

Functional anatomy of a common semantic system for words and pictures

R. Vandenberghe, C. Price, R. Wise, O. Josephs & R. S. J. Frackowiak

Wellcome Department of Cognitive Neurology, Institute of Neurology, 12 Queen Square, London WC1N 3BG, UK

THE relationship between the semantic processing of words and of pictures is a matter of debate among cognitive scientists^{1,2}. We studied the functional anatomy of such processing by using positron-emission tomography (PET). We contrasted activity during two semantic tasks (probing knowledge of associations between concepts, and knowledge of the visual attributes of these concepts) and a baseline task (discrimination of physical stimulus size), performed either with words or with pictures. Modality-specific activations unrelated to semantic processing occurred in the left inferior parietal lobule for words, and the right middle occipital gyrus for pictures. A semantic network common to both words and pictures extended from the left superior occipital gyrus through the middle and inferior temporal cortex to the inferior frontal gyrus. A picture-specific activation related to semantic tasks occurred in the left posterior inferior temporal sulcus, and word-specific activations related to semantic tasks were localized to the left superior temporal sulcus, left anterior middle temporal gyrus, and left inferior frontal sulcus. Thus semantic tasks activate a distributed semantic processing system shared by both words and pictures, with a few specific areas differentially active for either words or pictures.

Patients with visual agnosia sometimes perform semantic judgements correctly when concepts are visually presented as words, but not when presented as pictures, despite relatively spared perception of the stimulus surface structure³⁻⁶. Other patients do significantly better in semantic tasks with pictures than with the corresponding words, although they can read words aloud⁷. These and related dissociations⁸⁻¹¹, along with picture-word priming and interference studies in normal subjects¹², have led to divergent cognitive models¹. Semantic representation systems accessed by words and pictures may be segregated, and modality-specific semantic deficits may then be due to degradation of one of these systems⁹. In a contrasting view, only the procedure by which higher-level surface descriptions are mapped onto concepts differs between words and pictures, and the dissociations reflect a functional disconnection between modality-specific presemantic processing and an amodal knowledge system^{13,14}. According to a third model, the distinction between higher-level perceptual presemantic processing and modality-specific semantic processing is artificial, as even associative visual agnosics often also have higher-level perceptual deficits³.

In our experiment, subjects performed semantic or non-semantic judgements on triplets of either words or pictures¹⁵. During one task (Fig. 1a, d), the stimuli within each triplet belonged to the same category, and subjects matched the stimuli for meaning (associative semantics). During the second task (Fig. 1b, e), subjects matched stimuli for the real-life size of their referents (visual semantics)¹⁶. During a third task (Fig. 1c, f), they matched stimuli for physical size (baseline). The behavioural data are shown in Table 1.

We began by determining where word processing significantly differed from picture processing, irrespective of the task performed. We compared pictures with words separately for each of the three tasks, including the non-semantic task. Modality-specific presemantic processing was postulated where regional cerebral blood flow (rCBF) differed significantly ($P < 0.01$) between input modalities in each of these contrasts. Because of the orthogonality

TABLE 1 Behavioural data

	Reaction time (ms)		Accuracy (% correct)	
	mean	s.e.	mean	s.e.
Assoc.(W)	2,705	52	82.8	4
Assoc.(P)	2,635	57	79.5	5
Vis.sem.(W)	2,919	47	89.8	2
Vis.sem.(P)	2,698	50	77.9	3
Baseline(W)	2,639	60	71.0	6
Baseline(P)	2,818	50	71.4	6

The reaction times confirm the advantage of pictures to words during real-life size judgements, as well as the similarity in reaction time between pictures and words during semantic distance judgements¹⁶. A two-way repeated-measures ANOVA by items for the reaction times showed a significant main effect of task ($F(2,1331) = 3.4$; $P < 0.05$), no significant effect of materials, and a significant task by materials interaction ($F(2,1331) = 7.4$; $P < 0.001$). A two-way repeated-measures ANOVA by subject for the accuracies showed a significant main effect of task ($F(2,65) = 4.3$; $P < 0.05$), as the accuracies for the active conditions were higher than for the baseline. There was no significant effect of materials, nor a significant interaction. Abbreviations: Assoc., associative semantics; Vis.sem., visual semantics; P, pictures; W, words.

