BRAIN
A JOURNAL OF NEUROLOGY

Crossmodal integration of object features: Voxel-based correlations in brain-damaged patients

Kirsten I. Taylor, Emmanuel A. Stamatakis* and Lorraine K. Tyler

Centre for Speech, Language and the Brain, Department of Experimental Psychology, University of Cambridge, UK

*Present address: Division of Anaesthesia, School of Clinical Medicine, University of Cambridge, UK

Correspondence to: Kirsten I. Taylor,
Centre for Speech, Language and the Brain,
Downing Street, Cambridge CB2 3EB, UK
E-mail: ktaylor@csl.psychol.cam.ac.uk

How does the brain bind together the different sensory features of objects to form meaningful, multimodal object representations? Human functional imaging findings implicate the left posterior superior temporal sulcus/middle temporal gyrus (pSTS/MTG) in crossmodal integration, while animal ablation findings support a hierarchical object processing model in which outputs from each sensory stream are integrated in perirhinal cortex (PRc) of the anteromedial temporal lobe. To determine which neural regions are necessary for integrating audiovisual object features, and which regions are necessary for understanding the meaning of crossmodal objects, we administered crossmodal (audio-visual) and unimodal (auditory, visual) integration tasks to 16 brain-damaged patients. We correlated patients' behavioural performance with measures of neural integrity (signal intensity) of each voxel across the brains of each patient. The integrity of bilateral anteromedial and temporopolar regions, but not pSTS/MTG, was significantly correlated with poorer crossmodal compared with unimodal integration performance, and with meaningful aspects of crossmodal integration. Additional analyses confirmed the negative crossmodal integration findings in the pSTS/MTG: performance on a sentence-picture matching control task was significantly correlated with MTG/STG voxel signal intensities, suggesting that a truncated range of signals in this region could not have been responsible for the lack of a significant correlation between integrity and crossmodal integration performance, and individual analyses of three patients with lesions in pSTS/MTG but spared anteromedial temporal cortex revealed equivalent unimodal and crossmodal integration performance. These results extend findings from the non-human primate literature into the human domain by demonstrating that anteromedial temporal cortex is critically involved in crossmodal integration of object features. However, pSTS/MTG appears to play a supportive but non-essential role during crossmodal integration. Taken together, the present findings are consistent with a neurocognitive account of object representations which claims that anteromedial temporal lobe is critically involved in the formation and processing of complex, multimodal object representations.

Introduction

How does the brain bind together different sensory features to form meaningful, multimodal object representations? Human functional magnetic resonance imaging (fMRI) studies comparing activity for unimodal auditory and unimodal visual object features with simultaneously presented auditory and visual object features show increased activity for crossmodal compared with unimodal stimuli ('crossmodal integration effects') in the posterior superior temporal sulcus and middle temporal gyrus (pSTS/MTG) (Calvert, 2001; Beauchamp *et al.*, 2004*a*; Amedi *et al.*, 2005; Beauchamp, 2005). Correspondingly, some neurons in the non-human primate STS respond to both auditory and visual stimulation (Benevento *et al.*, 1977; Bruce *et al.*, 1981; Barraclough *et al.*, 2005), and demonstrate crossmodal integration effects for combined audiovisual stimulation (Barraclough *et al.*, 2005). Several theories have been proposed to account for crossmodal integration effects in pSTS/MTG. According to Beauchamp and colleagues (2004*a*), multimodal object representations are formed when unimodal sensory information reaching discrete unisensory pSTS/MTG sub-regions is integrated in intervening multisensory pSTS/MTG patches. This integration may transform different sensory inputs into a common code to facilitate further processing of the multi-modal object (Amedi *et al.*, 2005; Beauchamp, 2005). Other investigators have claimed that differences in the extent to which receptive fields overlap in individual polysensory neurons (e.g. differences in visual and auditory receptive fields coded for in non-human primate STS; Hikosaka *et al.*, 1988) are exploited to optimize multimodal integration and generate multidirectional sensory predictions (Pouget *et al.*, 2002; Avillac *et al.*, 2005).

An alternative account of multimodal object processing is based on a model of object processing developed in non-human primate research. This hierarchical object processing model claims that increasingly more complex combinations of object features are integrated within hierarchically organized sensory processing streams, and multimodal object representations are formed and processed where these streams converge (Desimone and Ungerleider, 1989; Murray *et al.*, 1998; Murray and Bussey, 1999; Bussey *et al.*, 2005). Perirhinal cortex (PRc), the first site in the ventral visual object processing stream to receive inputs from all sensory modalities via uni- or polymodal association areas (Suzuki and Amaral, 1994), is therefore hypothesized to play a critical role in processing multimodal object representations (Murray *et al.*, 1998; Murray and Richmond, 2001).

Similar models have been developed for the human system, most notably Damasio's convergence zone theory (Damasio, 1989, 1990) and related proposals (e.g. Barsalou, 1999; Simmons and Barsalou, 2003). In general, these models propose that features of objects are coded in the corresponding sensory and motor cortices, and that feature information is fed forward and bound in downstream convergence zones located at multiple hierarchical sites throughout the association cortices. Object perception is thought to require either reactivation of relevant sensory and motor features in the respective sensory-motor cortices (Damasio, 1989) or convergence zones (Simmons and Barsalou, 2003) via feedback connections. Damasio's original model (1989) proposed that while temporal and frontal lobe convergence zones

are involved in binding crossmodal information, crossmodal recognition and recall critically depends on the activation of features coded in early sensory cortices. Simmons and Barsalou's more recent approach suggested, based in part on non-human primate findings reviewed above, that PRc acts as the critical crossmodal convergence zone (Simmons and Barsalou, 2003).

