Unitary vs multiple semantics: PET studies of word and picture processing

P. Bright,[a,*] H. Moss,[a] and L.K. Tyler[a,b]

[a] *Department of Experimental Psychology, University of Cambridge, Cambridge, UK*
[b] *Wolfson Brain Imaging Unit, University of Cambridge, Cambridge, UK*

Accepted 23 January 2004

Abstract

In this paper we examine a central issue in cognitive neuroscience: are there separate conceptual representations associated with different input modalities (e.g., Paivio, 1971, 1986; Warrington & Shallice, 1984) or do inputs from different modalities converge on to the same set of representations (e.g., Caramazza, Hillis, Rapp, & Romani, 1990; Lambon Ralph, Graham, Patterson, & Hodges, 1999; Rapp, Hillis, & Caramazza, 1993)? We present an analysis of four PET studies (three semantic categorisation tasks and one lexical decision task), two of which employ words as stimuli and two of which employ pictures. Using conjunction analyses, we found robust semantic activation, common to both input modalities in anterior and medial aspects of the left fusiform gyrus, left parahippocampal and perirhinal cortices, and left inferior frontal gyrus (BA 47). There were modality-specific activations in both temporal poles (words) and occipitotemporal cortices (pictures). We propose that the temporal poles are involved in processing both words and pictures, but their engagement might be primarily determined by the level of specificity at which an object is processed. Activation in posterior temporal regions associated with picture processing most likely reflects intermediate, pre-semantic stages of visual processing. Our data are most consistent with a hierarchically structured, unitary system of semantic representations for both verbal and visual modalities, subserved by anterior regions of the inferior temporal cortex.
© 2004 Elsevier Inc. All rights reserved.

1. Introduction

Conceptual knowledge lies at the heart of the cognitive system, supporting a wealth of mental processes, including language comprehension and production, reasoning, and object recognition. This paper addresses a central issue concerning the functional and neural architecture of the conceptual system: are all these processes subserved by a unitary system of conceptual representations, or are there separate representations for the same concept for different modalities of input or output? This issue has a long history in cognitive psychology and neuropsychology, but no consensus has yet emerged. Some researchers have argued for a unitary semantics account, proposing that there are distinct conceptual representations for the verbal (word) and visual (object) input modalities (e.g., Paivio, 1971, 1991;

Shallice, 1988; Warrington & Shallice, 1984). Others have rejected this position, claiming that all processing routes converge on a single set of conceptual representations common to both modalities (e.g., Caramazza, Hillis, Rapp, & Romani, 1990).

It is important to note at the outset that many of the most influential neuropsychological studies that have been presented as evidence for modality-specific semantics have employed visually presented words and visually presented pictures (e.g., McCarthy & Warrington, 1986, 1988; Shallice, 1988, 1993; Warrington & McCarthy, 1994). Thus, in one sense, such investigations do not test whether or not the semantic system is "amodal" because both inputs are visual. Furthermore, "verbal" does not correspond in a straightforward manner to a sensory modality at all, but reflects some combination of content, context, and format (Plaut, 2002). Despite such criticisms, a large number of functional neuroimaging studies which have addressed the unitary vs multiple semantics controversy have also

---

* Corresponding author. Fax: +44-1223-766452.
E-mail address: pbright@csl.psychol.cam.ac.uk (P. Bright).

employed visually presented pictures and words (see below). The result is that, unlike visual vs auditory presentation, the visual presentation of pictures *and* words does not constitute an orthogonal comparison. It follows that the term "modality" might more accurately be replaced with "material," because the comparison in these cases is not of sensory systems of input (auditory, visual, and tactile) but of objects (or pictures of them) vs printed words. However, for the value of consistency with previous literature, we continue to adopt the term "input modality" in our considerations of whether the conceptual knowledge accessed by pictures and words form two neurally distinct components of the semantic system (modality-specific) or whether both stimulus types converge on to the same set of representations (unitary semantics).

The notion of modality-specific conceptual representation stems predominantly from Paivio's Dual Coding Theory. On this account, human cognition has developed to deal simultaneously with verbal and non-verbal objects and events, giving rise to two independent but interconnected symbolic systems. "One (the image system) is specialised for dealing with perceptual information concerning non-verbal objects and events. The other (the verbal system) is specialised for dealing with linguistic events" (Paivio, 1971, p. 379). The verbal system deals *directly* with linguistic input while also serving a symbolic function with respect to non-verbal input, which has direct access only to a non-verbal (or *imagery*) system. The two systems are assumed to be functionally and structurally distinct although interconnected by referential relations between representations in the two.

Neuropsychological evidence for separate visual and verbal semantic systems was first provided by Warrington (1975), in a study of two visual agnosic patients, one of whom (EM) appeared to be able to recognise a visually presented object, but not its name, while the other (AB) was able to recognise the name of an object but not its visual representation. Further cases of related dissociations in picture vs word based semantic judgements have supported the notion of neurally distinct, modality-specific semantic representations (e.g., Ferreira, Giusiano, Ceccaldi, & Poncet, 1997; Warrington & McCarthy, 1994; Warrington & Shallice, 1984). A strong version of the multiple, modality-specific semantics position implies the existence of distinct, self-contained semantic systems duplicated across modalities. However, most proponents propose that the systems are in constant communication. For example, Shallice (1988) favours the notion of distributed but interconnected modality-specific subsystems, each of which can only be accessed directly from the associated input modality.

A recent study by Saffran, Coslett, Martin, and Boronat (2003) presents data from a patient with progressive fluent aphasia (BA) who showed a clear performance deficit for words relative to pictures on a variety of semantic tasks. They interpret these findings as supporting a distributed semantic system, with knowledge distributed across distinct and neurally separable subsystems (e.g., colour is represented in the visual system, sounds in the auditory system, knowledge of object use in the kinaesthetic system, etc.). Thus, the impairment for verbal stimuli shown by BA reflected damage to a single component of the semantic system (the 'propositional/encyclopaedic' store). The type of knowledge tapped by a task is assumed to reflect the input route primarily involved in its acquisition. Pictures of objects initially access structural descriptions, and only subsequently engage motor, sensory, and propositional/encyclopaedic information. The authors argue that the semantic system is multimodal, with different types of information stored in different brain regions and in different representational formats. However, Lambon Ralph and Howard (2000), describing a semantic dementia patient (IW) with a similar performance profile, interpreted the disproportionate semantic impairment for words with respect to the theory of 'privileged accessibility' to a unitary semantic system (Caramazza et al., 1990) (see below).

Other researchers have presented evidence that semantic *processes* are lateralised, with only the left hemisphere supporting naming (e.g., Coslett & Saffran, 1992; Luzzatti, Rumiati, & Ghirardi, 1998). Patients with semantic impairments resulting from left-sided damage have relatively preserved comprehension of visually presented objects and their functions, but they are unable to name them. Thus, in the context of the multiple semantics account, the left hemisphere supports verbal semantics and the right hemisphere supports visual semantics. However, problematic for this account is the observation that semantic access and retrieval from visual input in optic aphasia (typically caused by lesions to left medial occipital cortex) is not as preserved as originally claimed (e.g., Hillis & Caramazza, 1995; Riddoch & Humphreys, 1987). This suggests that a distinction between right and left hemisphere processing on the basis verbal vs visual semantics may not be as straightforward as claimed by these authors.

A persistent difficulty for studies which propose modality- or material-specific semantics is to demonstrate unequivocally that the observed dissociation is, in fact, located at the level of conceptual representation, rather than within the pre-semantic representations or processes necessary for *access* to the conceptual system. Verbal and non-verbal tasks are not always comparable in terms of complexity or other relevant factors, which may have led to dissociations more apparent than real (e.g., Caramazza et al., 1990; Riddoch, Humphreys, Coltheart, & Funnell, 1988). Unless each task unequivocally taps into conceptual representations, modality-

specific effects may emerge because of impaired pre-semantic representations (e.g., visual-structural rather than visual-semantic).

Caramazza et al. (1990) claims that *all* modality-specific semantic deficits reported in the literature can be implemented within the architecture of unitary semantics theory, and argues that a common conceptual system will be recruited during the processing of an item, irrespective of modality, while allowing for an "asymmetry" in recruitment according to stimulus modality. Their organised unitary conceptual hypothesis (OUCH) assumes that perceptually salient features of a visually presented object will have privileged access to its semantic representation even though there is a single, undifferentiated representation for the visually presented object and the word for that object. Basic to this theory is that for any item, particular predicates are more important than others in defining its meaning. Thus, while the semantic representation of a visually presented object and its verbal description is the same, the procedure by which that representation is accessed may differ. A visually or aurally presented word will activate the lexicon, which will, in turn, activate the semantic properties which define its meaning, but a visually presented object will directly access those same semantic properties. These assumptions, of privileged accessibility and bias towards particular subcomponents of a semantic representation, are used to explain how modality-specific semantic effects can arise from damage to a unitary conceptual system.