of these three contrasts, the probability of reaching this criterion by chance is approximately 10^{-6} . Picture-specific processing, irrespective of the task performed, occurred in the right middle occipital gyrus (42, -72, 4; BA19/37; Fig. 2b). A nearby or identical area (34, -72, -8) has been reported to be more active when subjects passively view nonsense objects than when they view visual noise patterns¹⁷. Word-specific processing, irrespective of the task, occurred in left inferior parietal lobule (-34, -36, 24; BA40; Fig. 2a).

We then determined where the semantic network overlapped between words and pictures. Overlap was defined as a significant activation during associative or visual semantic judgements compared with baseline when the tasks were performed using words

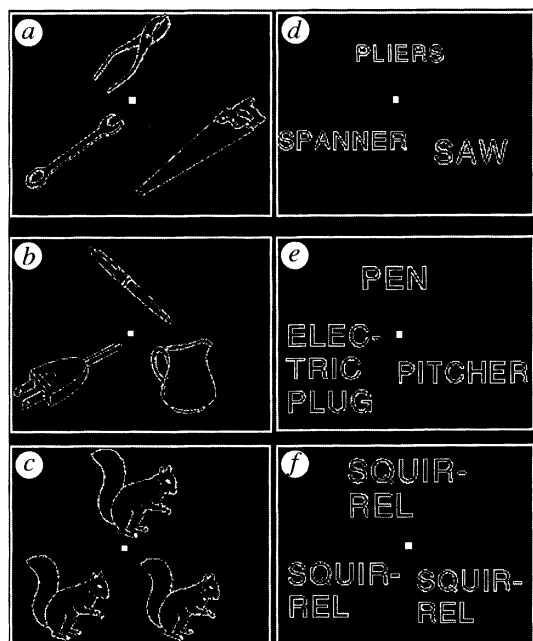


FIG. 1 Examples from stimuli for each of the picture (a-e) and word (d-f) conditions. a, d, Associative semantics; b, e, visual semantics; c, f, baseline conditions.

($P < 0.01$) and when using pictures ($P < 0.01$). Words and pictures shared a distributed semantic system (Fig. 2e and Table 2), consisting of many regions previously described in investigations of word generation and comprehension: the junction between parietal and temporal¹⁸, and between fusiform and inferior temporal cortex^{18–20}, the left middle temporal gyrus^{21,22}, and left inferior frontal gyrus (BA11/47)²³. Whatever the exact semantic operations performed by these areas, they are activated for both input modalities (Table 2). This shared semantic network was activated whichever type of semantic task subjects performed, although left hippocampal and right lateral cerebellar activations were stronger during associative than during visual semantics ($Z = 3.21$ and 3.66 , respectively). These were the only two areas in the entire brain volume searched where the two types of semantic task differed significantly ($P < 0.001$, uncorrected). The common activation pattern for the two types of semantic task may reflect anatomical overlap between an associative semantic and a visual semantic system, or activation of both systems during both semantic tasks.

