Differential Effects of Semantic Processing on Memory Encoding

Klaus Fliessbach*, Corinna Buerger, Peter Trautner,
Christian E. Elger, and Bernd Weber

*Department of Epileptology, University of Bonn Medical Center and Department of NeuroCognition,
Life & Brain Center, Bonn, Germany*

Abstract: Deeper semantic processing of words leads to enhanced memory encoding (depth of processing effect). The left inferior prefrontal cortex (LIPC) and the left hippocampus are known to be involved in this effect. We tested the hypothesis that different semantic encoding processes contribute qualitatively differently to memory encoding. In a memory experiment using functional magnetic resonance imaging, we compared three different encoding tasks: a nonsemantic alphabetical, an animacy decision, and a size comparison tasks. Recognition memory was tested subsequently. We hypothesized that the size comparison task would activate brain areas involved in the processing of object features and that this would be associated with successful memory encoding. Results showed that the size comparison task led to significantly better memory encoding than the two other tasks. As with the animacy decision task, it led to stronger activation of the LIPC and left hippocampus than the nonsemantic task. Both regions also had stronger activations for later remembered than for nonremembered words. The size comparison task additionally led to stronger activation in the left anterior fusiform gyrus, which was also associated with successful memory encoding. We conclude that different types of semantic processing affect memory encoding based on distinguishable brain processes. *Hum Brain Mapp* 31:1653–1664, 2010. © 2010 Wiley-Liss, Inc.

Key words: depth of processing effect; subsequent memory effect; dual coding theory; functional magnetic resonance imaging (fMRI)

INTRODUCTION

Words that undergo more intense semantic processing are better stored in episodic long-term memory than less intensely processed words. This observation prompted Craik and Lockhart to formulate their “levels of process-

ing” framework for memory research [Craik, 2002; Craik and Lockhart, 1972]. This led to a new direction in memory research: whereas before, the focus lay mostly on the characteristics of the material, now the way we process content has become a major subject of memory research. In its initial form, the levels of processing theory stated that the amount of semantic elaboration increases the duration of memory traces. Yet, exactly how semantic elaboration is defined and quantified remained unclear. Paivio [1975] suggested that different types of elaboration must be distinguished qualitatively. This assumption is based on dual-coding theory, which claims that information in the human brain is represented in a verbal and nonverbal (e.g. visual) code. According to Paivio’s hypothesis, processes that involve both the verbal and the visual channel

*Correspondence to: Klaus Fliessbach, Department of Epileptology, University of Bonn Medical Center, Sigmund Freud-Str. 25, D-53105 Bonn, Germany. E-mail: klaus.fliessbach@ukb.uni-bonn.de

Received for publication 18 August 2009; Revised 9 October 2009; Accepted 11 November 2009

DOI: 10.1002/hbm.20969

Published online 16 February 2010 in Wiley Online Library (wileyonlinelibrary.com).

are more elaborated than those involving only a single channel and should thus lead to better memory formation.

Functional imaging has made it possible to identify neurophysiological correlates of the depth of processing effect and to test hypotheses derived from dual-coding theory. A standard experimental procedure is to confront subjects with words and ask them to perform tasks that require different levels of semantic processing. Comparing responses for items that are remembered in a subsequent memory task with items that were forgotten then makes it possible to identify brain regions that putatively contribute to episodic memory encoding [subsequent memory effect (SME)]. By this method, the left inferior prefrontal cortex (LIPC) and the left medial temporal lobe have been most consistently linked to successful encoding of verbal material in general [Kirchhoff et al., 2000; Wagner et al., 1998]. Given the above-mentioned evidence for the influence of mental processes elicited by a study task on the effectiveness of memory encoding, several studies have addressed this issue by varying the type of task during the learning phase. Some of these studies have found a partial overlap between the regions involved in encoding in both semantic and nonsemantic study tasks within the LIPC and left medial temporal lobe [Baker et al., 2001; Fletcher et al., 2003; Otten et al., 2001]. However, there have also been demonstrations of encoding specificity, i.e. that specific brain regions show SMEs for one task but not another. This has been demonstrated for the task contrasts animacy judgments versus syllable counting [Otten et al., 2002; Otten and Rugg, 2001; Park et al., 2008], processing of meaning versus shape of words [Fletcher et al., 2002], and social versus nonsocial judgments [Mitchell et al., 2004]. These findings suggest that different study tasks may contribute to memory encoding by engaging different neural systems.

The present study was conducted to extend and sharpen these results by not only comparing deep semantic encoding tasks with nonsemantic tasks but by also directly comparing two different semantic tasks with each other. Most studies so far have induced semantic processing by having subjects complete categorization tasks such as living versus nonliving [Otten et al., 2001; Otten and Rugg, 2001; Park et al., 2008] or concrete versus abstract [Baker et al., 2001; Wagner et al., 1998] decisions. To solve tasks like these, the subject has to retrieve rather abstract information from semantic knowledge. As in these previous studies, we utilized an animacy decision task and a nonsemantic task. In addition, we wanted to use a task requiring the analysis of an object's visual features and thus selected a size comparison task. These two different semantic tasks¹ enable us to investigate brain processes involved in memory encoding with possible differences depending on the special nature of the semantic processing induced by the task. Our hypothesis was that the size

comparison task would induce activity in brain structures which are associated with the processing of visual object information. We assumed that this activation contributes to memory encoding and is lacking or less pronounced in the animacy task. The strong reference to visual object features was selected also to test an assumption derived from dual-coding theory: If a task elicits stronger visual analysis processes, this leads to better memory encoding.

Our subjects elaborated on words under three different encoding tasks: first, a nonsemantic alphabetical task, where subjects had to decide whether the first and last letters of a word were in alphabetical order or not, a typical "shallow" processing task; second, a semantic categorization task, where subjects had to decide whether a word refers to a living or a nonliving thing, i.e. an animacy decision; third, a size comparison task, in which subjects had to decide whether an object was larger or smaller than a shoebox. Subjects were then tested to see whether they remembered the words in a subsequent memory task. The hemodynamic response was measured using fMRI during the encoding phase (in all subjects) and during retrieval (in a subgroup of subjects). Scanning during retrieval was performed to test for any task-specific activation during successful recognition, given that previous results suggest that retrieval of memories is associated with a reactivation of brain structures that were active during encoding [Rugg et al., 2008]. However, as results did not provide any evidence for such reactivation, we will not report results for the retrieval data in detail.