In a recent fMRI study of audiovisual object feature integration in humans, we reconciled these sets of data by finding crossmodal integration effects in both pSTS/MTG and PRc (Taylor *et al.*, 2006). Because the purpose of crossmodal integration is to construct meaningful object representations (do we run from it or hug it?), this experiment also included two factors manipulating the meaning of the crossmodal stimuli: domain and congruency. Thus, stimuli represented either living or nonliving things (domain manipulation), and stimuli were either related in meaning (congruent, e.g. a barking sound with a picture of a dog) or unrelated in meaning (incongruent; e.g. a bird sound with a picture of an elephant). The fMRI analyses showed that only PRc activity was modulated by the meaning of the crossmodal stimuli, with greater responses for living compared with nonliving and with meaningfully unrelated (incongruent) compared with meaningfully related (congruent) crossmodal stimuli. These findings support a neurocognitive account, combining the hierarchical neural model of object processing with a cognitive model of conceptual representations, to explain how the meanings of familiar objects are neurally represented (Taylor *et al.*, 2006). Specifically, a number of cognitive models claim that object meanings are represented in terms of distributed features of different types, where objects differ in terms of the number and type of features (McRae, 1997; Tyler *et al.*, 2000, 2004, 2007; Tyler and Moss, 2001). For example, living things have many features which are shared by other living things (e.g. many animals have legs, eyes), making them more ambiguous. Conversely, nonliving things have relatively more features which are distinctive to the particular object (e.g. few tools have serrated edges) (McRae and Cree, 2002; Cree and McRae, 2003; Randall *et al.*, 2004). As object identification requires distinctive features or feature clusters, the neurocognitive account predicts that living things, because they are ambiguous by virtue of their similarity to each other, require more complex conjunctions of features to be uniquely identified than nonliving things, functions supported by more anterior and anteromedial ventral temporal regions. Thus, the finding of living/nonliving distinction in PRc but not pSTS/MTG supports this neurocognitive account of the neural representation of object meaning (see also Moss *et al.*, 2005), and further suggests that these regions play complementary functional roles during crossmodal feature integration to form meaningful, multimodal object representations, with sensory integration in pSTS/MTG, and semantic integration in PRc (Taylor *et al.*, 2006).

However, the fMRI data leave key questions unanswered: is the structural integrity of both pSTS/MTG and PRc necessary for integrating auditory and visual object features, and is only PRc integrity necessary for processing the meaningfulness of multimodal objects? While there is to our knowledge no evidence that pSTS/MTG damage leads to stable deficits in crossmodal integration performance (Ettlinger and Wilson, 1990), relevant lesion studies on the PRc have only been carried out in animals where

ablations of the PRc and entorhinal cortex do impair crossmodal integration abilities (Murray *et al.*, 1998). Also the critical question of whether the structural integrity of the human PRc is required to process the meaning of multisensory stimuli, i.e. the ability to determine whether the crossmodal stimuli represent a living or a nonliving thing (i.e. semantic domain) and whether or not they are related in meaning by virtue of belonging to the same object (congruency), remains unanswered.

The present study aims to determine whether the structural integrity of both the pSTS/MTG and PRc are necessary for integrating audiovisual object features, and whether PRc integrity is necessary to process the meaning of multimodal objects. We do this by testing the ability of brain-damaged patients to integrate crossmodal stimuli and by relating their performance to lesion location. Lesion-behaviour studies typically use the lesion-deficit approach in which targeted neural regions of interest are assessed as being damaged or intact and behavioural performance is classified as either normal or impaired. This method suffers from a number of shortcomings. First, brain–behaviour relationships are investigated within *a priori* selected regions of interest, ignoring the potential contribution of other regions in the network that are also involved in the cognitive function. Second, neural tissue in the region of interest is classified as either intact or damaged even though neural tissue may vary in the degree to which it is damaged, and this variation may carry significant functional ramifications. Third, the brain-lesion approach typically binarizes patients' behavioural performance as impaired or normal, a data reductionist approach which neglects informative variation in the extent to which behaviour is impaired. To overcome these limitations, we employed a voxel-based correlation method which correlates continuous measures of neural tissue integrity across the entire brain with continuous measures of behavioural performance on unimodal and crossmodal integration tasks (Tyler *et al.*, 2005*a*, *b*). The additional information carried by the two continuous variables increases the statistical power of this method and thus its sensitivity to brain–behaviour relationships. Moreover, because the analyses cover the entire brain and as patients with diverse sites of damage are included in the analyses, this method has the potential to provide a more complete understanding of the entire neural system involved in a particular cognitive task. In particular, it may detect the involvement of degenerated sites remote from the original lesion which had suffered from diaschisis (von Monakow, 1914). Three recent studies from our laboratory have demonstrated the sensitivity of this method to detect the neural systems involved in complex cognitive tasks (Tyler *et al.*, 2005*a*, *b*; Bright *et al.*, 2007).