Finally, we determined where semantic processing differed between words and pictures. We delineated a semantic processing network by comparing all associative semantic tasks to baseline ($P < 0.01$), and then determined within that network the areas where the difference between the semantic tasks and the baseline condition differs significantly between words and pictures ($P < 0.01$). The probability of reaching this conjoint criterion by chance is approximately 10^{-4} . Left superior temporal sulcus ($-58, -30, 0$), left anterior middle temporal gyrus ($-42, -10, -16$) and left inferior frontal sulcus ($-34, 26, 20$) were activated during associative semantics compared with baseline ($Z = 3.69$, 3.95 and 3.30 , respectively), and significantly more when these tasks were performed with words instead of pictures (interaction, $Z = 2.83$, 2.49 and 2.33) (Fig. 2c). Activation during semantic tasks specific to pictures was localized to the left posterior inferior temporal sulcus ($-38, -62, 4$) (main task effect, $Z = 2.37$; interaction, $Z = 2.57$) (Fig. 2d). No modality-specific semantic areas were revealed in the right hemisphere ($P > 0.05$, uncorrected).

None of the modality-specific areas showed any influence of the type of semantic task ($P > 0.05$, uncorrected). The absence of any effect of semantic task content contradicts the hypothesis that these areas would constitute semantic subsystems containing modality-specific knowledge. Instead, these input-specific, task-related increases in rCBF may arise because, compared with visual-perceptual tasks, semantic tasks may lead to enhanced activation of picture-specific structural descriptions²⁴ or word-specific lexical or phonological²⁵ representations. The blood flow increases may also relate to the process of accessing a shared semantic representation from these modality-specific representations^{13,14}.

Our results provide strong evidence for a distributed semantic system shared by both input modalities. In monkeys, inferotemporal cortex and ventral frontal convexity are known to be involved in object recognition^{26,27}. The anatomy and function of the common semantic processing stream we describe suggests that, phylogenetically, when primates acquired language, a pre-existing object-recognition system could have been adapted to attribute meaning to nouns. □

TABLE 2 Common semantic network

		rCBF increase (%)			
		Assoc. Words	Vis.sem.	Assoc. Pictures	Vis.sem.
L inf. front. g. (BA45)	-42,22,20	3.9 Z = 5.36	3.4	2.4 Z = 3.45	1.8
L inf. front. g. (BA11/47)	-16,30,-12	2.7 Z = 3.34	1.7	2.0 Z = 2.33	0.9
L mid. temp. g. (BA21)	-42,0,-28	2.9 Z = 3.40	2.4	1.9 Z = 2.37	1.1
L mid. temp. g. (BA21)	-58,-38,-4	3.1 Z = 4.42	2.3	1.7 Z = 2.58	2.3
L inf. temp. g. (BA20)	-44,-10,-28	3.3 Z = 4.12	2.7	1.9 Z = 2.36	1.4
L fusiform g. (BA21/37)	-46,-46,-20	2.9 Z = 3.64	2.8	2.4 Z = 3.03	2.0
L PT (BA19/39)	-40,-70,24	1.8 Z = 3.20	1.5	2.1 Z = 3.46	1.5
L sup. occ. g. (BA19)	-30,-76,40	3.3 Z = 4.00	1.9	2.2 Z = 2.75	2.4
L hippocampus (BA34)	-18,-16,-12	2.3 Z = 4.27	1.1	1.2 Z = 2.36	0.2
vermis	14,-82,-28	3.0 Z = 4.94	1.4	2.3 Z = 3.92	2.1
R cerebellum	38,-76,-48	4.8 Z = 3.80	2.5	4.4 Z = 3.41	1.9

In all of these areas, the probability of being activated in the contrast between associative semantics and baseline both for words and for pictures by chance was lower than $P < 10^{-5}$. First row: anatomical name, Talairach coordinates, percentage rCBF difference with respect to baseline, calculated by subtracting the adjusted rCBF in the respective baseline conditions from those in the semantic conditions and dividing by the respective baseline rCBF. Second row: the Z-scores for the subtraction between associative semantics and baseline. Abbreviations: Inf. front. g., inferior frontal gyrus; mid. temp. g., middle temporal gyrus; inf. temp. g., inferior temporal gyrus; PT, parietotemporal junction; sup. occ. g., superior occipital gyrus.