Specifically, our hypotheses are the following:

1. We predict that words from the animacy decision task and the size comparison task will have a greater rate of recognition than those from the alphabetical task. We predict that words from the size comparison task will have an equal or higher recognition rate than those from the animacy decision task.
2. Secondly, we predict that the animacy decision task will elicit stronger activity in the LIPC and left hippocampus than the alphabetical task and that this activity will be associated with successful memory encoding.
3. We also predict that this will also be the case for the size comparison task. In addition, we expect to find greater activation in brain regions involved in processing visual object features, especially the fusiform gyrus, when contrasting the size comparison task with the two other tasks. This hypothesis is based on previous studies which have shown that during processing of words, reference to visual object features such as color or form leads to enhanced activation in the fusiform cortex [Pulvermüller and Hauk, 2006]. In the present context, we expect this activation specifically to be associated with successful memory encoding. As a side hypothesis, we also predict that the right hippocampus will contribute more to memory encoding for this task. This hypothesis is derived from the known material specificity of the human

¹The term "semantic" refers to the fact that in contrast to the nonsemantic task both tasks require the processing of word meaning.

hippocampus. The left hippocampus is more strongly involved in the encoding of verbal material and the right hippocampus in the encoding of nonverbal, figural material [Golby et al., 2001]. Therefore, a task that induces the processing of visual information might involve the right hippocampus, although the learned material consists of words.

MATERIALS AND METHODS

Subjects

The study included 35 native German speakers without any history of neurological or psychiatric disease [18 females, mean age 26.2 years (± 4.4), range 22–40 years]. Nineteen subjects were enrolled in Experiment 1 (scanning was performed during encoding only) and 16 in Experiment 2 (scanning during encoding and retrieval). All subjects were right-handed according to the Edinburgh Handedness Inventory. All subjects gave written informed consent, and the study was approved by the Ethics Committee of the University of Bonn. Data from Experiment 1 has already been reported in a publication addressing a different question than the present study [Fliessbach et al., 2007].

Task

The experiment comprised an incidental encoding phase with three different encoding tasks (alphabetical task, animacy decision, and size comparison). Identical or highly similar tasks have been utilized in a number of fMRI studies assessing the depth of processing effect [Fletcher et al., 2003; Otten et al., 2001; Ranganath et al., 2004]. During the encoding phase, subjects saw a total of 252 words that were drawn randomly from a list of 378 words in a single session lasting ~ 39 min. The words were concrete German nouns with a word length between 3 and 10 characters and a word frequency between 1 per million and 150 per million according to the CELEX German word database. Two seconds before a word was displayed, it was indicated on the screen which one of three different tasks was to be performed for the forthcoming word. The three tasks were to decide (i) whether the first and the last letter of the word were in alphabetical order (alphabetical task); (ii) whether the word referred to a living or a nonliving thing (animacy decision); or (iii) whether the word referred to something larger or smaller than a shoebox (size comparison). For the size comparison task, subjects were instructed that in case of doubt, they should decide whether the respective object would fit into a shoebox. The screen indicating the task consisted of the words “alphabetical–nonalphabetical,” “living–nonliving,” or “smaller–larger” and was shown for 500 ms followed by a fixation cross for 1,500 ms. The cues for the different tasks

were equiprobable and randomly intermixed. This means that the assignment of a word to a task was random and varied across subjects. The word was then shown for 1 s and the subjects indicated their answers via response hand grips with their left (alphabetical, living, smaller) or right thumb (nonalphabetical, nonliving, larger). The words were divided evenly among the two possible answers in each task (i.e. 50% of the words had the first and last letter in alphabetical order, were living, or referred to an object smaller than a shoebox, respectively). For the latter aspect, only words that were judged concordantly by 9 out of 10 independent subjects in a pretest were used. The interstimulus interval between the word presentation and the screen indicating the task for the next trial was jittered between 4.5 and 7.5 s with a fixation cross displayed. After a short delay of ~ 10 min, an unexpected subsequent memory test was administered either outside the scanner (Experiment 1) or while subjects were being scanned in a second session lasting ~ 44 min (Experiment 2). The task was presented via video goggles (Nordic NeuroLab, Bergen, Norway) using E-prime presentation software (Psychology Software Tools; www.pstnet.com).

During the memory test, subjects saw all of the 252 words from the encoding period randomly intermixed with the remaining 126 words from the full list and had to decide whether they (i) definitely remembered the word; (ii) thought they recognized it but were not confident; or (iii) thought the word was new. This instruction was intended to differentiate highly confident hits from non-confident responses in which guessing could have occurred. While in Experiment 1 the recognition period was self-paced, in Experiment 2, items appeared every 4.5–7.5 s for 1,000 ms and the decision had to be made before the next word appeared.

Scanning procedure

Scanning was performed on a 1.5 T Avanto Scanner (Siemens, Erlangen, Germany) using a TIM 8-channel head coil. During encoding, 772 echo planar imaging (EPI) scans were acquired including six dummy scans. During retrieval (Experiment 2), 888 EPI scans were acquired. Forty-eight slices with an oblique orientation ($C > T$ 30°) were used. In both experiments, slices covered the whole brain including the cerebellum. Scan parameters in both experiments were as follows: slice thickness = 3 mm; interslice gap = 0.66 mm; matrix size = 64×64 ; field of view = 195×195 mm²; echo time (TE) = 50 ms; repetition time (TR) = 3 s.

fMRI data analysis

The data was preprocessed using FSL software Version 4.1.2 (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). The prestatistics processing included slice-timing correction using Fourier-space time-series phase-shifting; motion

Figure 1.