The present study relates continuous measures of behavioural data from 16 brain-damaged patients with continuous measures of structural brain integrity (i.e. voxel signal intensities) across their entire brains (Tyler *et al.*, 2005*a*, b). This validated method takes advantage of the range of information available in the variation in signal intensities and behaviour, providing increased statistical power and sensitivity, and covers the entire brain, increasing the likelihood of detecting the involvement of other brain regions in the task. Given our previous findings (Taylor *et al.*, 2006), we hypothesized that both pSTS/MTG and PRc integrity would be necessary for integrating sensory inputs from different modalities, whereas only PRc integrity would be associated with differential impairments depending on the meaning of the crossmodal stimuli (i.e. semantic domain and congruency). Thus, we hypothesized that neural integrity of pSTS/MTG and PRc would correlate with performance for multimodal compared with unimodal integration, irrespective of the meaningfulness of the stimuli (i.e. semantic domain or congruency), whereas only the neural integrity of the PRc would correlate with processing the meaning of the multimodal objects, such that decreased signal intensity in this region would be associated with poorer crossmodal integration performance for living compared with nonliving things and with incongruent compared with congruent crossmodal stimuli. To prefigure the results, only signal intensity in anteromedial temporal and temporopolar regions including the PRc were associated with crossmodal integration performance. We performed two additional analyses to investigate this result. First, we tested whether the absence of correlations involving pSTG/MTG may have been due to a truncated range of signal intensities in that region. We performed additional correlations of whole brain voxel signal intensities with patients' performance on a sentence-picture matching task known to involve the STS/MTG (Tyler and Marslen-Wilson, 2008). If a restricted range of STS/MTG signal intensities were responsible for the lack of correlation with crossmodal integration performance, then performance on the sentence-picture matching control task should likewise not correlate with STG/MTG signal intensities in these same patients. Second, we examined in detail the crossmodal integration performances of individual patients whose lesions included pSTS/MTG region.

Materials and Methods

Participants

Patients were selected from our long-term pool of neuropsychological patients based on the following criteria: English was the native language, availability of a 3D T_1-weighted MRI scan, the brunt of the lesion was in the left hemisphere (because fMRI activity associated with the crossmodal integration task was strongly left-lateralized; Taylor *et al.*, 2006), absence of severe visual and auditory dysfunctions (to ensure they could perform the tasks), and availability of data on a sentence-picture matching task (see Materials and Methods section). Sixteen patients fulfilled these criteria and participated in this study (10 men, 6 women; mean age = 58 years, SD = 11 years; 1 left-hander). Table 1 details each patient's lesion etiology, time since onset, patient age at onset and lesion location (extent). As is apparent from Table 1, the length of time since lesion onset was relatively large (mean: 11.6 years, range 2.0–35.4 years), and the mean age at onset relatively high (i.e. 46.4 years, range 15.8–74.0 years), indicating that at the time of behavioural testing, all patients were in a stable (chronic) stage. Every patient's lesion extended over several gyri, and included regions previously implicated in crossmodal audiovisual integration of meaningful objects (see Introduction, Discussion sections) as well as major post-primary sensory fields, indicating a sufficiently large neuroanatomic lesion extent for the voxel-based correlational analyses. The aetiologies of brain damage were aneurysm, haemorrhagic and ischaemic stroke, herpes simplex encephalitis and resected meningioma. In addition, a group of 12 healthy, right-handed individuals served as behavioural control participants (seven women, five men;

Table 1 Patient number with corresponding gender, lesion etiology, time since onset, age at onset and brain regions affected by lesions

Num	Gender	Etiology	Time since onset (year)	Age at onset (year)	Brain regions affected
1	F	Ischaemic stroke	10.0	58.7	Left inferior and middle frontal gyri, inferior orbital gyrus, insula, superior temporal pole and anterior superior temporal gyrus
2	M	Ischaemic stroke	12.5	41.0	Left inferior orbital gyrus, insula, middle and superior temporal poles, anterior inferior, middle and superior temporal gyri and anterior parahippocampal gyrus
3	M	Ischaemic stroke	35.4	15.8	Left inferior, middle and superior frontal gyri, inferior and middle orbital gyri, insula, rolandic operculum, superior temporal pole, middle temporal gyrus, anterior superior temporal gyrus, pre- and post-central gyri and supramarginal gyrus
4	M	Ischaemic stroke	2.0	74.0	Left superior temporal pole, posterior inferior temporal gyrus, middle and superior temporal gyri, angular gyrus and middle occipital gyrus
5	M	Ischaemic stroke	2.1	63.4	Left posterior middle and superior temporal gyri, angular gyrus and middle occipital gyrus
6	F	Ischaemic stroke	12.9	38.0	Left posterior middle and superior temporal gyri, supramarginal and angular gyri, superior parietal lobule and middle occipital gyrus
7	F	Ischaemic stroke	4.6	62.1	Left inferior and middle frontal gyri, inferior and middle orbital gyri, insula, rolandic operculum and precentral gyrus
8	M	Herpes simplex encephalitis	25.4	23.7	Bilateral middle and superior temporal poles, anterior inferior, middle and superior temporal gyri, fusiform gyrus, parahippo-campal gyrus, amygdala and hippocampus
9	F	Herpes simplex encephalitis	23.5	16.3	Bilateral inferior frontal gyri, insula, rolandic operculum, middle and superior temporal poles, anterior inferior, middle and superior temporal gyri, fusiform gyrus, parahippocampal gyrus, amygdala, hippocampi
10	M	Herpes simplex encephalitis	14.1	31.7	Bilateral insula, middle and superior temporal poles, anterior infer-ior, middle and superior temporal gyri, anterior fusiform gyrus, amygdala and hippocampus
11	F	Herpes simplex encephalitis	11.9	51.0	Bilateral inferior orbital gyri, insula, middle and superior temporal poles, anterior inferior temporal and fusiform gyri, parahippo-campal gyrus, amygdala and hippocampus
12	F	Haemorrhagic stroke	10.9	41.1	Left inferior, middle and superior frontal gyri, inferior and superior orbital gyri, insula, rolandic operculum, middle and superior temporal poles, parahippocampal gyrus, amygdala, hippocampus, superior temporal gyrus, pre- and post-central gyri, supramargi-nal gyrus
13	M	Haemorrhagic stroke	6.3	59.6	Left middle and superior temporal poles, inferior, middle and superior temporal gyri, parahippocampal gyrus, amygdala and hippocampus
14	M	Haemorrhagic stroke	3.6	70.7	Left posterior inferior temporal gyrus, fusiform gyrus and inferior occipital gyrus
15	F	Resected meningioma	7.4	34.6	Left rolandic operculum, posterior inferior, middle and superior temporal gyri and supramarginal gyrus
16	M	Resected meningioma	2.2	60.0	Left inferior, middle and superior frontal gyri, inferior and middle orbital gyri, rolandic operculum, superior temporal pole, superior temporal gyrus, pre- and post-central gyri and inferior parietal lobule

mean age = 62 years, SD = 10 years). Structural MRI scans were not collected from the control participants and they were therefore not included in the VBM analyses.