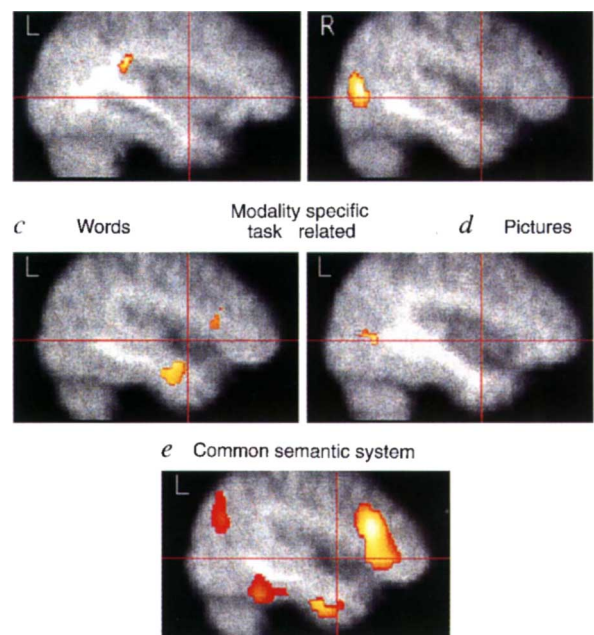


FIG. 2 Sagittal sections through the subjects' normalized averaged brains at $x = -34$ mm (left) (a), at $+42$ mm (right) (b), at -42 mm (left) (c, e), and at -38 mm (left) (d). a, Areas activated when subtracting pictures from words irrespective of task. b, Areas activated when subtracting words from pictures irrespective of task. c, Interaction between materials and task where the difference between associative semantics and baseline is larger for words. d, Interaction between materials and task where the difference is larger for pictures. e, Areas activated when the associative semantic task performed with words is compared with the non-semantic task performed with words and also activated when the same tasks are compared for pictures.

Methods

Subjects. Six right-handed male subjects, aged between 22 and 69 years, participated in this study. They gave their written, informed consent in accordance with the Declaration of Helsinki.

Stimuli. Nearly all pictures belonged to the Snodgrass–Vanderwart stimulus set²⁸. Stimulus triplets were presented on a screen 36.6 deg wide and 29.2 deg high, at 45 cm distance, for 4 s with 500 ms between presentations. Picture size was on average 11 deg, letters were 3.2 deg high and 2.5 deg wide. The sample stimulus was presented at 6.7 deg above the screen centre, the test stimuli at 5.6 (±2) deg below the screen centre and 10.3 (±2) deg to the left and the right.

Tasks. Subjects performed three different matching-to-sample tasks either with pictures (Fig. 1a–c) or with words (Fig. 1d–f). In one task (Fig. 1a, d), they selected the left or right key-press button, according to whether the left or right lower stimulus was closer in meaning to the sample stimulus above. Examples: cow: horse, bear; switch: light bulb, candle; coat: glove, sock; cucumber: tomato, corn. In a second semantic task (Fig. 1b, e), they performed a matching-to-sample task for real-life size¹⁶. Perceptual size cues were minimized. Examples: button: glass, bottle; suitcase: beaver, well; chisel: banana, kite; nail: thimble, salt cellar. The baseline task (Fig. 1c, f) consisted of a matching-to-sample task for physical size on the screen. In this baseline task, the distractor was 6% larger or smaller than the target, and the target was 6% larger or smaller than the sample size. Whether a triplet was presented as either words or pictures was entirely counterbalanced between subjects. Each subject saw each triplet of concepts only once, and each concept was used equally often in each condition.

Data acquisition. The brain was scanned with an ECAT EXACT HR+ PET scanner. The task order was counterbalanced. Each subject was scanned twice in each of 6 conditions. Each run consisted of 19 trials. Images of all 12 conditions were available for each of the 6 subjects.