Lambon Ralph, Graham, Patterson, and Hodges (1999) examined definitions provided by nine patients with semantic dementia to concepts presented in spoken word or picture form. They found that the extent of conceptual knowledge successfully tapped by verbal input (the concept label) closely predicted the quality of conceptual information produced following pictorial input, consistent with both inputs tapping a common semantic system. Although two patients were particularly poor in defining concepts relating to picture presentations, the authors claimed that this apparent modality effect was caused by additional pre-semantic visuo-perceptual deficits in these patients.

Thus, although neuropsychological dissociations have often been taken as evidence for modality-specific conceptual systems, the data are open to alternative explanations consistent with a unitary modality-independent system. Nevertheless, the unitary semantics model is not able to fully explain the size of the dissociations reported in task performance, such as the relative preservation of gesturing meaning relative to naming visually presented objects across all levels of severity among optic aphasia patients (Plaut, 2002). Recently, some authors have suggested a 'middle ground' on the unitary vs multiple semantics debate. For example, Plaut (2002) presents a computational account of optic aphasia which is consistent with the same semantic representation activated regardless of the mode of presentation, but some parts of the system becoming relatively specialised for particular types of input over others (leading to more selective deficits than could otherwise be observed).

Neuroimaging techniques offer the capacity to explore the functional neuroanatomy of the conceptual system, and, unlike the lesion-deficit approach, can directly address the extent to which neural activity in the healthy brain is differentiated as a function of the form of input. Furthermore, functional neuroimaging provides a systems-level approach whereby the distributed pattern of recruitment across the brain can be explored.

The first direct comparison of word and picture processing was undertaken by Vandenberghe, Price, Wise, Josephs, and Frackowiak (1996) using PET, the results of which have been interpreted as favouring a unitary conceptual system, undifferentiated by modality. However, a number of significant material-specific effects were in fact observed in this study. Healthy subjects carried out three match-to-sample tasks on triplets of either words or pictures. The tasks were based on (i) meaning (termed *associative semantics*), (ii) real-life size (*visual semantics*), or (iii) physical size (*baseline*). In both semantic tasks, a large area of left hemisphere activation, common to pictures and written words, was found which extended from superior occipital gyrus through middle, inferior temporal, and fusiform gyrus to inferior frontal gyrus. No differences between word and picture processing were found as a function of semantic task, and there were no modality-specific semantic effects in the right hemisphere. Modality-specific activations which occurred across all tasks (including baseline) were taken to represent pre-semantic processing effects. For pictures relative to words (irrespective of task), there was activation of right middle occipital gyrus (BA 19/37). The reverse contrast (words relative to pictures) produced activation of left inferior parietal lobe (BA 40). At face value, the results favoured an extensive distributed conceptual system underlying both words and pictures. Nevertheless, in addition to pre-semantic modality effects, there was also significant word-specific semantic recruitment for verbal materials in the left hemisphere (discussed below).

Further evidence for modality-independent conceptual activation comes from a PET study by Moore and Price (1999) in which the left fusiform gyrus (BA 20/37) was recruited during the processing of meaningful over non-meaningful stimuli, irrespective of input modality (using silent naming and viewing of words and objects). Other regions of common activation in this study included right inferior frontal gyrus (BA 47), and anterior cingulate (BA 24/6), most probably reflecting the greater overall demands of the semantic over non-semantic task conditions. A meta-analysis by Devlin et al. (2002a) on a

set of seven PET studies (including Moore & Price, 1999) also found that the same network was activated by words and pictures, although the specific regions of common recruitment were not described.

In addition to the neuroimaging studies that have directly contrasted conceptual activation for words vs pictures, a number have compared conceptual activations for a range of other modality contrasts. Buchel, Price, and Friston (1998) compared cortical activations in congenitally blind, late blind, and sighted subjects in a PET study during visual or tactile reading. All subject groups showed activation in overlapping regions of the left fusiform (BA 37) for visual and tactile reading for meaningful words relative to non-word letter strings. The results were consistent with this region being unaffected by variations in sensory input characteristics and linking converging inputs across different modalities, giving rise to conceptual representations. Other studies have also highlighted at least partially overlapping regions of the left fusiform gyrus across different modalities (e.g., generating mental images of object words relative to listening to abstract words (D'Esposito et al., 1997); making semantic relative to phonological decisions on auditorially presented words (Demonet et al., 1992); naming words and objects relative to letters and colours (Price, Moore, Humphreys, Frackowiak, & Friston, 1996); semantic classification of words characterised by auditory, visual or non-sensory features (Noppeney & Price, 2002)).

One of the most influential bodies of recent research on the structure and organisation of semantic memory has been developed by Martin and colleagues (Martin, 2001; Martin, Haxby, Lalonde, Wiggs, & Ungerleider, 1995; Martin, Ungerleider, & Haxby, 2000). On the basis of a series of neuroimaging investigations, these authors propose that qualitatively different forms of information are represented in lateral and ventral temporal cortices and that the location of these sites appears to parallel the organisation of the sensory and motor systems, a view also reflected in the recent neuropsychological work, for example by Saffran et al. (2003), described above. In an fMRI study which examined activations during viewing, match-to-sample, and naming of pictures of animals and tools they found distinct activations for these two categories of object (Chao, Haxby, & Martin, 1999). However, the same patterns of activity were observed when subjects read silently and answered questions about the written names of animals and tools, suggesting the semantic processing of words and pictures engage a common set of representations.

The inferior frontal gyrus (IFG) has also been consistently activated, irrespective of the modality of input in several neuroimaging studies (e.g., Thompson-Schill, Aguirre, D'Esposito, & Farah, 1999; Wagner, Desmond, Demb, Glover, & Gabrieli, 1997). However, it is unclear whether IFG activation is associated with conceptual representation or processing per se or whether it relates to more general differences in the overall cognitive demands between test and baseline conditions. There is no evidence that patients with damage restricted to IFG have semantic deficits. However, some studies which have explicitly attempted to match semantic and non-semantic conditions on overall "difficulty" have reproduced robust IFG activations during semantic relative to non-semantic conditions (e.g., Demb et al., 1995).

While several studies have been interpreted as consistent with a unitary conceptual system, as outlined above (e.g., Vandenberghe et al., 1996), all of these studies either (i) show *some* modality-specific effects in addition to the areas of common activation or (ii) present insufficient information to exclude such effects. Consideration of semantic modality effects requires an important distinction to be made between modality-specific activation of conceptual knowledge and modality-specific activations associated with earlier stages of input processing.

Two posterior regions have been identified as candidates for early, pre-semantic stages of input processing. First, the lateral occipital complex is preferentially activated by pictures of objects with clear shape interpretations relative to non-identifiable visual textures or noise patterns (e.g., Kanwisher, Chun, McDermott, & Ledden, 1996; Malach et al., 1995), a finding which is consistent with the deficits associated with damage to this region (e.g., Farah, Hammond, Mehta, & Ratcliff, 1989; Feinberg, Dyckes-Berke, Miner, & Roane, 1995). Malach et al. (1995) found no evidence for differential activation in the lateral occipital complex as a function of familiarity (pictures of real-life objects vs degraded, non-identifiable objects), suggesting that the region is not involved in semantic stages of representation. In a recent literature review, Grill-Spector, Kourtzi, and Kanwisher (2001) conclude that lateral occipital complex functions as a general-purpose system for the analysis of object shape and is not associated with a conceptual level of representation. They propose that it is hierarchically organised, with sensitivity to the local features of an object in more posterior, retinoptic areas, and with more global representations (whole or half objects) associated with activations in anterior-lateral areas.

A second region, located in the middle portion of the left fusiform gyrus (BA 37) and called the Visual Word Form Area (VWFA) responds maximally to visually presented words independently of their location on the retina (Cohen et al., 2000). This lies posteriorly to those regions of the fusiform gyrus which are associated with high over low semantic processing demands (Cohen et al., 2000). In contrast, within the VWFA, recruitment seems to be associated with a distinction between alphabetic material (e.g., real words *or* consonant strings) and non-alphabetic stimuli (e.g., false fonts or fixation) but not related to semantics. Furthermore, VWFA

recruitment does not seem to be differentiated as a function of whether the stimuli are real words or pseudowords (e.g., Dehaene, Le Clec, Poline, Le Bihan, & Cohen, 2002; Fiez & Petersen, 1998), although a relative reduction in activation is observed for consonant strings (Cohen et al., 2002) consistent with a sensitivity of this region to orthographic regularities. Thus, the data are consistent with functional differentiation along the extent of the fusiform gyrus, with higher level, semantic representation in anterior regions, and pre-semantic form and orthography-based representation in more posterior regions.

Once activation of these pre-semantic and/or intermediate level input regions is excluded from consideration, there is limited support for modality-specific semantic recruitment. In the Vandenberghe et al. (1996) study, word-specific semantic activations were found in left superior temporal sulcus, left anterior middle temporal gyrus, and left inferior frontal sulcus. The only picture-specific semantic activation was observed in left posterior inferior temporal sulcus. Consistent with the Vandenberghe study, Moore and Price (1999) found semantic activation for words over pictures in the left superior temporal gyrus (BA 22/41), extending to include the supramarginal gyrus, although this finding may have related more to differing phonological requirements among the tasks than to modality-specific semantic recruitment per se (e.g., Demonet, Price, Wise, & Frackowiak, 1994; Moore & Price, 1999). Picture-specific activation was found in ventral occipitotemporal cortices bilaterally (BA 19) including lateral occipital cortex, probably relating to pre-semantic object processing.