Crossmodal integration experiment

Materials

The crossmodal stimuli were a representative subset (60%) of those used in our fMRI study (Taylor *et al.*, 2006), selected to have equal numbers of stimuli in the unimodal auditory, unimodal visual and each crossmodal condition (*n* = 60 each), and within each condition, an equal number of congruent and incongruent trials, and within each

congruency condition, an equal number of living and nonliving things, as in the original experiment. Moreover, we ensured that the mean per cent correct and reaction times of each condition in the patient version matched those in the original version (i.e. differences <1 SD of original mean measure). Crossmodal stimuli consisted of 120 photographs of objects, each paired with a sound, half of which were spoken words and half sounds typically produced by target objects (e.g. a picture of a cat paired with the sound typically pro-duced by the target object, e.g. *'miao'*). All pictures were coloured, life-like images of living and nonliving things. The lemma frequencies (Baayen *et al.*, 1995) of the names of object pictures used in the crossmodal integration tasks were comparable for living and nonliving

things [$t(90=)=0.667$, $P=0.507$] and for congruent and incongruent trials [$t(90)=0.233$, $P=0.806$], and the frequencies of the words used in the word-picture crossmodal integration task were comparable for living and nonliving things [$t(45)=0.649$, $P=0.520$] and for congruent and incongruent trials [$t(45)=0.007$, $P=0.994$]. Sounds produced by the target objects were collected from archives available on the internet, and all were characteristic sounds made by the target living or nonliving thing in the absence of ambient sounds from the environment. All environmental sound stimuli were pretested to measure the identifiability of the sound stimuli and to ensure high levels of association between the sounds and objects. In a pretest prior to the experiment, an independent group of nine healthy individuals heard each environmental sound and wrote down the name of the object they thought produced the sound, and afterwards were told which object actually generated the sound and were asked to rate the sound's exemplarity as defined by 'how good an example you think a sound is of the object' on a seven-point scale where 1 = 'a bad example' and 7 = 'a good example'. The identifiability of environmental living and nonliving sounds, defined as the percentage of participants who could unambiguously identify the basic-level object producing the sound, did not significantly differ [$t(58)=0.908$, $P=0.367$]. Moreover, the exemplarity of living and nonliving sounds were not significantly different [$t(57)=1.646$, $P=0.105$; mean ± SD for living things = 4.89 ± 1.06, for nonliving things = 5.41 ± 1.34]. Spoken words were recorded in a sound-proof booth and digitized. The durations of sounds typically produced the target objects and spoken words were matched. Stimuli in the unimodal visual ($n=60$) and unimodal auditory ($n=60$) conditions were constructed by halving picture and auditory stimuli, respectively. Thus, in the unimodal tasks, two visual stimulus halves (unimodal visual condition) and two sounds produced by the target object or auditory word halves (unimodal auditory conditions) were presented. All 120 crossmodal and all 120 unimodal stimuli were unique. Half the trials in each condition were congruent (i.e. the sound and picture matched; e.g. a picture of a cow and the sound produced by the target object cow, i.e. 'moo') whereas in the other half they were incongruent [e.g. the sound produced by the target object dog ('woof') and a picture of an elephant]. Within each congruency condition, half of the stimuli represented objects from the domain of living things and half from the nonliving domain (Fig. 1). Congruency and domain were two variables measuring the meaningfulness of stimuli.

Procedure

Auditory and visual stimuli were presented simultaneously in the crossmodal conditions because crossmodal integration effects occur during simultaneous, but not serial, presentations of crossmodal stimuli (Calvert, 2001). In the unimodal conditions, participants were presented with two halves of stimuli. In the visual condition, these were presented simultaneously, one to the left and one to the right of screen centre, and in the auditory condition, the two monophonic halves were presented sequentially to both ears, separated by 750 ms of silence (Fig. 1). Pictures in the crossmodal conditions and the unimodal visual baseline stimuli were displayed for 1800 ms, and the mean duration of the unimodal auditory baseline was 1446 ms. The tasks were presented in the following fixed, blocked order: unimodal visual baseline, crossmodal words-pictures, unimodal auditory baseline with words, unimodal auditory baseline with sounds and crossmodal sounds-pictures. Participants were instructed to decide whether the two stimuli 'went together' and to press a corresponding button on a button box. Reaction times (RTs) were measured from the onset of the presentation of the picture or picture halves and, in the unimodal auditory condition, at the onset of the second sound half
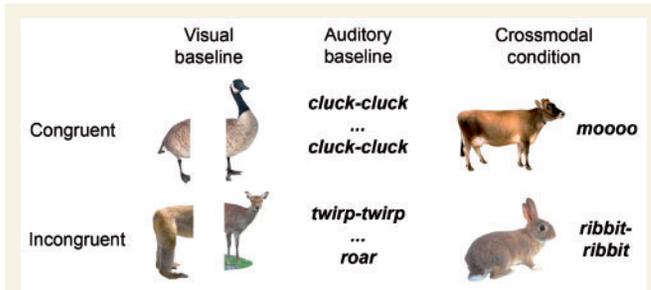

Figure 1 Example stimuli belonging to the living things domain.