Memory performance. Percentage of confidently recognized words in each task minus percentage of false alarms.

correction using MCFLIRT; grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor; registration to standard space EPI-template (resampled voxel size after registration $3 \times 3 \times 3 \text{ mm}^3$); smoothing with an 8-mm Gaussian kernel. The fMRI statistical analysis was performed using Statistical Parametric Mapping 5 (SPM5, www.fil.ion.ucl.ac.uk/spm/). Six event types were defined by task condition (alphabetical task, animacy decision, size comparison) and subsequent memory (hits and nonhits). Unconfident “old”-responses and “new”-responses were summarized as nonhits (see below). To model the BOLD time course in each voxel, onset vectors for these six event types were convolved with the SPM5 canonical hemodynamic response function (HRF) and its temporal and dispersion derivatives to account for latency shifts and shape differences of the BOLD response across brain regions. In the main analysis, all items that were shown during the encoding phase were included, to obtain as many observations as possible. In an additional analysis, items that were not correctly answered in the study task were excluded and modeled as an additional regressor of no interest. Because response data for the study phase was missing for three subjects due to technical reasons (all from Experiment 2), this analysis only included 32 subjects. For each subject, parameter images for the contrasts of each condition were generated and were subjected to a second-level random effects analysis using a 3×2 within-subjects ANOVA with the factors task (alphabetical task, animacy decision, and size comparison) and subsequent memory (hits and nonhits).

According to our a priori hypotheses, all analyses were performed for the following regions of interest: LIPC (1,329 voxels), left (318 voxels) and right hippocampus (314 voxels), and bilateral fusiform gyrus (1,115 voxels). The masks for these were generated by WFU-pick atlas Version 2.4 [Maldjian et al., 2003]. Within the regions of in-

terest, an inclusion threshold of $P < 0.005$, uncorrected, was set. To correct for multiple comparisons, only clusters with activations yielding a $P < 0.05$, family-wise error (FWE)-corrected for the search volume, are reported. Additionally, for exploratory analysis of activations outside these regions, a whole-brain analysis with a threshold of $P < 0.001$ and an extent threshold of six adjacent voxels were conducted.

The analysis for the encoding phase included the following steps:

1. Task effects: Each task was contrasted with the two other tasks separately, irrespective of the memory outcome, which yielded six directional comparisons.
2. SMEs: Separately for each task, hits were contrasted with nonhits.
3. Overlap of task effects and SME: For each area within the ROI that showed significant differences between tasks, we tested for SME by averaging across the whole area and testing for differences in activation between hits and nonhits in each task using a repeated measures *t*-test at a threshold of $P < 0.05$. On the whole-brain level, we identified regions with an overlap between task effects and SME by inclusively masking task effect maps ($P < 0.001$, uncorrected, extent threshold 6 voxels) with SME maps at a threshold of $P < 0.05$. This yields regions showing task-related effect and SME in the respective task.
4. Interaction between task effects and SME: We also tested whether in the subregions displaying significant task effects there were significant differences in the SME between the tasks. Here, we used one-tailed tests, because of our a priori hypothesis concerning the strength of SME for the different tasks in different regions.

RESULTS

Performance

Mean accuracy for the encoding tasks was 0.86 ± 0.03 (alphabetical task), 0.92 ± 0.01 (animacy decision), and 0.96 ± 0.01 (size comparison). Reaction times were significantly longer for the alphabetical task ($1,742 \pm 568 \text{ ms}$) than for the two other tasks ($P < 0.001$), which did not differ significantly in reaction times ($1,245 \pm 446 \text{ ms}$ for animacy decision, $1,303 \pm 476 \text{ ms}$ for size comparison, $P > 0.05$).

Memory performance (hits minus false alarms) is depicted in Figure 1. Memory performance was significantly greater than chance (hits minus false alarms > 0) for all three tasks. Subsequent memory performance showed a strong dependence on the encoding task (one-way ANOVA, main effect of task, $F_{33,2} = 184.4$, $P < 10^{-5}$) with significant improvement from the alphabetical task (0.18 ± 0.10) to the animacy decision task (0.47 ± 0.14)

TABLE I. Reaction times in ms (RT) and accuracy for the three tasks, divided by misses and hits

		Misses		Hits		T	P
		Mean	SD	Mean	SD		
RT							
Alphabetical task		1,707	587	1,750	581	−1.1	0.280
Animacy decision task		1,203	436	1,281	465	−3.6	0.001
Size comparison task		1,271	456	1,321	488	−1.9	0.069
Accuracy							
Alphabetical task		0.87	0.15	0.85	0.15	1.2	0.243
Animacy decision task		0.95	0.05	0.90	0.05	5.6	<0.001
Size comparison task		0.96	0.05	0.96	0.04	0.5	0.589

and the size comparison task (0.52 ± 0.14) (pairwise dependent measures *t*-tests, all $P < 0.01$). This means that among the two semantic encoding tasks the size comparison task led to significantly better subsequent memory performance than the animacy decision task. “Unconfident” responses were given to learned items significantly less often than to new words (0.21 ± 0.12) for the animacy decision task (0.17 ± 0.11) and the size comparison task (0.16 ± 0.10) and significantly more often for the alphabetical task (0.25 ± 0.13). This means that “unconfident” responses did not indicate positive recognition for the two semantic tasks and did not contribute to the observed levels of processing effect. On the basis of these findings and in accordance with previous studies [Duverne et al., 2009; Morcom et al., 2003; Otten et al., 2001] we categorized words as remembered (hits) only when they attracted *confident* old-responses and otherwise regarded them as non-hits. Pooling of “unconfident” and “new” responses was also done to increase the robustness of parameter estimation for the subsequent memory contrasts for the two semantic tasks, where otherwise in several subjects the number of events in the “new”-answers to old items category would have been critically low. Because for the alphabetical task “unconfident” answers discriminated between old and new items above chance level, we additionally calculated parameter estimates based on the three separated answer categories for this task, to test whether this would affect our results.

Accuracy for the study task did not differ between remembered and forgotten words for the alphabetical and the size comparison task, but it did for the animacy decision task. Here, accuracy for words remembered later was lower (0.90 ± 0.05) than for words that were forgotten (0.95 ± 0.06). There was also a significant difference in reaction times for the animacy decision task between subsequently remembered and forgotten words, but not for the two other tasks (Table I). These findings indicate that task difficulty influenced subsequent memory performance at least for the animacy decision task. To ensure that results were not influenced by these differences, we ran an analysis excluding incorrectly answered words during encoding in addition to our main fMRI analysis, which included all

items. In a further analysis, we included reaction times as a possible nuisance factor in the second-level analysis.

