J Gen Psychol 49:229–240.

Brassen S, Weber-Fahr W, Sommer T, Lehmbeck JT, Braus DF (2006) Hippocampal-prefrontal encoding activation predicts whether words can be successfully recalled or only recognized. Behav Brain Res 171:271–278.

Bunge SA (2004) How we use rules to select actions: a review of evidence from cognitive neuroscience. Cogn Affect Behav Neurosci 4:564–579.

Christoff K, Keramatian K (2007) Abstraction of mental representations: theoretical considerations and neuroscientific evidence. In: Perspectives on rule-guided behavior (Bunge SA, Wallis JD, eds), pp 107–126. New York: Oxford UP.

Cinan S (2003) Executive processing in free recall of categorized lists. Learn Motiv 34:240–261.

Curtis CE, D'Esposito M (2003) Persistent activity in the prefrontal cortex during working memory. Trends Cogn Sci 7:415–423.

Dale AM (1999) Optimal experimental design for event related fMRI. Hum Brain Mapp 8:109–114.

Delis DC, Kramer JH, Kaplan E, Ober BA (1987) The California Verbal Learning Test. New York: Psychological Corporation.

D'Esposito M, Postle BR, Ballard D, Lease J (1999) Maintenance versus manipulation of information held in working memory: an event-related fMRI study. Brain Cogn 41:66–86.

Dobbins IG, Wagner AD (2005) Domain-general and domain-sensitive prefrontal mechanisms for recollecting events and detecting novelty. Cereb Cortex 15:1768–1778.

Eldridge LL, Engel SA, Zeineh MM, Bookheimer SY, Knowlton BJ (2005) A dissociation of encoding and retrieval processes in the human hippocampus. J Neurosci 25:3280–3286.

Gershberg FB, Shimamura AP (1995) Impaired use of organizational strategies in free recall following frontal lobe damage. Neuropsychologia 13:1305–1333.

Gold BT, Balota DA, Jones SJ, Powell DK, Smith CD, Andersen AH (2006) Dissociation of automatic and strategic lexical-semantics: functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. J Neurosci 26:6523–6532.

Gonsalves BD, Kahn I, Curran T, Norman KA, Wagner AD (2005) Memory strength and repetition suppression: multimodel imaging of medial temporal cortical contributions to recognition. Neuron 47:751–761.

Hildebrandt H, Brand A, Sachsenheimer W (1998) Profiles of patients with left prefrontal and left temporal lobe lesions after cerebrovascular infarctions on California Verbal Learning Test-like indices. J Clin Exp Neuropsychol 20:673–683.

Hirshorn EA, Thompson-Schill SL (2006) Role of the left inferior frontal gyrus in covert word retrieval: neural correlates of switching during verbal fluency. Neuropsychologia 44:2547–2557.

Holyoak KJ, Thagard P (1995) Mental leaps: analogy in creative thought. Cambridge, MA: Massachusetts Institute of Technology.

Howard MW, Kahana MJ (2002) A distributed representation of temporal context. J Math Psychol 46:269–299.

Hunt RR, Einstein GO (1981) Relational and item-specific information in memory. J Verbal Learn Verbal Behav 20:497–514.

Janowsky JS, Shimamura AP, Kritchevsky M, Squire LR (1989) Cognitive impairment following frontal lobe damage and its relevance to human amnesia. Behav Neurosci 103:548–560.

Jetter W, Poser U, Freeman RB Jr, Markowitsch HJ (1986) A verbal long term memory deficit in frontal lobe damaged patients. Cortex 22:229–242.

Moscovitch M (1994) Cognitive resources and dual-task interference effects at retrieval in normal people: the role of the frontal lobes in medial temporal cortex. Neuropsychology 8:524–534.

Moscovitch M, Winocur G (2002) The frontal cortex and working with memory. In: The frontal lobes (Stuss DT, Knight RT, eds), pp 188–209. Oxford, UK: Oxford UP.

Murray LJ, Ranganath C (2007) The dorsolateral prefrontal cortex contributions to successful relational memory encoding. J Neurosci 27:5515–5522.

Nessler D, Mecklinger A, Penney TB (2001) Event related brain potentials and illusory memories: differential encoding. Brain Res Cogn Brain Res 10:283–301.