(i.e. the point in time where two auditory stimuli could likewise be integrated). DMDX software (Forster and Forster, 2003) controlled stimulus timing and presentation as well as the collection of accuracy data.

Sentence-picture matching task

Successful syntactic processing has been linked with left inferior frontal cortex as well as middle to posterior regions of the STG/MTG and inferior parietal regions (Caplan and Hildebrant, 1988; Amici *et al.*, 2007; Tyler and Marslen-Wilson, 2008). To acquire an independent measure of STS/MTG function to correlate with whole brain voxel signal intensities, thereby testing for an adequate range of voxel signal intensities in the STS/MTG region, we collected behavioural data from all patients on a sentence-picture matching task stressing syntactic processing (Tyler *et al.*, 2002). This task presents participants with syntactically simple and complex spoken sentences, and asks participants to select from an array of three pictured scenes the one which depicts the sentence's meaning (Tyler *et al.*, 2002). Each patient's performance in terms of per cent correct ($n=34$) was used in voxel-based correlation analyses.

Voxel-based correlation analyses

High-resolution T_1-weighted MRI scans of 1×1 mm in-plane resolution and 1–2 mm slice thickness were available for all patients, and during spatial normalization these were resampled to 1 mm³ isotropic voxel resolution. Preprocessing involved spatial normalization of the patients' skull-stripped T_1 images to the MNI template using SPM2 (Wellcome Institute of Cognitive Neurology, London, UK). The same linear (12 affine transformations) and non-linear ($7 \times 8 \times 7$ basis functions) transformations were used for all patients. Images were not segmented into grey and white matter as this procedure results in substantial misclassifications when brains contain large lesions (Stamatakis and Tyler, 2005). Moreover, because all images are normalized to a standard 3D space, the neuroanatomical location of significant correlations between specific voxel coordinates and behavioural performance is known. A manual control of the normalization quality by visual inspection revealed errors in periventricular regions. To avoid erroneous findings in these regions resulting from misalignment, a ventricular mask was created to remove the ventricles from the statistical modelling. The outcome of each normalization procedure was carefully visually inspected to ensure that there were no anatomical distortions. Images were smoothed with a 10 mm Gaussian kernel to account for minor anatomical variations and to increase the signal-to-noise ratio (Friston *et al.*, 1994).

Voxel-based covariate analyses were conducted with the general linear model approach (Friston *et al.*, 1995) implemented in SPM2

Table 2 Mean (±SD) control participants' and patients' percent correct performance on the unimodal and crossmodal integration tasks and corresponding performance indices

Performance measure	Control participants (*n* = 12)	Patients (*n* = 16)	Mann-Whitney U	P
Percent correct				
Auditory baseline, sounds	96.1 ± 3.7	95.0 ± 4.4	83.0	NS
Auditory baseline, words	91.1 ± 4.8	77.3 ± 22.6	41.0	<0.05
Visual baseline	95.0 ± 3.8	93.0 ± 4.7	72.5	NS
Crossmodal integration, sounds+pictures	92.9 ± 3.0	85.8 ± 14.0	83.5	NS
Crossmodal integration, words+pictures	97.9 ± 3.1	83.4 ± 25.6	47.0	<0.05
Performance indices				
Crossmodal integration index	1.01 ± 0.02	0.98 ± 0.13	89.0	NS
Crossmodal integration index for living versus nonliving things	1.02 ± 0.06	0.95 ± 0.12	58.5	NS
Crossmodal integration index for incongruent versus congruent stimuli	0.97 ± 0.08	0.98 ± 0.12	74.0	NS

(Wellcome Institute of Cognitive Neurology, London, UK). Task performance measures of each patient were correlated with the corresponding signal intensity values of each voxel in each preprocessed brain image (Tyler *et al.*, 2005*a*, *b*). Global mean signal from each scan was included in the model as a confounding covariate. The *t*-statistic was employed to test for regional effects in each analysis, and associated *P*-values were corrected for search volume with the Gaussian random field theory. Significant results therefore represent brain regions (i.e. clusters of voxels) in which voxel-level signal intensities in the cluster were significantly correlated with the behavioural scores. Statistical parametric maps were thresholded at the standard $P < 0.001$ uncorrected at the voxel level and the coordinates of the peak significant voxels in clusters surviving a random field corrected *P*-value of <0.05 adjusted for the entire brain are reported. We further investigate regions approaching significance at the $P < 0.001$ threshold by reporting highly significant clusters (i.e. random field corrected $P < 0.000001$ adjusted for the entire brain) at a voxel threshold of $P < 0.01$. Coordinates are reported in MNI space and with corresponding Talairach and Tournoux labels (Talairach and Tournoux, 1988).

Results

The performance of control participants and patients on each task, and corresponding values on the crossmodal integration indices described below, are shown in Table 2. Patients' per cent correct performance differed from that of controls only on the auditory baseline task with words, and correspondingly on the crossmodal integration task with words and pictures. Functional imaging studies identify regions involved in crossmodal integration by testing for crossmodal integration effects, i.e. those regions responding more during crossmodal compared with unimodal integration trials, to control for baseline unimodal performance (Taylor *et al.*, 2006). Correspondingly, we calculated an index of relative crossmodal to unimodal integration performance [(% correct on crossmodal integration tasks)/(% correct on unimodal integration tasks); hereafter referred to as 'crossmodal integration index'], where smaller values indicate worse crossmodal relative to unimodal integration performance. Because the crossmodal integration index controls for baseline unimodal integration performance,

this measure did not significantly differ for patients and control participants (Table 2), nor did crossmodal integration indices for living versus nonliving things and for incongruent versus congruent stimuli (described below).