The behavioral results are in accordance with Hypothesis (i): Both types of deep semantic processing led to markedly better memory performance compared with the alphabetical task. From the two deep semantic processing tasks, size comparison led to significantly better memory encoding than animacy decision. It is of note that these differences are not explained by reaction time differences between the tasks.

ROI Analyses

Left inferior prefrontal cortex

There was a stronger activation of the LIPC in both the animacy decision and the size comparison task than in the alphabetical task in a region corresponding to Broca’s area. Activation here was associated with successful memory encoding of words processed with the animacy decision ($T_{34} = 3.0$, $P = 0.005$) and the size comparison task ($T_{34} = 4.8$, $P < 0.001$), but not with the alphabetical task ($T_{34} = 0.3$, $P > 0.9$), yielding a significant task $\times$ memory interaction ($F_{2,68} = 3.5$, $P = 0.037$) (see Fig. 2) with both the animacy decision task ($T_{34} = 1.8$, $P = 0.042$, one-sided) and the size comparison task ($T_{34} = 2.4$, $P = 0.012$, one-sided) having a significantly larger SME than the alphabetical task.

Hippocampus

There was stronger left hippocampal activation for the animacy decision and the size comparison task than for the alphabetical task. In this area, the only significant SME was present for the size comparison task ($T_{34} = 3.2$, $P = 0.003$), whereas the animacy decision task showed a trend for higher activation for later remembered versus forgotten words ($T_{34} = 1.8$, $P = 0.079$), and the alphabetical task showed no positive SME ($T_{34} = -0.3$) (Fig. 2b). The SME for the size comparison task (but not for the animacy decision task) was significantly larger than that for the alphabetical decision task ($T_{34} = 2.5$, $P = 0.010$, one-sided).

Figure 2.

(a) Results for the LIPC: In the LIPC (peak voxel coordinates $X = -45$, $Y = 27$, $Z = -15$), both semantic tasks led to stronger activation than did the alphabetical task, and for these tasks higher activations were associated with successful memory encoding. The SPM is thresholded at $P < 0.005$, uncorrected, for both contrasts AD>AT and SC>AT (i.e. inclusive masking). The activation is significant at $P < 0.05$, FWE-corrected for search volume LIPC. AT = alphabetical task, AD = animacy decision, SC = size comparison. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. (b) Results for the left hippocampus: In the left hippocampus (peak voxel coordinates $X = -21$, $Y = -15$, $Z = -21$), both semantic tasks led to stronger activation than did the alphabetical task. There was a significant SME for the size comparison task and a tendency for higher activation for remembered versus forgotten words for the animacy decision. The SPM is thresholded at $P < 0.005$, uncorrected,

for both contrasts AD>AT and SC>AT (i.e. inclusive masking). The activation is significant at $P < 0.05$, FWE-corrected for the search volume comprising the left hippocampus. AT = alphabetical task, AD = animacy decision, SC = size comparison. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. (c) In the left anterior fusiform gyrus (peak voxel coordinates $X = -33$, $Y = -27$, $Z = -21$), both semantic tasks led to stronger activation than did the alphabetical task. Activation for the size comparison task was significantly higher than that for the animacy decision task. There was a significant SME for the size comparison task and not for the other tasks. The SPM is thresholded at $P < 0.005$, uncorrected, contrast SC>AT. The activation is significant at $P < 0.05$, FWE-corrected for the search volume comprising the bilateral fusiform gyrus. AT = alphabetical task, AD = animacy decision, SC = size comparison. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

Figure 3.

In the right anterior fusiform gyrus (peak voxel coordinates $X = 36$, $Y = -42$, $Z = -24$), there was a task-specific SME for the size comparison task. Both the alphabetical and the size comparison tasks activated this region stronger than the animacy decision task. The SPM is thresholded at $P < 0.001$, uncorrected, contrast hits > misses for the size comparison task.

The results for the LIPC and the left hippocampus are in accordance with Hypothesis (ii), except that the SME for the animacy decision task did not reach statistical significance.

Stronger activation in the right hippocampus for the two semantic tasks were observed in a region homologous to the left-sided activation spots, but no significant SME was observed for any task in this area (all $P < 0.4$) (not shown).

Fusiform gyrus

There was a stronger activation of the left anterior fusiform gyrus in the size comparison task than in the alphabetical task. In this region, there was also (i) stronger activation for the size comparison task than for the animacy decision task ($T_{34} = 2.3$, $P = 0.032$) and (ii) a significant SME in the size comparison task ($T_{34} = 3.0$, $P = 0.005$), but not for the two other tasks. The SME for the size comparison task was significantly larger than that for the alphabetical task ($T_{34} = 2.7$, $P = 0.006$, one-sided). The comparison between the SME for size comparison and animacy decision revealed a trend toward a stronger effect in the size comparison task ($T_{34} = 1.5$, $P = 0.074$, one-sided).

There was also an area within the left anterior fusiform gyrus where the animacy decision task elicited higher activations than the alphabetical task. Activation here did not differ from that in the size comparison task and no SME was observed in any task (all $P > 0.2$). The alphabetical task was associated with stronger activations in bilateral posterior fusiform regions than the two other tasks, adjacent to the extended occipital activations present during processing of this task. These activations were also not predictive of subsequent memory for any task (not shown).

These results demonstrate additional memory-relevant activation in the left anterior fusiform gyrus by the size comparison task, which is in accordance with Hypothesis (iii).

Neither excluding items that yielded inaccurate answers in the encoding tasks nor including reaction times as a covariate did significantly alter these results. The separate analysis of “unconfident” and “new”-answers for the alphabetical task did not reveal any differences in the BOLD responses to these conditions in the ROI (repeated measures t -tests, all $P > 0.2$).