The overall picture that has emerged from the neuroimaging literature remains unclear with respect to input modality effects. Although it is relatively well established that the cortical recruitment for words and pictures differs in more posterior, pre-semantic processing areas (perhaps reflecting the distinction between visual form and orthography-based representations), whether there is modality-specificity at the level of semantic representations remains poorly understood. In this study, we directly address the question of whether there are distinct (separable) neural regions that underlie the semantic representation of objects and words by comparing brain activations associated with each type of material. If activation of anterior temporal regions associated with semantic representation is essentially undifferentiated by the modality of input (objects or words), it will be most consistent with unitary semantics theory (Caramazza, 2000). Conversely, a finding that recruitment in semantic regions differs according to the type of material presented would be more consistent with Dual Coding Theory (Paivio, 1971, 1986) and multiple semantics theory described by Warrington, Shallice and others (e.g., Shallice, 1988; Warrington & Shallice, 1984).

We present a meta-analysis of four PET studies (three semantic categorisation tasks and one lexical decision task), two using pictures as stimuli and two using words. Although in comparison to semantic categorisation, lexical decision is likely to place weaker demands on the semantic system, the task has been shown to robustly recruit regions of the brain involved in semantic processing (see Bookheimer, 2002 for a recent review). In a comparison of a lexical decision and a semantic categorisation task (using words), we found additional frontal and cerebellar recruitment during the latter (presumably relating to working memory, attentional, and semantic demands not present during lexical decision). However, a direct comparison of the two tasks revealed no significant differences in the pattern or extent of recruitment (Devlin et al., 2002b). On this basis, we chose to include the lexical decision task. Methodologies and procedures, stimulus sets, and scanner settings were largely held constant across all tasks. In taking this approach we believe we can explore and compare regional recruitment for word and picture processing with a high level of conviction that differential activations relate to differences in stimulus-specific processing rather than to methodological or task differences per se. Although it might be argued that, by restricting our study to categorisation or lexical decision conditions, we compromise our capacity for exploring the neural correlates of conceptual processing at different levels or types of demand, we know that even subtle methodological differences in a task paradigm can have wide ranging effects on regional activations. Consequently, we have placed our emphasis on examining modality differences in words and pictures within a relatively unified task structure. However, we should note that, although words, by definition, evoke naming responses we do not know whether names are automatically evoked for pictures. This is an important issue which we will return to in the discussion.

There are two competing predictions regarding the pattern of recruitment during conceptual processing of words and pictures: (a) if conceptual representations for words and pictures are separable and non-overlapping, we would expect them to involve distinct semantic processing regions; (b) if the unitary conceptual system position is correct, there should be extensive coactivation for word and picture categorisation in the more anterior, semantic regions, although there may be differential activations in posterior areas relating to modality-specific pre-semantic effects.

2. Materials and methods

2.1. Subjects

A total of 38 subjects (mean age = 30 years; range = 21–48 years; 37M, 1F) were tested across the four tasks described in this study. All were right-handed,

native English speakers without any known history of neurological or psychiatric illness. No subject participated in more than one task. Each gave informed consent and was medically screened for PET prior to entering the scanning room.

2.2. Words 1: Lexical decision (Devlin et al., 2002b)

Twelve participants (mean age 32, range 21–44 years) performed a lexical decision task on visually presented words. Ten scans[2] were acquired for each subject; two for each of four semantic categories (animals, fruit/vegetables, vehicles, and tools) and 2 baseline scans. The baseline task required subjects simply to "find the *x*" in orthographically illegal letter strings (e.g., RFSTEN). Words were matched across scans, categories, and domains on six variables known to affect lexical or semantic processing: (i) familiarity [MRC psycholinguistic database (Coltheart, 1981) and CSL norms]; (ii) concreteness (MRC database and CSL norms); (iii) neighbourhood size (the number of alternative words which differ from the target word by only one letter in any position; Macquarie neighbourhood database program); (iv) written word frequency in British English usage (Celex database; Baayen & Pipenbrook, 1995); (v) number of letters; and (vi) number of syllables. Baseline stimuli matched the non-lexical and non-semantic properties of the test stimuli (orthographic visual stimulation, motor response) and the same presentation and response time windows were used to maximise the likelihood that both tasks placed similar demands on working memory and attentional processes. In the first 45 s of each scan, the time most relevant to signal acquisition (Silbersweig, Stern, Frith, & Cahill, 1993), 10 words from a single category appeared in a pseudorandom order, intermixed with five non-words. Each item was presented for 500 ms, with 2500 ms between successive items. The same words were repeated in a different order for the remaining 45 s of the scan, intermixed with five different pseudowords. No words were replicated across scans. Subjects responded with either a right button press (word) or left button press (non-word). In the baseline task, subjects were asked to press a right mouse button if the string contained an "*x*" and the left hand button if it did not.

2.3. Words 2: Semantic categorisation (Devlin et al., 2002b)

Eight participants (mean age 28, range 21–46 years) performed a semantic categorisation task on visually presented words (animals, fruits/vegetables, vehicles, and tools). There were two semantic conditions [living and non-living things]. Twelve scans were acquired for each subject, four for each domain (living and non-living) and four baseline scans. For the semantic condi-

tions, a trial comprised three lower case cue-words presented for 200 ms each, followed by a target in capital letters, also for 200 ms. The interstimulus duration was 400 ms, with 1500 ms between successive trials. Participants were required to indicate whether the target was a member of the set defined by the three cue items or not by pressing the right (SAME) or left (DIFFERENT) mouse button (e.g., "crow," "owl," "stork," and "WHEAT"). Twelve trials (cue triplets plus target) were presented for the initial 45 s of the scan during the rising phase of the head curve followed by a blank screen for the remaining 45 s, where the participants were instructed to rest. The words were matched for word frequency, familiarity, and letter length (MRC Psycholinguistic Database; (Coltheart, 1981), and Celex databases (Baayen & Pipenbrook, 1995), and CSL norms). Trials in the baseline condition consisted of three variable length strings of the same letter (e.g., "kkkk," "kkk," and "kkkkk") and a target string of the same capitalised letter ("KKKKK") or a different capitalised letter (e.g., "BBBBB"). With this procedure, we were able to maintain the same stimulus and response characteristics as the semantic categorisation task, but with no lexical or semantic component. There were a total of 96 semantic categorisation trials and 48 baseline trials.

2.4. Pictures 1: Semantic categorisation (previously unpublished data)

We used the same semantic categorisation with pictures as stimuli instead of written words. Ten scans were acquired for each of nine participants (mean age 29, range 21–43 years), two for each of four semantic conditions (animals, fruit/vegetables, vehicles, and tools) and 2 baseline scans. For all semantic conditions, a trial consisted of three pictures presented sequentially for 400 ms each, followed by a framed target picture. The interstimulus duration was 200 ms, with 2217 ms between the offset of the target and the first cue of the following trial. Participants were required to indicate whether the target was a member of the set defined by the three cue items or not by pressing the left (SAME) or right (DIFFERENT) mouse button (e.g., butterfly, caterpillar, beetle, ant). Subjects made a 'same' response by pressing the left mouse button and a 'different' response by pressing the right mouse button. In each condition, the stimuli were always drawn from the same main category, but, for 'different' trials, the target was from a different subcategory (e.g., pig, horse, sheep, and wasp). In all conditions there were equal numbers of 'same' and 'different' trials. The cue triples and targets in each category were matched on the following variables: (i) familiarity, using the MRC Psycholinguistic databases (Coltheart, 1981) and CSL norms; (ii) visual complexity (matched across living and non-living do-

mains); and (iii) semantic relatedness (a measure of the relatedness of each cue triplet to its target), based on pre-test data collected at the CSL.

The baseline task employed simple two-dimensional shapes (each of which varied in shape and colour) and required participants to indicate whether the target was of the same 'family' as the preceding three cues. We chose this baseline task as a means of extracting out non-semantic visual processing activations from our test conditions. However, as some of the shapes could be named (e.g., square, circle), it remained possible that semantic processing areas were also recruited during the task. The baseline stimuli were tested for relatedness in the same way as the test items (see above). In each condition, 12 trials were presented during the initial 53 s of the scan followed by a blank screen for the remaining 37 s, where the participants were instructed to rest. All items (including baseline stimuli) were scaled to the same size. There were a total of 96 test trials and 48 baseline trials.

2.5. Pictures 2: Semantic categorisation (Tyler et al., 2003)

Nine participants (mean age 28, range 21–48 years) performed a semantic categorisation task on pictures. The experimental design, test stimuli, and the timings employed were identical to those described in the picture semantic categorisation task above, except that a different baseline was employed. The new baseline consisted of meaningless, simple shapes made up of combinations of small squares, which varied in terms of the colour and number of squares in each object. Sets of three simple shapes were presented sequentially, followed by a target shape which was either from the same "family" of shapes or from a different family. All baseline items were pre-tested for relatedness in the same way as the test items.