In a first analysis, crossmodal integration indices were correlated with voxel signal intensities across the entire brains of all patients. Significant positive correlations, i.e. regions where lower signal intensities were associated with relatively poorer crossmodal integration performance were found in bilateral temporopolar regions which included the bilateral temporal poles (BAs 38, right 20, 21, 36), anterior inferior and middle temporal lobes (BA 20, 21) and the right fusiform gyrus (BA 36), as well as the right anteromedial temporal lobe including the amygdala and hippocampus. A cluster approaching significance in the left anterior and anteromedial temporal lobe ($P = 0.094$; peak coordinates −35, −30, −15) included the parahippocampal cortex (BA 37), hippocampus and inferior temporal lobe (BA 20; see Table 3). At a reduced threshold ($P < 0.01$), these clusters extended into the bilateral anteromedial temporal lobe including the perirhinal, entorhinal and parahippocampal cortices (left: BAs 20, 30, 36, 37; right: BAs 20, 28, 35, 36, 37), the left hippocampus, and bilateral anterior superior temporal lobes (BA 21) (Fig. 2). To explore these findings, we plotted the per cent correct on unimodal and crossmodal integration trials against the peak voxel signal intensities of each cluster. While the correlations of unimodal and crossmodal integration performance with peak voxel signal intensities did not differ in the right temporopolar and left inferior temporal lobe clusters (William's test: $t = 1.55$, $P = NS$ and $t = 1.827$, $P = 0.09$, respectively), these correlations significantly differed for the cluster containing the left anteromedial temporal lobe including the perirhinal cortex (William's test: $t = 2.287$; $P < 0.05$), reflecting a stronger dependence of crossmodal than unimodal integration performance on neural integrity of the left anteromedial and anterior inferior temporal lobe (Fig. 3A). Thus, these results show that only decreased neural integrity in bilateral temporopolar and anteromedial temporal lobe regions—in particular the left anteromedial and inferior anterior temporal lobe—was associated with poor crossmodal integration performance. In contrast, the pSTS/MTG does not seem to be necessary for processing crossmodal stimuli.
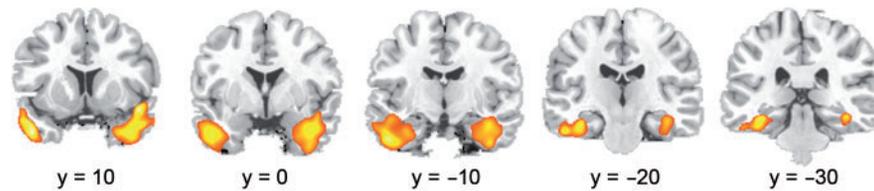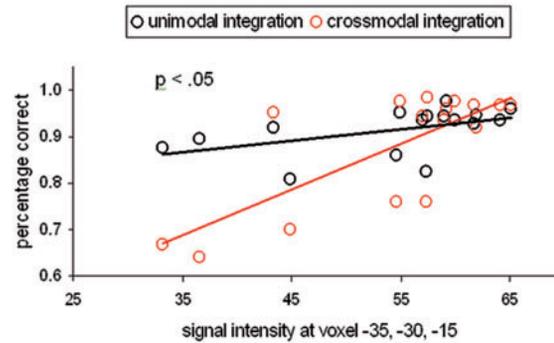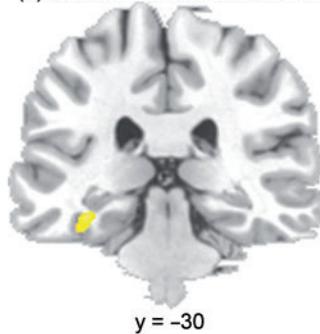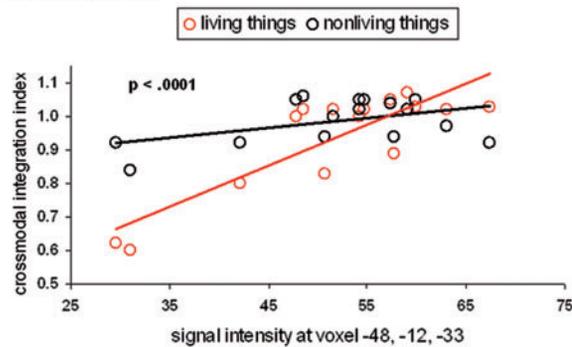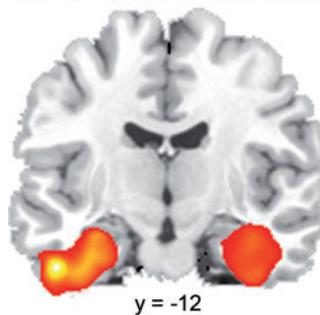
Figure 2 Neuroanatomical regions where relatively poorer crossmodal compared with unimodal integration performance correlated with decreased measures of neural integrity (voxel signal intensities; results are thresholded at $P < 0.01$; see text for details). MNI coordinates are reported, and L = L.
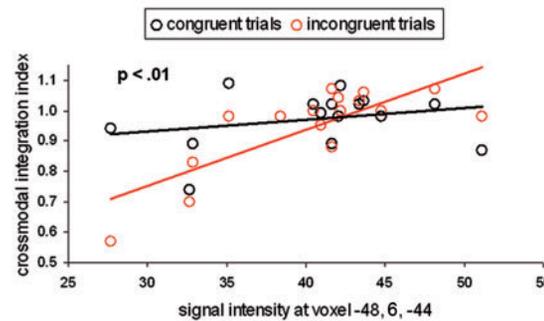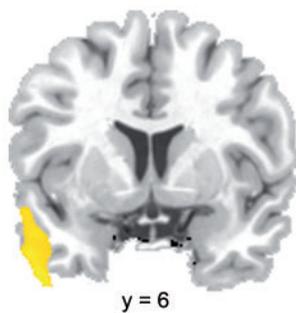

Figure 3 (Left panel) Neuroanatomical regions (based on results reported in Table 3) where voxel signal intensities were significantly correlated with crossmodal compared with unimodal integration performance (**A**), crossmodal integration of living compared with nonliving things (**B**) crossmodal integration of meaningfully unrelated (incongruent) versus meaningfully related (congruent) stimuli (**C**). The right panel shows the corresponding relationships between the peak voxel signal intensities and behavioural performance measures (see text for details). MNI coordinates are reported, and L = L.