Whole-Brain Analysis

Task effects

First, we analyzed which brain regions were specifically involved during the performance of a task, irrespective of the subsequent memory outcome. Contrasting fMRI signals associated with the alphabetical task with both other tasks separately yielded activation of a widespread network of bilateral parietooccipital cortex and medial and dorsolateral prefrontal cortex, much as described previously [Otten et al., 2001]. The animacy decision task conversely led to stronger activations than the alphabetical tasks predominantly in the left hemisphere, with major clusters in the middle and inferior temporal gyrus, inferior frontal gyrus, and medial prefrontal areas. Compared to the size comparison task, stronger activations for the animacy decision task were present in the right posterior fusiform gyrus and left caudate head. Compared with the alphabetical task, the size comparison task led to stronger activation of left frontal inferior and middle temporal gyrus, medial prefrontal areas, and left and right medial

TABLE II. Results of the whole brain analysis

Contrast	Location			Zmax	N	Region
	x	y	z			
Subsequent memory effects (hits > misses)						
Alphabetic decision task	12	-63	-30	4.48	38	R. cerebellum (culmen, declive, pyramis, uvula)
	27	-72	-39	3.53	7	R. pyramis
Animacy decision task	-51	30	0	4.59	125	L. inferior frontal gyrus, middle frontal gyrus
	-12	36	45	3.89	88	L. superior frontal gyrus, medial frontal gyrus, cingulate gyrus
	18	-45	3	3.83	12	R. parahippocampal gyrus
	-9	-51	-39	4.06	8	L. anterior cerebellum (nodule)
Size comparison task	-57	-45	-18	3.58	8	Inferior temporal gyrus, l. middle temporal gyrus
	-51	12	-15	4.11	87	L. inferior frontal gyrus
	-9	24	45	4.11	47	L. supplementary motor area
	24	-72	-48	4.33	17	R. inferior semilunar lobule
	-51	12	-15	4.20	11	L. superior temporal gyrus
	-9	18	0	3.37	11	L. superior temporal gyrus
	36	-42	-24	3.63	10	R. fusiform gyrus
	-39	-24	24	3.88	10	L. insula, inferior parietal lobule
	-54	-54	-18	3.52	9	Inferior temporal gyrus, l. middle temporal gyrus, l. fusiform gyrus
	-45	-48	-21	3.29	7	L. fusiform gyrus, culmen
Task effects inclusively masked with subsequent memory effects						
Alphabetical task > animacy decision	-27	-75	33	Inf	167	L. middle occipital gyrus, precuneus
	6	-75	21	Inf	59	R. precuneus
	9	-75	-24	4.88	34	R. cerebellum
	-45	-66	-12	7.38	20	L. inferior occipital gyrus
	-42	9	27	5.11	19	L. inferior frontal gyrus (pars opercularis)
	6	-60	6	4.63	19	R. lingual gyrus, posterior cingulate
	63	-48	-12	5.50	16	R. inferior temporal gyrus, middle temporal gyrus
	36	24	21	4.48	15	R. middle frontal gyrus
	42	-63	39	5.37	14	R. angular gyrus, inferior parietal lobule
	-6	6	57	5.18	13	L. supplementary motor area, superior frontal gyrus
	18	-57	48	4.78	12	Precuneus
	9	-90	0	4.63	11	R. calcarine sulcus, lingual gyrus
	-6	24	30	4.12	9	L. anterior cingulum, cingulate gyrus
	-33	-72	-33	4.59	7	L. cerebellar crus, uvula
	-12	-15	9	4.34	7	L. thalamus
	24	-66	-51	3.81	7	R. cerebellum, inferior semilunar lobule
	54	-57	-21	5.36	6	R. inferior temporal gyrus
	-9	-78	-33	4.05	6	R. cerebellar crus, pyramis
	15	9	30	3.40	6	Cingulate gyrus
Alphabetical task > size comparison	-27	-63	42	Inf	146	L. superior parietal lobe, precuneus
	6	-75	51	Inf	51	R. precuneus
	9	-75	-24	4.57	42	R. cerebellum, declive
	-12	-12	12	4.81	25	L. thalamus
	3	-69	-42	3.63	23	Vermis, uvula
	15	-24	-12	4.28	17	R. parahippocampal gyrus
	-45	-69	-12	6.30	16	L. inferior occipital gyrus
	-6	24	30	4.49	16	L. anterior cingulum, cingulate gyrus
	27	-63	-51	4.37	16	R. cerebellum, inferior semilunar lobule
	-3	6	57	6.24	15	L. supplementary motor area, superior frontal gyrus
	18	-57	48	4.10	13	R. superior parietal lobe, precuneus
	12	-60	-30	3.44	12	Culmen
	9	-90	0	5.60	11	R. calcarine sulcus, lingual gyrus
	3	-63	6	4.75	10	R. lingual gyrus, posterior cingulate gyrus
	-6	-72	-18	3.99	9	L. cerebellum, declive
	-33	-72	-33	4.23	8	L. cerebellum, uvula
	42	21	27	3.88	7	R. inferior frontal gyrus (pars triangularis), middle frontal gyrus
	0	-84	27	3.78	7	L. cuneus
	6	33	21	3.77	6	R. anterior cingulum
	-39	6	30	5.06	6	L. precentral gyrus, inferior frontal gyrus

TABLE II. (Continued)

Contrast	Location			Zmax	N	Region
	x	y	z			
Animacy decision > alphabetical task	-51	6	-24	4.22	47	L. middle temporal gyrus
	-48	33	-6	4.24	45	L. orbitofrontal cortex , inferior frontal gyrus
	-3	45	-21	5.17	40	L. rectal gyrus , orbital gyrus
	-9	48	42	4.77	37	L. superior medial frontal gyrus , medial frontal gyrus
	-45	27	-15	4.86	23	L. orbitofrontal cortex , inferior frontal gyrus
	-6	60	24	4.38	16	L. superior medial frontal gyrus , l. medial frontal gyrus
Size comparison > alphabetical task	-54	27	12	3.78	10	L. inferior frontal gyrus (pars triangularis) , inferior frontal gyrus
	-54	27	-3	4.63	93	L. inferior frontal gyrus (pars triangularis)
	-3	42	-21	4.64	24	L. rectal gyrus , orbital gyrus
	-9	39	45	4.19	23	L. superior medial frontal gyrus , medial frontal gyrus
	-33	-30	-21	4.31	11	L. fusiform gyrus
	-21	-12	-21	4.27	15	L. hippocampus , parahippocampal gyrus
Size comparison > animacy decision	-51	-60	-12	3.65	13	L. inferior temporal gyrus

Region of the peak activation is written in bold letters.