2.6. Data acquisition and analysis

All studies were performed at the Wolfson Brain Imaging Centre in Cambridge, England on a GE Advance PET Scanner (General Electric Medical Systems, Milwaukee, Wisconsin). It comprises 18 rings of crystals, which results in 35 image planes, each 4.25 mm thick. The axial field-of-view is 15.3 cm thus allowing for whole brain acquisition. In each experiment, participants received a bolus of 300 MBq of H_2O^{15} before each scan for a total radiation exposure of 4.2 mSv. The emission data was acquired with the septa retracted (3D mode) and reconstructed using the PROMIS algorithm (Kinahan & Rogers, 1989) with an unapodised Colsher filter. Corrections were applied for randoms, scatter, attenuation, and dead time. The voxel sizes were 2.34, 2.34, and 4.25 mm.

Stimuli were presented to subjects, and behavioural data collected, using DMDX software (Forster & Forster, 1991). The functional images were realigned using SPM (Friston, Holmes, Worsley, & Poline, 1995); SPM99, Wellcome Institute of Cognitive Neurology, (www.fil.ion.ucl.ac.uk). Translation and rotation corrections did not exceed 4 mm and 3°, respectively for any of the participants. The mean image created by the realignment procedure was used to determine the parameters for transforming the images onto the Montreal Neurological Institute (MNI) mean brain. These parameters were then applied to the functional images (Ashburner & Friston, 1997) and the image was resampled into isotropic $2 \, mm^3$ voxels. Finally, each image was smoothed with a 16 mm at full-width half-maximum (FWHM) Gaussian filter.

The SPM software was used to compute a within-subjects analysis (fixed-effects) using the general linear model (Friston et al., 1995). All main effects were calculated separately for each independent study and then combined using a conjunction analysis, to ensure that activations were common to all contrasts entered into the analysis (Price & Friston, 1997). A conjunction analysis, now routinely implemented for PET and fMRI studies, calculates main effects by summing simple main effects and excluding regions where there are significant differences between the simple main effects (e.g., Friston, Holmes, Price, Buchel, & Worsley, 1999; Lee, Robbins, Graham, & Owen, 2002; Noppeney & Price, 2002; Tyler, Russell, Fadili, & Moss, 2001). Scan-to-scan variability within a PET session and the session-by-contrast interactions are approximately the same in PET, resulting in the difference between inferences based on first-level (fixed effects) and second-level (random effects) greatly reduced in comparison to fMRI (Friston et al., 1999). The advantage of conjunctions is that they allow robust inferences to be drawn about group effects, which in turn reflect aspects of functional anatomy that may be typical of the population from which the group was sampled (Friston et al., 1999). A statistical threshold appropriate for each conjunction was chosen by computing .001 to the power of $1/n$ (where n equals the number of contrasts entered into the analysis). Thus, for a conjunction of four contrasts (Results, Section 3.1), a p value of .178 (familywise error corrected) was used. For a conjunction of two contrasts (Results, Section 3.2) a p value of .032 was used. In each case, these values provide a *conjunction p* value (uncorrected) of .001. This choice of statistical thresholding was informed by comments from K. Friston, C. Price, and J. Andersson in the SPM mailing list archives (http://www.jisc-mail.ac.uk/lists/spm.html).

Masking procedures (Results, Section 3.2) implemented in SPM were conducted to distinguish between common and specific activated clusters when comparing conditions of interest (words and pictures). All masks

were thresholded at $p < .05$ uncorrected, except where specified. With this masking procedure, results are restricted either to those voxels that fall within the area determined by the mask (inclusive masking) or to regions outside the mask area (exclusive masking), enabling regions of activation overlap and differentiation among contrasts to be identified (e.g., Chaminade & Decety, 2002; Dolan, Morris, & de Gelder, 2001; Morcom, Good, Frackowiak, & Rugg, 2003; Pochon et al., 2002). Analysis of the clusters of interest was performed using condition-specific parameter estimates which reflects adjusted regional cerebral blood flow in the different conditions relative to the fitted mean expressed as a percentage of whole brain mean blood flow. Since SPM coordinates are given in MNI space the results were converted to Talairach space with a non-linear transform and mapped onto Brodmann areas with reference to the Talairach and Tournoux (1988) brain atlas.

3. Results

3.1. Semantic activations

To determine the semantic network activated by words and pictures, we entered all semantic conditions against baseline for each study in a conjunction analysis. Table 1 presents cluster extents and voxel-level statistics. In the left hemisphere, the largest cluster encompassed inferior frontal gyrus (BA 47, extending superiorly to BA 45) and anterior portions of the insula and temporal pole (BA 38). A second cluster encompassed the length of the fusiform gyrus (BA 20/36), extending to include

middle temporal gyrus (BA 21) and inferior temporal gyrus (BA 20). A smaller cluster was found in superior frontal gyrus (BA 8/9). In the right hemisphere, there were significant activations in inferior frontal gyrus (BA 47) and cerebellum.

3.2. Regional specificity of word and picture processing

To examine regional specificity and/or overlap of word vs picture processing, we conducted three analyses. First, we entered, as a conjunction, the semantic categories against baseline contrast for both word-based tasks (1 & 2) and then *inclusively* masked it (at $p = .05$, uncorrected for multiple comparisons) with the semantic categories against baseline contrast for the pictures tasks (3 & 4). This provides a more conservative estimate of activation overlap across modalities and a more appropriate procedure for statistical comparison with modality-specific activations than the conjunction of all four studies described in (1). In reporting this masked analysis, the z values and peak coordinates refer to the results for the masked contrast only (words vs baseline) (see Fig. 1).

Second, to identify word-specific semantic activations, we entered the word-based contrasts as before, but *exclusively* masked them with the picture-based semantic activations (at $p = .05$ uncorrected). Third, we reversed this procedure to produce picture-specific semantic activations. The results are presented in Table 2 and Fig. 2. The top panel of Fig. 2 presents semantic activations common to both input modalities (green), word-specific semantic activations (blue) and picture-specific semantic activations (yellow) superimposed on inferior axial brain slices. The bottom panel presents the

Table 1
A conjunction of the four tasks showing brain areas of activity for the contrast of semantics minus baseline

Regions	Cluster extent	Voxel level		Coordinates		
		$P_{corrected}$	Z	x	y	z
Left Hemisphere						
Inferior frontal gyrus (BA 47)	3827	0.000	>8.0	−38	30	−12
Inferior frontal gyrus (BA 47)		0.000	6.24	−48	26	8
Superior temporal gyrus (BA 38)		0.000	5.90	−30	20	−32
Fusiform gyrus (BA 36/37)	2240	0.000	7.29	−32	−38	−20
Fusiform gyrus (BA 36)		0.000	7.16	−34	−14	−30
Parahippocampal gyrus (BA 36)		0.000	6.99	−30	−24	−26
Superior frontal gyrus (BA 8)	547	0.001	5.67	−14	40	50
Superior frontal gyrus (BA 9)		0.007	5.20	−8	48	44
Right Hemisphere						
Cerebellum	500	0.008	5.18	28	−88	−32
Cerebellum		0.021	4.97	22	−82	−34
Cerebellum		0.023	4.96	32	−82	−36
Inferior frontal gyrus (BA 47)	442	0.014	5.06	36	24	−12

All voxels significant at $p < .178$ ($Z > 5.0$) after statistical correction are reported. Cluster extents are presented at an uncorrected threshold of .001. Multiple peaks within an extent are shown on subsequent lines.

A Lexical Decision

Tools

| bucket | helfkt | pliers | trrupe |

Baseline

| pnjjyx | bjhdzpt | wjkxl | lksx |

B Domain specificity - Words   **SAME**   **DIFFERENT**

Artefacts

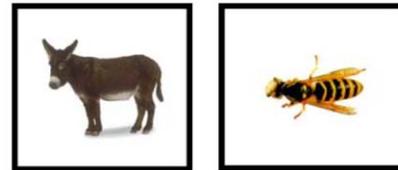| ship | yacht | raft | CANOE | PRAM |

Baseline

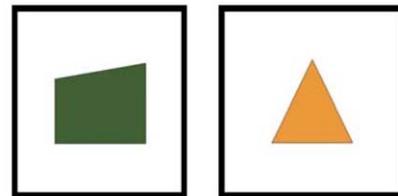| ddd | dddddd | ddddd | DDDD | RRRR |

Semantic categorisation I & II - Pictures   **SAME**   **DIFFERENT**

Animals

Baseline: Semantic categorization I
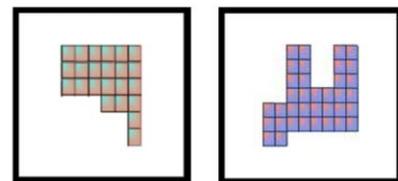
Baseline: Semantic categorization II

Fig. 1. Examples of the experimental stimuli.

location of peak activations for the significant clusters in each contrast, together with rCBF plots of semantics minus baseline activations for the four tasks.