Data analysis. The scans from each subject were realigned using the first image as a reference. A T1-weighted magnetic resonance image (MRI) was co-registered to the mean PET image for each subject and then stereotactically transformed²⁹ to a standard MRI template in the Talairach and Tournoux space, and the same transformation matrix subsequently applied to the PET images. Data were analysed using a randomized block design with replication and global brain activity as two covariates of no interest³⁰. The search volume went from

$z = -48$ mm to $z = 60$ mm, with a final image resolution (full width of half-maximum) $x = 16.9$ mm, $y = 18.9$ mm, $z = 20.8$ mm.

Received 22 April; accepted 24 July 1996.

1. Caplan, D. *Language: Structure, Processing and Disorders* 63–104 (MIT Press, Cambridge, MA, 1992).
2. Shallice, T. In *The Cognitive Neuropsychology of Language* (eds Coltheart, M., Sartori, G. & Job, R.) 111–127 (Lawrence Erlbaum, Hove, 1987).
3. Farah, M. *Visual Agnosia: Disorders of Object Recognition and What They Tell Us About Normal Vision* (MIT Press, Cambridge, MA, 1990).
4. Gomori, A. & Hawryluk, G. *Neurology* **34**, 947–950 (1984).
5. Albert, M., Soffer, D., Silverberg, R. & Reches, A. *Neurology* **29**, 870–879 (1979).
6. Mack, J. & Bollen, F. *Neuropsychologia* **15**, 345–348 (1977).
7. Bub, D., Black, S., Hampson, E. & Kertesz, A. *Cogn. Neuropsychol.* **5**, 27–66 (1988).
8. Beauvois, M. *Phil. Trans. R. Soc. Lond. B* **298**, 35–47 (1982).
9. Warrington, E. & McCarthy, R. *Neuropsychologia* **32**, 1465–1473 (1994).
10. Warrington, E. & Shallice, T. *Brain* **107**, 829–853 (1984).
11. McCarthy, R. & Warrington, E. *Nature* **334**, 428–430 (1988).
12. Glaser, W. *Cognition* **42**, 61–105 (1992).
13. Caramazza, A., Hillis, A., Rapp, B. & Romani, C. *Cogn. Neuropsychol.* **7**, 161–189 (1990).
14. Butterworth, B., Howard, D. & McLoughlin, P. *Neuropsychologia* **22**, 409–426 (1984).
15. Howard, D. & Patterson, K. *Pyramids and Palm Trees: a Test of Semantic Access from Pictures and Words*. (Thames Valley, Bury St Edmunds, 1992).
16. te Linde, J. J. *Exp. Psychol. Learn. Mem. Cogn.* **8**, 584–598 (1982).
17. Martin, A., Wiggs, C., Ungerleider, L. & Haxby, J. *Nature* **379**, 649–652 (1996).
18. Demonet, J.-F. et al. *Brain* **115**, 1753–1768 (1992).
19. Price, C., Moore, C. & Frackowiak, R. S. J. *Neuroimage* **3**, 40–52 (1996).
20. Martin, A., Haxby, J., Lalonde, F., Wiggs, C. & Ungerleider, L. *Science* **270**, 102–105 (1995).
21. Price, C., Wise, R. & Frackowiak, R. S. J. *Cereb. Cort.* **6**, 62–70 (1996).
22. Damasio, H., Grabowski, T., Tranel, D., Hichwa, R. & Damasio, A. *Nature* **380**, 499–505 (1996).
23. Petersen, S., Fox, P., Snyder, A. & Raichle, M. *Nature* **331**, 585–589 (1988).
24. Ridgway, J. & Humphreys, G. *Cogn. Neuropsychol.* **4**, 131–185 (1987).
25. Coltheart, V., Patterson, K. & Leahy, J. Q. J. *Exp. Psychol.* **47A**, 917–955 (1994).
26. Mishkin, M., Ungerleider, L. & Macko, K. *Trends Neurosci.* **6**, 414–417 (1983).
27. Wilson, F., Scalaidhe, S. O. & Goldman-Rakic, P. *Science* **260**, 1955–1958 (1993).
28. Snodgrass, J. & Vanderwart, M. J. *Exp. Psychol. Hum. Learn. Mem.* **6**, 174–215 (1980).
29. Friston, K. et al. *Hum. Brain Mapp.* **2**, 189–210 (1995).
30. Friston, K. et al. *Hum. Brain Mapp.* **2**, 165–189 (1995).