Oztekin I, Long NM, Badre D (2010) Optimizing design efficiency of free recall events for fMRI. J Cogn Neurosci 22:2238–2250.

Petrides M (1994) Frontal lobes and working memory: evidence from investigations of the effects of cortical excisions in nonhuman primates. Amsterdam: Elsevier.

Petrides M (2000) The role of the mid-dorsolateral prefrontal cortex in working memory. Exp Brain Res 133:44–54.

Polyn SM, Natu VS, Cohen JD, Norman KA (2005) Category-specific cortical activity precedes retrieval during memory search. Science 310:1963–1966.

Polyn SM, Norman KA, Kahana MJ (2009) A context maintenance and retrieval model of organizational processes in free recall. Psychol Rev 116:129–156.

Postle BR, Berger JS, Taich AM, D'Esposito M (2000) Activity in the human frontal cortex associated with spatial working memory and saccadic behavior. J Cogn Neurosci 12:2–14.

Raaijmakers JGW, Shiffrin RM (1981) Search of associative memory. Psychol Rev 88:93–134.

Roediger HL, McDermott KB (1995) Creating false memories: remembering words not presented in lists. J Exp Psychol 21:803–814.

Rohrer D, Wixted JT (1994) An analysis of latency and interresponse time in free recall. Mem Cognit 22:511–524.

Savage CR, Deckersbach T, Heckers S, Wagner AD, Schacter DL, Alpert NM, Fischman AJ, Rauch SL (2001) Prefrontal regions supporting spontaneous and directed application of verbal learning strategies. Brain 124:219–231.

Schneider W, Shiffrin RM (1977) Controlled and automatic human information processing. I. Detection, search, and attention. Psychol Rev 84:1–66.

Sederberg PB, Kahana MJ, Howard MW, Donner EJ, Madsen JR (2003) Theta and gamma oscillations during encoding predict subsequent recall. J Neurosci 23:10809–10814.

Sederberg PB, Schulze-Bonhage A, Madsen JR, Bromfield EB, Litt B, Brandt A, Kahana MJ (2007) Gamma oscillations distinguish true from false memories. Psychol Sci 18:927–932.

Shrager Y, Kirwan CB, Squire LR (2008) Activity in both hippocampus and perirhinal cortex predicts the memory strength of subsequently remembered information. Neuron 59:547–553.

Staresina BP, Davachi L (2006) Differential encoding mechanisms for subsequent associative recognition and free recall. J Neurosci 26:9162–9172.

Stark CE, Squire LR (2001) When zero is not zero: the problem of ambiguous baseline conditions in fMRI. Proc Natl Acad Sci U S A 98:12760–12766.

Strange BA, Otten LJ, Josephs O, Rugg MD, Dolan RJ (2002) Dissociable human perirhinal, hippocampal, and parahippocampal roles during verbal encoding. J Neurosci 22:523–528.

Stricker JL, Brown GG, Wixted J, Baldo JV, Delis DC (2002) New semantic and serial clustering indices for the California Verbal Learning Test, Ed 2: Background, rationale, and formulae. J Int Neuropsychol Soc 8:425–435.

Stuss DT, Alexander MP, Palumbo CL, Buckle L, Sayer L, Pogue J (1994) Organizational strategies of patients with unilateral or bilateral frontal lobe injury in word list learning tasks. Neuropsychology 8:355–373.

Troyer AK, Moscovitch M, Winocur G, Alexander MP, Stuss D (1998) Clustering and switching on verbal fluency: the effects of focal frontal- and temporal-lobe lesions. Neuropsychologia 36:499–504.

Tulving E (1962) Subjective organization in free recall of "unrelated" words. Psychol Rev 69:344–354.

Wager TD, Smith EE (2003) Neuroimaging studies of working memory: a meta-analysis. Cogn Affect Behav Neurosci 3:255–274.

Wagner AD, Schacter DL, Rotte M, Koutstaal W, Maril A, Dale AM, Rosen BR, Buckner RL (1998) Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. Science 281:1188–1191.

Wagner AD, Maril A, Bjork RA, Schacter DL (2001) Prefrontal contributions to executive control: fMRI evidence for functional distinctions within lateral prefrontal cortex. Neuroimage 14:1337–1347.