3.2.1. Common word and picture activations

There was a large, common cluster of activation encompassing the left inferior frontal gyrus (BA 47), which extended to include the most anterior portion of the superior temporal gyrus (BA 38). A second large cluster involved anterior and medial portions of the fusiform gyrus (BA 36/20/37) and the parahippocampal gyrus. Finally, a small region of activation was found in left middle temporal gyrus (BA 21), although this was not significant after correcting for multiple comparisons. See Fig. 2.

3.2.2. Word-specific activations

These were confined to two large clusters. The first included the left anterior and superior portions of the superior temporal gyrus (BA 38) and the insula, extending posteriorly to middle temporal gyrus (BA 21) and medially to the parahippocampal gyrus (BA 35/34) and the hippocampus. A second cluster extended from the right temporal pole to the right amygdala and anterior fusiform gyrus (BA 36).

3.2.3. Picture-specific activations

Two large picture-specific activation clusters were located in more posterior regions. One extended from the right inferior occipital gyrus (BA 19) to encompass a large portion of the right fusiform gyrus (BA 18/19/37).

Table 2
Brain areas of activity for the contrast of semantics minus baseline: (i) Words-related activations inclusively masked with pictures-related activations (at $p = .05$); (ii) Words-related activations exclusively masked with pictures-related activations (at $p = .05$); (iii) Pictures-related activations exclusively masked with words-related activations (at $p = .05$)

Regions	Cluster extent	Voxel level		Coordinates		
		$P_{corrected}$	Z	x	y	z
(i) Common activations (all left hemisphere)						
Inferior frontal gyrus (BA 47)	1970	0.001	6.85	−34	30	−10
Superior temporal gyrus (BA 38)		0.001	5.97	−30	18	−26
Fusiform gyrus (BA 36)	806	0.001	6.35	−32	−10	−30
Parahippocampal gyrus (BA 36)		0.001	5.63	−30	−32	−22
(ii) Word-specific activations						
Left hemisphere						
Superior temporal gyrus (BA 38)	1127	0.002	5.33	−30	10	−22
Parahippocampal gyrus (BA 36)		0.005	5.13	−22	−34	−20
Right hemisphere						
Superior temporal gyrus (BA 38)	852	0.010	5.01	38	16	−28
(iii) Picture-specific activations						
Left hemisphere						
Fusiform gyrus (BA 37)	4332	0.001	7.24	−34	−50	−20
Lingual gyrus (BA 18)		0.001	6.51	−16	−78	−10
Fusiform gyrus (BA 19)		0.001	6.35	−38	−78	−14
Right hemisphere						
Inferior occipital gyrus (BA 19)	2863	0.001	7.58	42	−78	−14
Fusiform gyrus (BA 37)		0.001	7.58	40	−52	−20
Cerebellum		0.001	5.40	44	−78	−30

All voxels significant at $p < .032$ ($Z = 5.0$) after statistical correction are reported. Cluster extents are presented at an uncorrected threshold of .001. Multiple peaks within an extent are shown on subsequent lines.

A second, larger cluster was found in analogous regions of the left hemisphere, additionally involving the lingual gyrus (BA 18/19) and precuneus (BA 31).

We conducted a second analysis of picture-specific semantic activations, increasing the mask threshold for words from .05 to .01. This removed those voxels that did not reach the more stringent level of significance in the masking contrast, thus reducing its effect on the pictures-related activations. We found more extensive activation in occipitotemporal regions, but further recruitment of anterior temporal regions was not observed. Thus, the more extensive anterior temporal activations observed for words relative to pictures, did not appear to relate to differences in the statistical power of the tasks.

4. Discussion

In this study we found that both words and pictures robustly activated a common region of the left fusiform gyrus (BA 36, anterior BA 37), left inferior frontal gyrus (BA 47) and the most anterior aspect of the left temporal pole (BA 38) during semantic task conditions relative to a low-level baseline. A conjunction of the four tasks also produced right hemisphere activation of inferior frontal gyrus (BA 47) and cerebellum, although when we inclusively masked words with pictures, no right-sided activation was observed. There was regionally extensive recruitment of anterior temporal lobes during semantic judgements of words but not pictures. Picture-specific activations were primarily restricted to occipital and posterior temporal areas, bilaterally, including inferior occipital gyrus (BA 19), fusiform gyrus (BA 19/37) and lingual gyrus (BA 11).

4.1. Common effects for words and pictures

There is increasing evidence for successive object processing stages that occur from early, retinoptic areas into anterior aspects of the inferior temporal cortex (Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999; Grill-Spector et al., 1998; Tyler et al., 2004). Our findings, while consistent with this position, also indicate that anterior fusiform involvement in semantic processing is not differentiated by form of input (words or pictures). Several earlier reports have distinguished anterior from posterior regions of the fusiform gyrus during the processing of meaningful over non-meaningful stimuli, irrespective of input modality (Demonet et al., 1992; Moore & Price, 1999). However, the peak fusiform activations in the present study were mostly
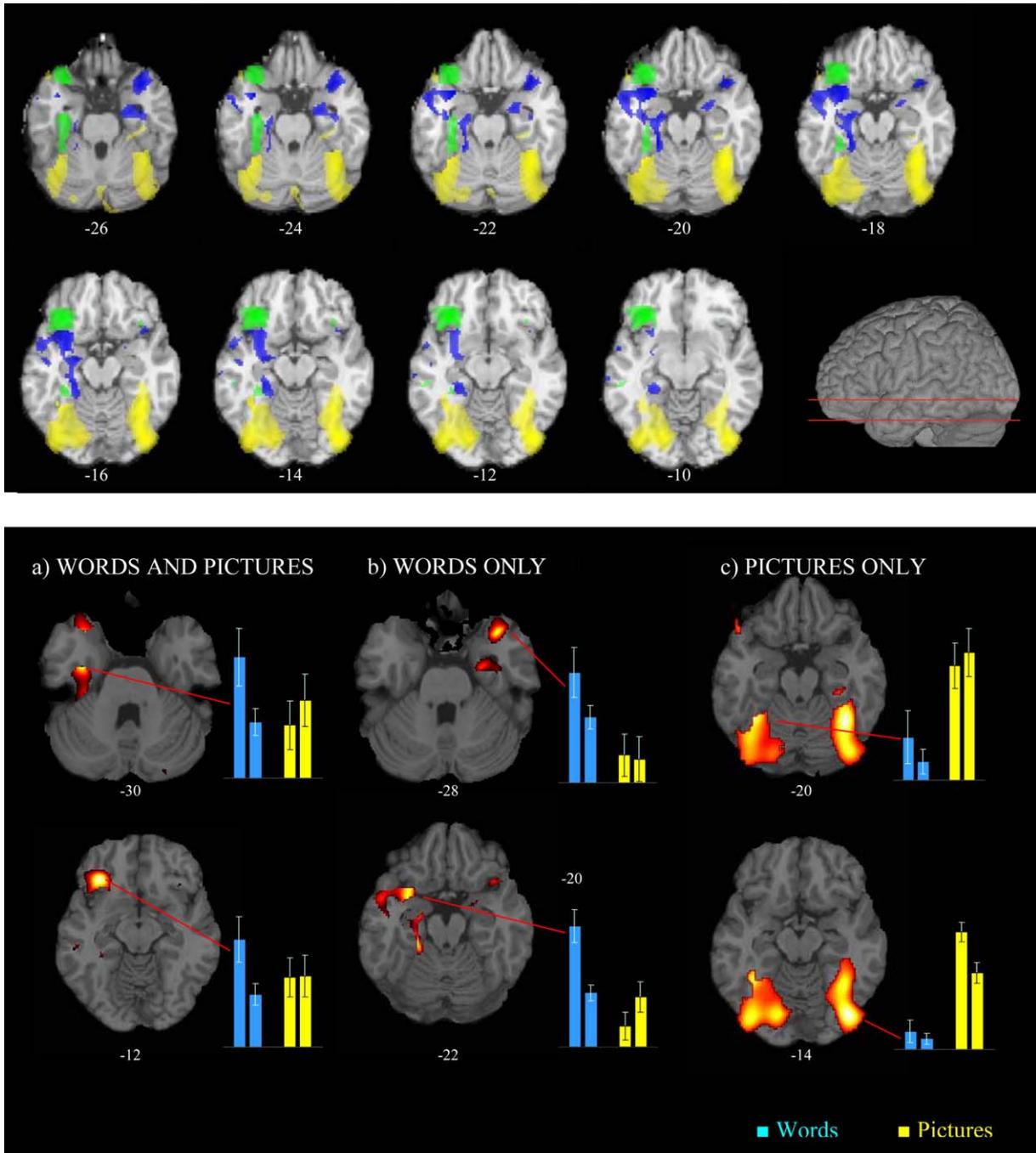
Fig. 2. Modality-specific and modality-independent activations for the semantics minus baseline contrasts, using inclusive and exclusive masking (at $p = .05$). Top panel shows activations superimposed on axial inferior brain sections, presented in MNI space. Green, activations common to words and pictures; blue, word-specific activations; and yellow, pictures-specific activations. Bottom panel presents slices showing areas of maximal activation for each analysis, together with peak rCBF comparison charts for the two modalities. All cluster extents are presented at an uncorrected threshold of .001. Images use neurological convention (L = L).

anterior to those previously reported. Activation in posterior regions (BA 19) appears not to differentiate meaningful from non-meaningful stimuli (consistent with a pre-conceptual level of representation), whereas anterior and medial regions (BA 37/20/36) become significantly more active during the processing of meaningful stimuli, irrespective of task (e.g., Bar et al., 2001;

Moore & Price, 1999). The present results identify left anterior fusiform gyrus (BA 20 and anterior BA 37), left anterior medial temporal cortex (BA 35/36) and the temporal poles (BA 38) as critically involved in the representation of conceptual knowledge.