Zmax: Z-value of activation peak; N: cluster size.

temporal lobe (hippocampus/parahippocampal gyrus and left fusiform gyrus). Compared with the animacy decision task, we observed stronger activation during size comparison in bilateral temporoparietal regions including inferior and lateral parietal lobe, the precuneus and temporal cortices as well as bilateral middle frontal gyri. These results are not reported in detail.

Subsequent memory effects

We proceeded to calculate SMEs for each task separately by contrasting subsequent hits versus nonhits. Using a threshold of $P < 0.001$, the only region displaying a SME for the alphabetical task was the right cerebellum. This was also observed when confident hits were contrasted against “new” answers alone (“unconfident” responses excluded). SMEs for the animacy decision task were confined to the left frontal and temporal lobe and the right posterior parahippocampal gyrus. For the size comparison task SMEs were also observed in the LIPC and additionally there were SMEs in the left and right fusiform gyri as well as an area in the left supplementary motor area. Thus, in addition to the ROI analysis, this contrast identified a SME for the size comparison task not only in the left but also in the right fusiform gyrus (see Fig. 3). This area was significantly activated by the alphabetical and the size comparison task but not by the animacy decision task.

Overlap between task and subsequent memory effects

The main purpose of this study was to identify brain regions which are specifically activated by the different tasks and which additionally subserve memory encoding. We identified those regions by inclusively masking the task contrasts ($P < 0.001$, uncorrected, extent threshold 6 voxels) with the

SME maps with a mask threshold of $P < 0.05$. Within the regions showing stronger activations for the alphabetical task than for the two other tasks, there were several, mostly posterior, regions which also predicted successful memory encoding. Among the regions showing stronger activations by the animacy decision task than by the alphabetical task, left lateral temporal and left inferior frontal activations, as well as medial prefrontal activations predicted memory encoding. Among the regions showing stronger activations by the size comparison than by the alphabetical task, activations in the LIPC, medial prefrontal cortex, left hippocampus, and left fusiform gyrus predicted memory encoding. Finally, stronger activation in the left lateral inferior temporal gyrus by the size comparison than by the animacy decision task was associated with successful encoding. The opposite contrast did not show significant results. Thus, the results of the whole-brain analysis are essentially consistent with the ROI analyses: There was concordant memory-relevant activation of the LIPC (and the medial prefrontal cortex, which was not subject of the ROI analysis) by the two semantic tasks as compared to the nonsemantic task. The size comparison task led to additional memory-relevant activation of the left hippocampus (in which activation did not survive the extent threshold for the animacy task) and the left anterior fusiform gyrus. Additionally, in direct comparison of the two semantic tasks, left (lateral) inferior temporal lobe activation was stronger in the size comparison task. Results of the whole-brain analysis are summed up in Table II.

DISCUSSION

The present study demonstrates improved recognition memory for words that were processed with a size

comparison task compared to a nonsemantic task and an animacy decision task. This behavioral effect was paralleled by higher activation in the left anterior fusiform gyrus during the processing of the size comparison task compared with the two other tasks and a specific SME for this task in the same area. There was also a SME specifically for this task in the right fusiform gyrus. These findings suggest that additional recruitment of object-processing brain areas in a task requiring the analysis of visual object features contributes to enhanced memory encoding.

The choice of a task explicitly requiring the analysis of a visual feature such as size was motivated by theoretical considerations. Dual-coding theory claims that there are different codes or channels by which items are stored in memory, i.e. a verbal and nonverbal codes [Paivio, 1975]. As a consequence, items that are encoded via both verbal and nonverbal channels should be better kept in memory. This theory delivers a possible explanation for the fact that highly visualizable, concrete words are easier to memorize than abstract words, because there presumably is a stronger nonverbal (in this case visual) representation of them [Fließbach et al., 2006]. In this example, it is a property of the processed items that we assume leads to enhanced encoding by recruitment of different encoding channels. Another hypothesis would be that the type of processing of items during encoding should lead to different encoding success depending on the amount of recruitment of the two assumed routes. Therefore, a task that depends on the analysis of visual object features might lead to better encoding than a task that places lesser demands on processing visual information. We assume that this is the case in the size comparison task used in the present study in contrast to the animacy decision task. This assumption remains somewhat speculative because the degree to which the subjects did or did not visualize the objects or engage visual processing circuits in the size comparison task or even the animacy decision task cannot be inferred from our data. One could argue that the animacy decision task, like any kind of semantic processing, automatically induces visual information processing, where concrete words are concerned. This is in line with our finding that the animacy decision task also led to stronger anterior fusiform activation than the nonsemantic task. This activation was, however, not related to memory encoding success. In terms of operationalization, it can be stated that the size comparison task explicitly requires processing of a visual feature, while the animacy decision task does not. Therefore, the assumption that the size comparison task makes greater demands on visual information processing than the animacy decision task seems justified.

Our behavioral results show that recognition memory significantly improves from the nonsemantic task to the animacy decision task and the size comparison task. It is important that this finding is not simply explained by reaction time differences between the tasks: subjects completed both semantic tasks with a *shorter* reaction time (RT) than the nonsemantic task and did not differ signifi-

cantly in their reaction time. Thus, there is no evidence that *longer* processing during the study phase is responsible for the memory differences between the tasks. This finding is in line with classical work conducted under the levels of processing-framework [Craik and Tulving, 1975]. Within the tasks, however, there was a significantly longer RT for the animacy decision task for items later remembered versus items forgotten. In addition, for this task, accuracy differed between later hits and misses, where more items that were processed incorrectly were remembered. Both findings suggest that in this task the level of difficulty (or more generally time of processing during encoding) influenced memory encoding: Words that required longer processing and more often led to mistakes showed a tendency to be remembered better. Similar findings have been reported before [Duverne et al., 2009; Uncapher and Rugg, 2005; Wagner et al., 1998], but there are also several studies that did not produce such an effect [Otten et al., 2001; Otten and Rugg, 2001; Park et al., 2008]. It might mean that for the animacy decision task, part of the SMEs seen in the fMRI analysis is a consequence of more demanding processing as a function of task difficulty. For the two other tasks, there were only insignificant trends in the same direction for reaction times, and no obvious effect of accuracy, so that we assume that task difficulty did not have a major impact on memory encoding in these tasks. To ensure that our fMRI findings were not driven by effects of difficulty, we calculated the results both with and without items with inaccurate responses and observed no qualitative differences in the results. Given that significant RT differences for remembered and forgotten words were confined to the animacy decision task and that fMRI studies have found similar SME without consistently observing this RT difference, we do not expect that our results are affected by this circumstance. This is underpinned by the observation that the inclusion of RT as a covariate did not alter our results.