ACKNOWLEDGEMENTS. We thank the volunteers who participated in this study. R.V. is supported by the Belgian National Research Council.

CORRESPONDENCE and requests for materials should be addressed to C.P. (e-mail: cprice@fil.ion.ucl.ac.uk).

Contribution of human hippocampal region to novelty detection

Robert T. Knight

Department of Neurology, Center for Neuroscience, University of California, Davis, Veterans Medical Center, 150 Muir Road, Martinez, California 94553, USA

THE ability to respond to unexpected stimuli (the ‘orienting response’) is a fundamental characteristic of mammalian behaviour¹, but the brain mechanisms by which novelty is detected remain poorly defined. Electrophysiological recordings of scalp and intracranial event-related potentials (ERPs) have shown that novel stimuli activate a distributed network involving prefrontal and posterior association cortex^{2–6}. In addition, ERP^{7,8} and single-neuron^{9,10} recordings, as well as neuroimaging¹¹ and modelling¹² studies, have suggested that temporal cortical regions, including the hippocampus, are also involved. To examine further the role of the medial temporal lobe in novelty processing, I measured physiological responses to novel auditory and tactile stimuli in patients with damage to the posterior hippocampal region. In normal control subjects, unexpected novel stimuli produce a characteristic ERP signal, accompanied by an autonomic skin response. Both responses are reduced in hippocampal lesion patients, whereas the response to expected control stimuli is unaffected. Thus the hippocampal region, in addition to its known role in memory formation, is an essential component of the distributed limbic–cortical network that detects and responds to novel stimuli.

Distinctive ERPs are recorded from the scalp of humans in

response to stimuli that stand out from others in the perceptual stream. Such events may be salient either because they are relevant signals requiring a response or because they are novel and surprising. The ERP scalp signature for these two types of signals exhibits different latencies and topographies. Novel stimuli or stimuli sufficiently deviant from the ongoing sensory environment trigger a positive brain potential maximal over fronto-central regions (250–550 ms; P3a) which is proposed to be a central marker of the orienting response^{13,14}. Voluntary detection of a task-relevant target stimulus elicits a longer latency positivity (300–600 ms; P3b) which is maximal over parietal sites. Target stimuli may also elicit overlapping, early latency P3a activity dependent on the degree of target deviancy from background events.

I performed experiments in which target and novel stimuli (tones, complex tones; finger taps, wrist shocks), embedded in trains of repetitive background stimuli, were delivered randomly to subjects engaged in detection tasks. Target stimuli differed only slightly from the background stimuli, whereas novel stimuli differed substantially in order to be highly distinctive. Subjects were patients with unilateral brain lesions centred in the posterior hippocampal region (Fig. 1) and age-matched controls. They sat in a sound-attenuated booth and pressed a button upon detection of a designated target stimulus during separate auditory and somatosensory testing sessions. ERPs were recorded from scalp electrodes to background, target and novel stimuli.

In controls, target stimuli elicited a prominent parietal–maximal P3b in both modalities and a small, shorter latency fronto-central P3a. Novel stimuli in both modalities generated enhanced amplitude P3a potentials which had more fronto-central scalp distributions ($P < 0.001$ for target versus novel scalp topography; Figs 2 and 3) and shorter latencies than the target P3bs (novels at Fz: auditory, 365 ms; somatosensory, 354 ms; targets at Pz: auditory, 431 ms, somatosensory, 410 ms; $P < 0.002$ for both modalities).

Hippocampal lesions had differential effects on P3b and P3a