The importance of anterior medial temporal cortex in object identification has recently been shown in an fMRI

study where participants named pictures of common objects at a domain (i.e., living or non-living) or basic level (i.e., identify the object itself as, for example, a monkey). Domain-level naming produced activation of posterior regions of the inferior temporal cortex while basic-level naming produced additional anterior recruitment of left perirhinal and entorhinal cortices (Tyler et al., 2004). The involvement of the temporal poles could not be evaluated because of susceptibility artefacts in fMRI (Devlin et al., 2002b). We argue that when an object is to be identified at a high level of specificity, it requires extraction of more detailed information in order to differentiate between similar objects and these additional processes engage anterior medial regions of temporal cortex. The meta-analysis results are consistent with the existence of a hierarchical object processing structure, with more anterior temporal regions necessary for representation of objects at high levels of specificity and posterior regions more associated with intermediate and pre-semantic stages of object representation. Rather than reflecting modality-specific semantic representations, both verbal and visual input routes appear to converge on these anterior temporal sites, consistent with their being part of the verbal basis for a unitary semantic system (although see below for consideration of additional word-specific activation of the temporal poles).

Recently, Simmons and Barsalou (2003) have proposed a model, based on Damasio's (1989) convergence zone (CZ) theory, which describes feature-binding neurons arranged hierarchically from posterior to anterior brain regions. Neurons in more posterior CZs proximal to a particular modality-specific area (e.g., visual processing or motor processing areas) capture the content held there. In more anterior regions, patterns of activation *across* modalities are indexed. The authors stress Damasio's assertion that a CZ points back to the sensori-motor content held in feature maps but does not store or represent that information (i.e., it indexes information held elsewhere). However, although this proposal implies a common mechanism for capturing modality-specific content, in a subsequent paper, the authors suggest that some conjunctive neurons may have modality-specific 'tunings' (Barsalou, Simmons, Barbey, & Wilson, 2003). Our data offer limited support for this proposal. If correct, there is likely to be quite extensive overlap in regions of activation for words and pictures in semantic processing areas, with differential activations mostly restricted to the feature-map sites in the posterior regions. We found that the areas of overlap were primarily restricted to anterior fusiform gyrus (37/20) and parahippocampal gyrus (BA 36), with more widespread anterior temporal activity found only for words and extensive bilateral occipitotemporal recruitment specific to pictures.

There is an increasing body of evidence suggesting that the types of features that define and distinguish different objects and object categories may reflect the cortical organisation of conceptual representations in the brain (e.g., Martin, 2001; Martin et al., 2000; Tyler & Moss, 2001), and that the regions activated by an object processing task relate to the *type* or *level* of semantic processing required, rather than the intrinsic properties of the presented objects (Tyler et al., 2004). Nevertheless, although based on only one primary type of task (semantic categorisation), our results suggest that the anterior fusiform gyrus involvement in representation of information at a semantic or conceptual level is not differentially affected by the mode of input (words or pictures).

In addition to temporal regions, we found a large cluster of common activation encompassing the entire extent of the inferior frontal gyrus (BA 45/47), consistent with previous reports suggesting its general role in semantic processing (e.g., Vandenberghe et al., 1996; Wagner et al., 1997). Various theories have been suggested to account for activation of LIFG during semantic relative to non-semantic tasks: for example, increased competition among candidate responses (Thompson-Schill, 2003), semantic 'working memory' or central executive system (e.g., Gabrieli et al., 1996; Wagner et al., 1997) or complex syntactic operations (Carpentier et al., 2001). The present data preclude a straightforward interpretation of the role of LIFG in semantic tasks, but suggest that it may be involved in tasks that require effortful, explicit processing, or the active holding in memory of information required for task success since these are processes common to all word and picture semantic judgment tasks.

However, our interpretation of the data can be challenged on at least two counts. First is the possibility that both semantic categorisation tasks using words *and* pictures present demands that are essentially word based. Whether the stimuli employed are words or pictures, the chosen strategy may be to generate a category name for the cues (e.g., animals) and then to rate whether the category to which the target belongs matches it. In this case, a word-based strategy would be chosen for both stimulus types and the common fusiform activation may reflect word-specific representations. We cannot discount this interpretation strictly on the basis of our own results, but in the context of evidence associating the temporal poles (BA 38) with word retrieval, we argue that the area of common activation (BA 20/36/37) is unlikely simply to reflect naming or word retrieval. It is unlikely that *any* task of visual object processing at a semantic level can be conceived which rules out the possibility of implicit naming or word retrieval. Nevertheless, we accept that our findings are based upon a particular type of task and that further confirmatory evidence, based on a wider range of tasks and contexts may allow us to apply our results to the semantic system in general with a greater level of confidence.

Second, with respect to Paivio's dual code theory (1971), it is possible that both words and pictures activate both visual and verbal codes. For example, the frontal activation may relate to the processing of verbal codes and the fusiform activation might be associated with the processing of non-verbal codes. Thus, a common set of dual codes are activated in both types of task (for words, the visual codes are automatically triggered by the verbal codes, and the opposite pattern would take place for picture presentations). It follows that the results might better reflect a 'dual-code' process, with verbal and non-verbal representation reflected by LIFG and fusiform gyrus, respectively. However, while the nature of frontal lobe involvement in semantic tasks remains controversial, we argue that there is very little (if any) neuropsychological evidence which indicates that semantic-level representations are stored within the LIFG. Instead, we suggest that it is the anterior fusiform gyrus that is the most plausible region for holding representations, which are accessed both by pictures and words.

Finally, we should point out that, even if either of these two interpretations were correct, the observation that processing converges on this antero-medial left temporal region indicates that the neural correlates of semantic representations engaged in our tasks are not dependent upon the format in which a stimulus is presented.

4.2. Word-specific effects

We found more extensive anterior temporal activation for words relative to pictures which involved a broad region of bilateral temporal poles (BA 38). Damasio (1992) has reported that patients with primarily left temporal pole damage have a difficulty with proper noun retrieval. PET studies have associated the temporal poles, either unilaterally or bilaterally, with face recognition (e.g., Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996) and general aspects of comprehension (Maguire, Frith, & Morris, 1999). Semantic dementia patients (in which damage to the temporal poles is a hallmark feature) show preserved ability to retrieve general information about persons or objects, but are unable to uniquely identify those objects (e.g., Hodges, Graham, & Patterson, 1995). Thus, neuroimaging and lesion studies suggest that anterior temporal regions may be involved in processing detailed aspects of object attributes. Consistent with this position, our data suggest that the temporal poles are recruited when fine-grained discrimination among similar objects is required, but not when discriminating among semantically meaningless stimuli.

Why is the extensive recruitment of the temporal poles word-specific? Semantic categorisation of pictures requires category-level knowledge. Participants can perform this task on the basis of overall similarity between objects and do not need to access the unique properties that differentiate one object from another. In contrast, in the tasks using words subjects were given basic-level object labels, which access unique semantic representations. If, as Grabowski et al. (2001) have argued, the left temporal pole is more engaged in naming tasks as the level of specificity of word retrieval increases, this may explain the more widespread left hemisphere word-specific activations found in the present study. We also found word-specific activation in the right temporal pole, consistent with the view that semantic system is bilaterally represented (Beeman, Bowden, & Gernsbacher, 2000; Perani et al., 1999).

In summary, although the temporal poles activation was word-specific, we propose that these regions may be involved in semantic representation in both modalities but that their engagement might be primarily determined by the level of processing required for the task. However, this is a provisional interpretation and further work is needed to explore the extent to which the additional temporal pole activation for words reflects either material-specific semantic representations or differences in the task processes which operate in these regions.