Table 3 Voxel-based correlations of voxel signal intensities across the entire brains of 16 patients with measures of relative performance on crossmodal compared with unimodal integration tasks (crossmodal integration effects), and relative crossmodal integration performance with incongruent compared with congruent and with living compared with nonliving stimuli (BA = Brodmann's area; MNI coordinates are reported)

Cluster level		Voxel level		Coordinates			Location	
$p_{corrected}$	Extent	$p_{corrected}$	Z	X	y	z	BA	Region
Crossmodal integration index								
0.000	4884	0.896	3.92	49	13	−15	20, 21, 36, 38	Right amygdala, hippocampus, temporal pole, inferior, middle and
		0.999	3.50	39	12	−35		superior temporal lobes and fusiform gyrus
		1.000	3.40	41	−1	−26		
0.000	4321	0.953	3.81	−45	11	−34	20, 21, 36, 38	Left temporal pole, inferior and middle temporal lobes
		0.987	3.69	−47	−9	−34		
		0.988	3.68	−51	6	−26		
0.094	1086	0.998	3.54	−35	−30	−15	20, 37	Left parahippocampal cortex, hippocampus, inferior temporal lobe
		0.999	3.50	−36	−23	−20		and fusiform gyrus
Crossmodal integration of living versus nonliving things								
0.000	25 433	0.009	5.45	−48	−12	−33	20, 21, 30, 35,	Left perirhinal and parahippocampal cortices, amygdala, hippocam-
		0.025	5.20	−49	−20	−28	36, 37, 38	pus, temporal pole, inferior and middle temporal lobes, fusiform
								gyrus
		0.074	4.92	−42	9	−39		
0.000	26 066	0.732	4.10	42	−9	−37	20, 21, 28, 30,	Right perirhinal, entorhinal and parahippocampal cortices, amyg-
		0.777	4.06	39	4	−43	34, 35, 36,	dala, hippocampus, temporal pole, inferior and middle temporal
		0.814	4.02	39	−27	−13	37, 38, 48	lobes, fusiform gyrus
Crossmodal integration of incongruent versus congruent stimuli								
0.000	3639	0.848	3.97	−48	6	−44	20, 21, 38	Left temporal pole and inferior and middle temporal lobes
		0.863	3.95	−56	0	−36		
		0.997	3.56	−52	9	−24		

In a further analysis, we focussed on the effects of our meaning manipulations— semantic domain and congruency. Our previous fMRI study (Taylor *et al.*, 2006) found that only PRc, and not pSTS/MTG, activity was modulated by the meaning of the crossmodal stimuli. Specifically, PRc responses were greater during the crossmodal integration of living compared with nonliving things and with meaningfully unrelated compared with meaningfully related (i.e. incongruent versus congruent) stimuli. To determine whether the neural integrity of the left PRc and potentially other regions are necessary to develop the meaning of crossmodal object features, we performed two additional analyses. First, we performed a correlation analysis to identify neural regions sensitive to semantic category (i.e. living versus nonliving) during crossmodal integration. In this analysis, the same voxel signal intensity data used in the previous analysis were correlated with a new measure of crossmodal integration performance for living compared with nonliving things [i.e. (crossmodal integration index for living things)/(crossmodal integration index for nonliving things), where lower values indicate a selective crossmodal integration deficit for living things]. Significant positive correlations (i.e. lower signal intensities correlated with selective difficulty integrating crossmodal features of living things) were found in the bilateral anteromedial temporal lobes including the perirhinal cortices (BAs 20, 30, 35, 36, 37, hippocampi and right 28, 34), temporal poles (BAs 20, 21, 36, 38), anterior inferior and middle temporal lobes (BAs 20, 21, 36) and fusiform gyri (BAs 20, 36, 37; see Table 3). To further explore these findings, we plotted the crossmodal integration indices for living and nonliving things

against the peak voxel signal intensities in each cluster. These correlations of living and nonliving integration indices with peak voxel signal intensities were significantly different in both the left and right hemisphere clusters (William's test: $t = 6.466$, $P < 0.0001$ and $t = 3.617$, $P < 0.01$, respectively), indicating significantly stronger relationships between voxel signal intensities and crossmodal integration performance for living compared with nonliving things (see Fig. 3B, where right panel depicts behavioural correlations with signal intensities in peak voxel of left hemisphere cluster).

Second, we correlated whole brain voxel signal intensities with a measure of the patients' relative performance on incongruent (not meaningfully related) compared with congruent (meaningfully related) crossmodal integration trials [i.e. (crossmodal integration index for incongruent trials)/(crossmodal integration index for incongruent trials)], where lower values indicate poorer incongruent crossmodal integration performance. Significant positive correlations, i.e. those regions where lower signal intensities were associated with a relative performance deficit with incongruent crossmodal stimuli, were found in the left temporal pole (BAs 21, 38) and left anterior inferior and middle temporal lobe (BAs 20, 21). At a reduced threshold ($P < 0.01$), correlations involved more extensive left temporal regions including the parahippocampal gyrus (BA 36), hippocampus, amygdala and fusiform gyrus (BAs 20, 37) as well as homologous but less extensive regions in the right hemisphere (Table 3). We plotted the crossmodal integration indices for incongruent compared with congruent pairs against the peak voxel signal intensity. As expected, the crossmodal integration of stimulus pairs which were not meaningfully

related (incongruent) showed stronger association with voxel signal intensities than the crossmodal integration of meaningfully related stimulus pairs (William's test: $t = 4.203$, $P < 0.01$; see Fig. 3C). Thus, reduced signal intensities in the temporopolar and anteromedial temporal lobe regions were associated not only with relative impairments in crossmodal compared with unimodal integration of object features, but with selective impairments in processing the *meaning* of the crossmodal object features, a pattern consistent with fMRI data (Taylor *et al.*, 2006).