Our fMRI results extend previous findings on the neural basis of depth of processing effects. Previous studies have shown that deep semantic processing involves the LIPC [Kapur et al., 1994]. In a study comparing the animacy decision with the same nonsemantic task that we used, Otten et al. [2001] found evidence that LIPC and left hippocampal activation were associated with both the depth of processing and memory success. This makes these structures key candidates for the mediation of depth of processing effects on memory encoding. Our study essentially replicates these findings (with the exception that the left hippocampal SME for the animacy decision task did not reach statistical significance) and additionally shows that the same structures contribute to memory encoding during a different semantic task, in this case size comparison. This overlap suggests that the effect of deeper semantic processing is mediated by these two structures, independent of the exact nature of the task. Besides this overlap in the activations by both tasks, we observed that within a region in the left anterior fusiform gyrus, a specific enhancement

of activation occurred during the size comparison task. This region further shows a specific SME for the size comparison task. SMEs in the fusiform gyrus have been inconsistently described previously [Baker et al., 2001; Otten and Rugg, 2001; Wagner et al., 1998; Weis et al., 2004]. The exact nature of these effects remains unclear, given that they occur under different types of semantic [Baker et al., 2001; Wagner et al., 1998] and nonsemantic [Baker et al., 2001; Otten and Rugg, 2001] processing as well as if the stimuli are pictures and not words [Weis et al., 2004]. The present study's results suggest that the left anterior fusiform gyrus in particular contributes to memory encoding when the processing of visual object features is required by the study task. The left anterior fusiform gyrus has been consistently linked to the retrieval of visual information from memory [Wheeler and Buckner, 2003, 2004; Woodruff et al., 2005]. We therefore assume that this activation results from more intense visual information processing under the size comparison task. This region contributes to memory encoding specifically for this task, and one might speculate that this additional activation is a reason for the superiority of memory encoding for that task compared with the two other tasks. The SME for the size comparison task in a correspondent region of the right anterior fusiform gyrus in the whole-brain analysis further supports this interpretation. Whether our results are really a demonstration of "dual-coding" must remain speculative, given the ambiguous character of fMRI results (see below). Our behavioral and fMRI results are, however, both consistent with the assumption that specific neural processes underlie memory encoding of words that undergo semantic analysis of visual features. In the same sense we would expect such specific neural processes underlying memory encoding if the orienting task would focus on other sensory features such as acoustic or somatosensory properties.

In addition to our region of interest analysis (based on a priori hypotheses), we conducted a further exploratory whole-brain analysis to be able to detect relevant effects in other brain areas than the a priori chosen ones. This analysis revealed an overlap between task-related activity and a SME for the alphabetical task predominantly in posterior brain regions including the right cerebellum. Similar effects in posterior brain regions have been previously reported for a syllable counting task [Otten and Rugg, 2001], and we have reported and commented on a cerebellar SME [Fliessbach et al., 2007]. Presumably, these activations reflect phonological working memory processes that are necessary to solve such structural tasks, and more intense processing is associated with better long-term memory encoding. In addition to the activations in the LIPC, hippocampus and fusiform gyrus that were subject to the regional analyses, task-related activity that predicted memory encoding for the two semantic tasks was found in medial prefrontal areas for the two semantic tasks. Again, similar results have been reported [Otten and Rugg, 2001], and it has been suggested that these are correlates of preparatory states induced by the cues signaling the task.

This assumption is backed by electrophysiological data showing that such activity appears prior to the onset of the actual stimuli [Otten et al., 2006].

As for fMRI data in general, our results do not allow inferences to be drawn as to the mechanisms by which the mentioned brain areas contribute to memory encoding. They do not even allow the inference that these areas are necessary for memory encoding. More causal inferences will have to be provided by lesion studies or studies using transcranial magnetic stimulation. As for the hippocampus, it appears undisputable that it is causally involved in successful memory encoding, given the clear evidence that hippocampal damage leads to memory encoding deficits [Fletcher et al., 1997] and more direct evidence from electrophysiological studies for its functional role in memory encoding [Fernandez et al., 1999]. The role of the LIPC and other brain areas, such as the fusiform gyrus, remains less clear. It has been suggested that LIPC activity represents working memory engagement during the processing of semantic tasks and that a high working memory engagement during the processing of words is a prerequisite for effective encoding in the left hippocampus [Moscovitch, 1992]. This assumes a serial process, where working memory processes precede memory formation in the hippocampus. Similarly, fusiform activity could reflect visual processing operations in the ventral visual stream during processing of visual object features, and these processes could precede memory encoding in the hippocampus. Again, fMRI data cannot be used to test these assumptions because of its correlative nature and its poor temporal resolution. Task-specific SME as described in our study could, however, create foci for future studies, which could clarify the functional role of the involved structures in memory formation.

As a side hypothesis, we formulated that memory encoding during the size comparison task might also involve the right hippocampus. This hypothesis relied on the known material specificity of the human hippocampus, where the left hippocampus is more strongly involved in the encoding of verbal material (in left language dominant individuals [Weber et al., 2007]) and the right hippocampus in the encoding of nonverbal, figural material [Golby et al., 2001]. Therefore, a task that induces the processing of visual information might also involve the right hippocampus, even when the learned material consists of words. We did not find any evidence for this hypothesis. Instead, our results suggest that additional recruitment during the size comparison task with relevance for memory formation occurs predominantly on the left side. However, the SME for the size comparison task in the right fusiform gyrus indicates an involvement of the right hemisphere in encoding words under such a task. For more causal inferences on the lateralization of different effects of semantic processing on memory encoding, lesion studies might provide helpful data.