4.3. Picture-specific effects

Picture-specific recruitment was restricted to more posterior regions, including inferior occipital gyrus (BA 19), lingual gyrus (BA 18) and fusiform gyrus (BA 37). Bilateral occipitotemporal cortex activation is routinely found when objects are compared to non-objects (Bly & Kosslyn, 1997) and when objects are compared to words (Moore & Price, 1999). Although the nature of the representations along the extent of the fusiform gyrus remains controversial, neuroimaging, and event-related potential studies of visual object recognition have suggested that the more posterior regions (BA 19) are associated with prerecognition analysis such as feature extraction and intermediate shape processing (e.g., Bar et al., 2001). Anterior fusiform gyrus (BA 37), in contrast, appears to mediate explicit identification (Bar et al., 2001; Buchel et al., 1998). The extensive picture-specific activation found in the present study involved primarily posterior occipito-temporal regions, but did extend to involve more posterior aspects of BA 37 (lateral occipital complex). In the left hemisphere, this modality-specific recruitment was situated adjacent to an area of anterior fusiform gyrus shared by both input modalities. Buchel et al. (1998) have argued that BA 37 is multimodal such that its response properties should not be altered by the modality of input. Our data suggest that, while this proposal may be correct for more anterior aspects of the fusiform gyrus, there may be important functional differences throughout the

posterior–anterior extent of this region, with modality-specificity (pictures) recruitment of posterior BA 37. Anterior regions of the lateral occipital cortex (posterior and mid fusiform gyrus) are activated more strongly by whole, intact objects than by scrambled object stimuli. However, the region appears not to distinguish familiar from novel shapes (Malach et al., 1995), suggesting that the picture-specific activations in fusiform gyrus found in the present study reflect an intermediate or pre-semantic stage of visual processing.

4.4. Conclusions

In this study, we found evidence for a critical role of the anterior extent of the fusiform gyrus (BA 37/20) in the representation of conceptual knowledge. This region does not appear to be modulated by the modality of visual input (pictures or words), suggesting that it holds unitary semantic representations formed via converging inputs from more posterior areas. We also found that the left parahippocampal and perirhinal cortex are recruited when a semantic level of representation is required (irrespective of modality), consistent with a role for this region in the integration of sensory information into semantically meaningful polymodal feature combinations. Finally, we found common recruitment of the left inferior frontal gyrus, which we suggest reflects an executive or working memory role rather than conceptual representation per se.

Our findings are concerned with the semantic processing of two types of stimulus (words and pictures). Further work is needed in order to extend our knowledge of whether semantic processing across different primary sensory modalities (visual, auditory, and tactile) or different types of task also converge on the same regions associated with conceptual representations. A number of recent studies (PET and fMRI) have indicated significant gender effects during lexical semantic and naming tasks (e.g., Grabowski, Damasio, Tranel, & Eichhorn, 2003; Rossell, Bullmore, Williams, & David, 2002), although these have usually reported such differences to be smaller than the task effects. However, the relevance of gender in language-related activations should be considered in future studies.

In summary, our data are most consistent with those accounts which propose a single, common system of semantic representations (Caramazza et al., 1990; Tyler & Moss, 2001), rather than with the view that there are distinct conceptual representations for the verbal and visual modalities (Beauvois, 1982; Paivio, 1971, 1991; Warrington & Shallice, 1984). However, we also found word-specific activations in anterior temporal cortex and picture-specific activations in occipitotemporal cortex. The posterior recruitment (including posterior BA 37/19), associated with pictures most likely relates to intermediate, non-semantic levels of representation. Our

data suggests that the temporal poles are an important part of a distributed system subserving semantic representations, but that their involvement may depend upon the level of specificity at which an object is represented.

References

Ashburner, J., & Friston, K. (1997). Multimodal image coregistration and partitioning—a unified framework. *NeuroImage, 6*, 209–217.

Baayen, R., & Pipenbrook, R. (1995). *The Celex Lexical Database. Philadelphia Linguistic Data Consortium*. Philadelphia: University of Pennsylvania.

Bar, M., Tootell, R. B. H., Schacter, D. L., Greve, D. N., Fischl, B., Mendola, J. D., Rosen, B. R., & Dale, A. M. (2001). Cortical Mechanisms Specific to Explicit Visual Object Recognition. *Neuron, 29*, 529–535.

Barsalou, L. W., Simmons, W. K., Barbey, A., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences, 7*, 84–91.

Beauvois, M. F. (1982). Optic aphasia: A process of interaction between vision and language. *Philosophical Transactions of the Royal Society of London. Series B. Biological Sciences, 298*(1089), 35–47.

Beeman, M. J., Bowden, E. M., & Gernsbacher, M. A. (2000). Right and left hemisphere cooperation for drawing predictive and coherence inferences during normal story comprehension. *Brain and Language, 71*, 310–336.

Bly, B. M., & Kosslyn, S. M. (1997). Functional anatomy of object recognition in humans: Evidence from positron emission tomography and functional magnetic resonance imaging. *Current Opinion in Neurology, 10*(1), 5–9.

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience, 25*, 151–188.

Buchel, C., Price, C., & Friston, K. (1998). A multimodal language region in the ventral visual pathway. *Nature, 394*(6690), 274–277.

Caramazza, A. (2000). Minding the facts: A comment on Thompson-Schill et al.'s "A neural basis for category and modality specificity of semantic knowledge". *Neuropsychologia, 38*(7), 944–949.

Caramazza, A., Hillis, A. E., Rapp, B. C., & Romani, C. (1990). The multiple semantics hypothesis: Multiple confusions? *Cognitive Neuropsychology, 7*, 161–189.

Carpentier, A., Pugh, K. R., Westerveld, M., Studholme, C., Sknnja, O., Thompson, J. L., Spencer, D. D., & Constable, R. I. (2001). Functional MRI of language processing: Dependence on input modality and temporal lobe epilepsy. *Epilepsia, 42*(10), 1241–1254.

Chaminade, T., & Decety, J. (2002). Leader or follower? Involvement of the inferior parietal lobule in agency. *Neuroreport, 13*, 1975–1978.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience, 2*, 913–919.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., & Michel, F. (2000). The visual word form area. Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123*, 291–307.

Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain, 125*(5), 1054–1069.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33A*, 497–505.

Coslett, H. B., & Saffran, E. M. (1992). Optic aphasia and the right hemisphere—a replication and extension. *Brain and Language, 43*, 148–161.

Damasio, A. R. (1989). The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation, 1*, 123–132.

Damasio, A. R. (1992). Aphasia. *New England Journal of Medicine, 326*, 531–539.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature, 380*(6574), 499–505.

Dehaene, S., Le Clec, H. G., Poline, J. B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: A prelexical representation of visual words in the fusiform gyrus. *Neuroreport, 13*(3), 321–332.

D'Esposito, M., Detre, J. A., Aguirre, G. K., Stallcup, M., Alsop, D. C., Tippet, L. J., & Farah, M. J. (1997). A functional MRI study of mental image generation. *Neuropsychologia, 35*, 725–730.

Demb, J. B., Desmond, J. E., Wagner, A. D., Vaidya, C. J., Glover, G. H., & Gabrieli, J. D. (1995). Semantic encoding and retrieval in the left inferior prefrontal cortex: A functional MRI study of task difficulty and process specificity. *Journal of Neuroscience, 15*(9), 5870–5878.

Demonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., Rascol, A., & Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain, 115*(Pt 6), 1753–1768.

Demonet, J. F., Price, C., Wise, R., & Frackowiak, R. S. J. (1994). A PET study of cognitive strategies in normal subjects during language tasks. Influence of phonetic ambiguity and sequence processing on phoneme monitoring. *Brain, 117*(4), 671–682.

Devlin, J. T., Moore, C. J., Mummery, C. J., Gorno-Tempini, M. L., Phillips, J. A., Noppeney, U., Frackowiak, R. S., Friston, K. J., & Price, C. J. (2002a). Anatomic constraints on cognitive theories of category specificity. *Neuroimage, 15*(3), 675–685.

Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Moss, H. E., Fadili, M. J., & Tyler, L. K. (2002b). Is there an anatomical basis for category-specificity? Semantic memory studies in PET and fMRI. *Neuropsychologia, 40*(1), 54–75.

Dolan, R. J., Morris, J. S., & de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proceedings of the National Academy of Sciences of the United States of America, 98*, 10006–10010.

Farah, M. J., Hammond, K. M., Mehta, Z., & Ratcliff, G. (1989). Category-specificity and modality-specificity in semantic memory. *Neuropsychologia, 27*(2), 193–200.

Feinberg, T. E., Dyckes-Berke, D., Miner, C. R., & Roane, D. M. (1995). Knowledge, implicit knowledge and metaknowledge in visual agnosia and pure alexia. *Brain, 118*(Pt 3), 789–800.

Ferreira, C. T., Giusiano, B., Ceccaldi, M., & Poncet, M. (1997). Optic aphasia: Evidence of the contribution of different neural systems to object and action naming. *Cortex, 33*(3), 499–513.

Fiez, J. A., & Petersen, S. E. (1998). Neuroimaging studies of word reading. *Proceedings of the National Academy of Sciences of the United States of America, 95*, 853–860.

Forster, K. I., & Forster, J. C. (1991). *The DMASTR display system for mental chronometry*. Tucson: University of Arizona.

Friston, K. J., Holmes, A. P., Worsley, K. J., & Poline, J. P. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping, 2*(4), 189–210.

Friston, K. J., Holmes, A. P., Price, C. J., Buchel, C., & Worsley, K. J. (1999). Multisubject fMRI studies and conjunction analyses. *NeuroImage, 10*, 385–396.

Gabrieli, J. D. E., Desmond, J. E., Demb, J. B., Wagner, A. D., Stone, M. V., Vaidya, C. J., & Glover, G. H. (1996). Functional Magnetic Resonance Imaging of Semantic Memory Processes in the Frontal Lobes. *Psychological Science, 7*(5), 278–283.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform 'face area' increases with expertise in recognising novel objects. *Nature Neuroscience, 2*, 568–573.

Grabowski, T. J., Damasio, H., Tranel, D., & Eichhorn, G. E. (2003). Effects of gender on blood flow correlates of naming concrete entities. *NeuroImage, 20*, 940–954.

Grabowski, T. J., Damasio, H., Tranel, D., Ponto, L. L., Hichwa, R. D., & Damasio, A. R. (2001). A role for left temporal pole in the retrieval of words for unique entities. *Human Brain Mapping, 13*, 199–212.

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research, 41*, 1409–1422.

Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). A sequence of object processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping, 6*, 316–328.

Hillis, A. E., & Caramazza, A. (1995). Cognitive and neural mechanisms underlying visual and semantic processing: Implications from "optic aphasia". *Journal of Cognitive Neuropsychology, 7*, 457–478.

Hodges, J. R., Graham, N., & Patterson, K. (1995). Charting the progression in semantic dementia: Implications for the organisation of semantic memory. *Memory, 3*(3–4), 463–495.

Kanwisher, N., Chun, M. M., McDermott, J., & Ledden, P. J. (1996). Functional imagining of human visual recognition. *Cognitive Brain Research, 5*(1–2), 55–67.

Kinahan, P. E., & Rogers, J. G. (1989). Analytic 3D image reconstruction using all detected events. *IEEE Transactions on Nuclear Science, 35*, 680–684.

Lambon Ralph, M. A., Graham, K. S., Patterson, K., & Hodges, J. R. (1999). Is a picture worth a thousand words? Evidence from concept definitions by patients with semantic dementia. *Brain and Language, 70*, 309–335.

Lambon Ralph, M. A., & Howard, D. (2000). Gogi aphasia or semantic dementia? Simulating and assessing poor verbal comprehension in a case of progressive fluent aphasia. *Cognitive Neuropsychology, 17*, 437–465.

Lee, A. C. H., Robbins, T. W., Graham, K. S., & Owen, A. M. (2002). "Pray or Prey?" Dissociation of semantic memory retrieval from episodic memory processes using photon emission tomography and a novel homophone task. *NeuroImage, 16*, 724–735.

Luzzatti, C., Rumiati, R. I., & Ghirardi, G. (1998). A functional model of visuo-verbal disconnection and the neuroanatomical constraints of optic aphasia. *Neurocase, 4*, 71–87.

McCarthy, R. A., & Warrington, E. K. (1986). Visual associative agnosia: A clinico-anatomical study of a single case. *Journal of Neurology, Neurosurgery, and Psychiatry, 49*, 1233–1240.

McCarthy, R. A., & Warrington, E. K. (1988). Evidence for modality-specific meaning systems in the brain. *Nature, 334*, 428–430.

Maguire, E. A., Frith, C. D., & Morris, R. G. M. (1999). The functional neuroanatomy of comprehension and memory: The importance of prior knowledge. *Brain, 122*(10), 1839–1850.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., & Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America, 92*, 8135–8139.

Martin, A. (2001). Functional neuroimaging of semantic memory. In R. Cabeza & A. Kingstone (Eds.), *Handbook of functional neuroimaging of cognition* (pp. 153–186). Cambridge: MIT Press.

Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science, 270*, 102–105.

Martin, A., Ungerleider, L. G., & Haxby, J. V. (2000). Category-specificity and the brain: The sensory-motor model of semantic representations of objects. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (2nd ed., pp. 1023–1036). Cambridge: MIT Press.

Moore, C. J., & Price, C. J. (1999). Three distinct ventral occipito-temporal regions for reading and object naming. *Neuroimage, 10*, 181–192.

Morcom, A. M., Good, C. D., Frackowiak, R. S. J., & Rugg, M. D. (2003). Age effects on the neural correlates of successful memory encoding. *Brain, 126*, 213–229.

Noppeney, U., & Price, C. J. (2002). Retrieval of visual, auditory, and abstract semantics. *Neuroimage, 15*(4), 917–926.

Paivio, A. (1971). *Imagery and verbal processes*. London: Holt, Rinehart & Winston.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press: Oxford.

Paivio, A. (1991). Dual Coding Theory: Retrospect and current status. *Canadian Journal of Psychology, 45*, 255–287.

Perani, D., Cappa, S. F., Schnur, T., Tettamanti, M., Collina, S., Rosa, M. M., & Fazio, F. (1999). The neural correlates of verb and noun processing—a PET study. *Brain, 122*, 2337–2344.

Plaut, D. C. (2002). Graded modality-specific specialisation in semantics: A computational account of optic aphasia. *Cognitive Neuropsychology, 19*, 603–639.

Pochon, J. B., Levy, R., Fossati, P., Lehericy, S., Poline, J. B., Pillon, B., Le Bihan, D., & Dubois, B. (2002). The neural system that bridges reward and cognition in humans: An fMRI study. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 5669–5674.

Price, C. J., & Friston, K. J. (1997). Cognitive conjunction: A new approach to brain activation experiments. *Neuroimage, 5*(4/Pt1), 261–270.

Price, C. J., Moore, C. J., Humphreys, G. W., Frackowiak, R. S. J., & Friston, K. J. (1996). The neural regions sustaining object recognition and naming. *Proceedings of the Royal Society of London. Series B., 263*, 1501–1507.

Rapp, B. C., Hillis, A. E., & Caramazza, A. (1993). The role of representations in cognitive theory: More on multiple semantics and the agnosias. *Cognitive Neuropsychology, 10*, 235–249.

Riddoch, M. J., & Humphreys, G. W. (1987). Visual object processing in optic aphasia: A case of semantic access agnosia. *Cognitive Neuropsychology, 4*, 131–185.

Riddoch, M. J., Humphreys, G. W., Coltheart, M., & Funnell, E. (1988). Semantic systems or system. Neuropsychological evidence re-examined. *Cognitive Neuropsychology, 5*, 3–25.

Rossell, S. L., Bullmore, E. T., Williams, S. C. R., & David, A. S. (2002). Sex differences in functional brain activation during a lexical visual field task. *Brain and Language, 80*, 97–105.

Saffran, E. M., Coslett, H. B., Martin, N., & Boronat, C. B. (2003). Access to knowledge from pictures but not words in a patient with progressive fluent aphasia. *Language and Cognitive Processes, 18*, 725–757.

Shallice, T. (1988). Specialisation within the semantic system. *Cognitive Neuropsychology, 5*, 133–142.

Shallice, T. (1993). Multiple semantics: Whose confusions? *Cognitive Neuropsychology, 10*, 251–261.

Silbersweig, D. A., Stern, E., Frith, C. D., & Cahill, C. (1993). Detection of thirty-second cognitive activations in single subjects with positron emission tomography: A new low-dose H2(15)O regional cerebral blood flow three-dimensional imaging technique. *Journal of Cerebral Blood Flow and Metabolism, 13*(4), 617.

Simmons, K., & Barsalou, L. W. (2003). The similarity-in-topography principle: Reconciling theories of conceptual deficts. *Cognitive Neuropsychology, 20*, 451–486.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. Stuttgart: Georg Thieme Verlag.

Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: Inferring "how" from "where". *Neuropsychologia, 41*, 280–292.

Thompson-Schill, S. L., Aguirre, G. K., D'Esposito, M., & Farah, M. J. (1999). A neural basis for category and modality specificity of semantic knowledge. *Neuropsychologia, 37*(6), 671–676.

Tyler, L. K., Bright, P., Dick, E., Tavares, P., Pilgrim, L., Fletcher, P. C., Greer, M., & Moss, H. E. (2003). Do semantic categories activate distinct cortical regions. PET studies of picture processing. *Cognitive Neuropsychology, 20*, 541–559.

Tyler, L. K., & Moss, H. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Science, 5*, 244–252.

Tyler, L. K., Russell, R., Fadili, J., & Moss, H. E. (2001). The neural representation of nouns and verbs: PET studies. *Brain, 124*, 1619–1634.

Tyler, L. K., Stamatakis, E. A., Bright, P., Acres, K., Abdallah, S., Rodd, J. M., & Moss, H. E. (2004). Processing objects at different levels of specificity. *Journal of Cognitive Neuroscience, 16*, 1–12.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature, 383*, 254–256.

Wagner, A. D., Desmond, J. E., Demb, J. B., Glover, G. H., & Gabrieli, J. D. E. (1997). Semantic repetition priming for verbal and pictorial knowledge: A functional MRI study of left inferior prefrontal cortex. *Journal of Cognitive Neuroscience, 9*, 714–726.

Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology, 27*, 635–657.

Warrington, E. K., & McCarthy, R. A. (1994). Multiple meaning systems in the brain: A case for visual semantics. *Neuropsychologia, 32*(12), 1465–1473.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107*(Pt 3), 829–854.