In conclusion, our results demonstrate a task-specific depth of processing effect associated with enhanced subsequent memory in the left fusiform gyrus in a size comparison task. Additionally, we observed a SME in the right

fusiform gyrus which was specific for this task. These results demonstrate that although different types of semantic processing share common brain processes, there are also qualitative differences between semantic processes and their influence on episodic memory encoding which are based on distinct brain structures.

ACKNOWLEDGMENTS

This research was supported by the German Federal Ministry of Education and Research (BMBF 01GW0511).

REFERENCES

- Baker JT, Sanders AL, Maccotta L, Buckner RL (2001): Neural correlates of verbal memory encoding during semantic and structural processing tasks. *Neuroreport* 12:1251–1256.
- Craik FIM (2002): Levels of processing: Past, present ... and future? *Memory* 10:305–318.
- Craik FIM, Lockhart RS (1972): Levels of processing—Framework for memory research. *J Verbal Learn Verbal Behav* 11:671–684.
- Craik FIM, Tulving E (1975): Depth of processing and retention of words in episodic memory. *J Exp Psychol Gen* 104:268–294.
- Duvernois S, Motamedinia S, Rugg MD (2009): The relationship between aging, performance, and the neural correlates of successful memory encoding. *Cereb Cortex* 19:733–744.
- Fernandez G, Effern A, Grunwald T, Pezer N, Lehnertz K, Dimpelmann M, Van Roost D, Elger CE (1999): Real-time tracking of memory formation in the human rhinal cortex and hippocampus. *Science* 285:1582–1585.
- Fletcher PC, Frith CD, Rugg MD (1997): The functional neuroanatomy of episodic memory. *Trends Neurosci* 20:213–218.
- Fletcher PC, Palomero-Gallagher N, Zafiris O, Fink GR, Tyler LK, Zilles K (2002): The influence of explicit instructions and stimulus material on lateral frontal responses to an encoding task. *Neuroimage* 17:780–791.
- Fletcher PC, Stephenson CM, Carpenter TA, Donovan T, Bullmore ET (2003): Regional brain activations predicting subsequent memory success: An event-related fMRI study of the influence of encoding tasks. *Cortex* 39:1009–1026.
- Fliessbach K, Trautner P, Quesada CM, Elger CE, Weber B (2007): Cerebellar contributions to episodic memory encoding as revealed by fMRI. *Neuroimage* 35:1330–1337.
- Fliessbach K, Weis S, Klaver P, Elger CE, Weber B (2006): The effect of word concreteness on recognition memory. *Neuroimage* 32:1413–1421.
- Golby AJ, Poldrack RA, Brewer JB, Spencer D, Desmond JE, Aron AP, Gabrieli JD (2001): Material-specific lateralization in the medial temporal lobe and prefrontal cortex during memory encoding. *Brain* 124:1841–1854.
- Kapur S, Craik FI, Tulving E, Wilson AA, Houle S, Brown GM (1994): Neuroanatomical correlates of encoding in episodic memory: Levels of processing effect. *Proc Natl Acad Sci USA* 91:2008–2011.
- Kirchhoff BA, Wagner AD, Maril A, Stern CE (2000): Prefrontal-temporal circuitry for episodic encoding and subsequent memory. *J Neurosci* 20:6173–6180.
- Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH (2003): An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 19:1233–1239.
- Mitchell JP, Macrae CN, Banaji MR (2004): Encoding-specific effects of social cognition on the neural correlates of subsequent memory. *J Neurosci* 24:4912–4917.
- Morcom AM, Good CD, Frackowiak RS, Rugg MD (2003): Age effects on the neural correlates of successful memory encoding. *Brain* 126:213–229.
- Moscovitch M (1992): Memory and working-with-memory—A component process model based on modules and central systems. *J Cogn Neurosci* 4:257–267.
- Otten LJ, Henson RN, Rugg MD (2001): Depth of processing effects on neural correlates of memory encoding: relationship between findings from across- and within-task comparisons. *Brain* 124:399–412.
- Otten LJ, Henson RNA, Rugg MD (2002): State-related and item-related neural correlates of successful memory encoding. *Nat Neurosci* 5:1339–1344.
- Otten LJ, Quayle AH, Akram S, Ditlew TA, Rugg MD (2006): Brain activity before an event predicts later recollection. *Nat Neurosci* 9:489–491.
- Otten LJ, Rugg MD (2001): Task-dependency of the neural correlates of episodic encoding as measured by fMRI. *Cereb Cortex* 11:1150–1160.
- Paivio A (1975): Coding distinctions and repetition effects in memory. In: Bower GH, editor. *The Psychology of Learning and Motivation*, Vol. 9. New York: Academic Press.
- Park H, Uncapher MR, Rugg MD (2008): Effects of study task on the neural correlates of source encoding. *Learn Mem* 15:417–425.
- Pulvermuller F, Hauk O (2006): Category-specific conceptual processing of color and form in left fronto-temporal cortex. *Cereb Cortex* 16:1193–1201.
- Ranganath C, Yonelinas AP, Cohen MX, Dy CJ, Tom SM, D'Esposito M (2004): Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia* 42:2–13.
- Rugg MD, Johnson JD, Park H, Uncapher MR (2008): Encoding-retrieval overlap in human episodic memory: A functional neuroimaging perspective. *Prog Brain Res* 169:339–352.
- Uncapher MR, Rugg MD (2005): Encoding and the durability of episodic memory: A functional magnetic resonance imaging study. *J Neurosci* 25:7260–7267.
- 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.
- Weber B, Fliessbach K, Lange N, Kugler F, Elger CE (2007): Material-specific memory processing is related to language dominance. *Neuroimage* 37:611–617.
- Weis S, Klaver P, Reul J, Elger CE, Fernandez G (2004): Temporal and cerebellar brain regions that support both declarative memory formation and retrieval. *Cereb Cortex* 14:256–267.
- Wheeler ME, Buckner RL (2003): Functional dissociation among components of remembering: control, perceived oldness, and content. *J Neurosci* 23:3869–3880.
- Wheeler ME, Buckner RL (2004): Functional-anatomic correlates of remembering and knowing. *Neuroimage* 21:1337–1349.
- Woodruff CC, Johnson JD, Uncapher MR, Rugg MD (2005): Content-specificity of the neural correlates of recollection. *Neuropsychologia* 43:1022–1032.