These results suggest that integrity of bilateral anteromedial temporal and temporopolar regions is necessary for preserved ability to integrate audiovisual object features. Our previous fMRI study (Taylor *et al.*, 2006) with healthy individuals and those of Beauchamp *et al*. (2004*a*, *b*) additionally found crossmodal integration effects in the pSTS/MTG. Because we did not find any correlations between crossmodal integration performance and signal intensity in this region we carried out two additional sets of analyses.

To further investigate the lack of a significant correlation between signal integrity in the STG/MTG and our indices of crossmodal integration, we tested the possibility that the range of signal intensities in the pSTS/MTG was too small to result in a significant correlation. Thus, we correlated behavioural scores on a task known to be associated with the STG/MTG (but not with PRc) integrity—a sentence-picture matching task—with voxel signal intensities across the whole brains of all patients. We chose the sentence-picture matching task because it is sensitive to syntactic processing and is associated with activity in regions of the STG/MTG known to be involved in syntactic analysis (Caplan *et al.*, 1996; Tyler and Marslen-Wilson, 2008). If the range in pSTS/MTG signal intensity values across all patients was too narrow to result in significant correlations with behavioural crossmodal integration performance, then signal intensities in this region should not correlate with performance on any task—including the sentence-picture matching task. We found that per cent correct on the sentence-picture matching task was positively correlated with signal intensity in the LMTG (BAs 20, 21; peak coordinates: −51, −31, −9; BA 20), suggesting that lower signal intensity (i.e. reduced neural integrity) was related to poorer sentence-picture performance. At a reduced threshold ($P < 0.01$), the correlations involved extensive left temporal regions including the superior temporal lobe (BAs 20, 21, 22, 41, 42), Heschl's gyrus, the rolandic operculum and inferior temporal lobe (BA 20) (Fig. 4). Thus, these findings confirm the significant relationship between STG/MTG and RO with syntactic processing (e.g. Caplan *et al.*, 1996; Amici *et al.*, 2007; Tyler and Marslen-Wilson, 2008), and show that the range in STG/MTG signal intensities was large enough to detect this functional–anatomic relationship. Moreover, an examination of the behavioural scores of three patients whose lesions included the pSTS/MTG cluster identified in our previous fMRI study (Taylor *et al.*, 2006) (Fig. 5) showed that these patients did not have poorer performance on the crossmodal compared with the unimodal tasks. Behavioural scores of Patients 5 and 6 were within normal limits as measured by z-scores (standard scores). Patient 4 was impaired on both the unimodal and crossmodal integration tasks but performance on the crossmodal integration task was not significantly different from that on the
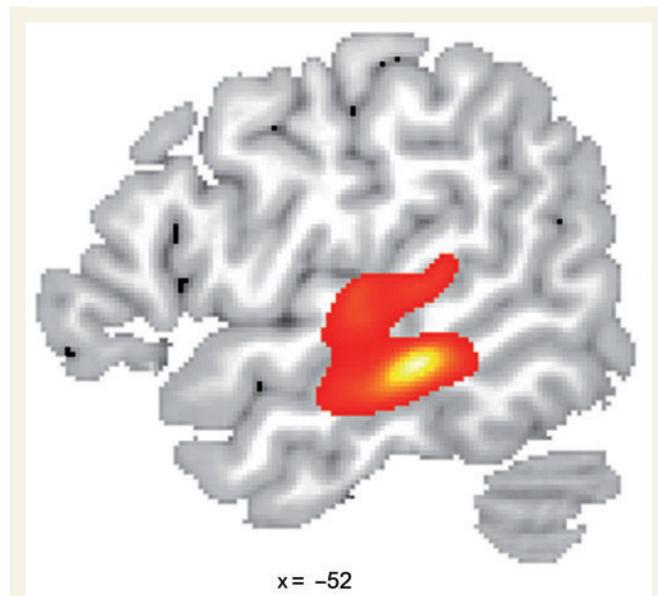

$x = -52$

Figure 4 Neuroanatomical left hemisphere regions where voxel signal intensities significantly correlated with performance on a sentence-picture matching task (MNI coordinates are reported).

unimodal integration task [χ^2 (df = 1) = 1.62, $P = $ NS] (see Fig. 5 and Table 1 for lesion descriptions).

Discussion

We found that reduced structural integrity of bilateral temporal polar and anteromedial temporal lobe regions, including the PRc, was associated with poorer crossmodal compared with unimodal integration scores in a series of brain-damaged patients. These results suggest that anteromedial temporal and temporopolar regions including the PRc are necessary for crossmodal integration of audiovisual object features, consistent with the non-human primate findings (Murray *et al.*, 1998). However, the neural integrity of pSTS/MTG was not associated with crossmodal integration performance. To test whether a truncated range of pSTS/MTG voxel signal intensities was responsible for this negative finding, we correlated voxel signal intensities with behavioural scores on a syntactic control task known to be associated with STS/MTG (Tyler and Marslen-Wilson, 2008). The resulting significant correlation between syntactic performance and neural integrity of the STS/MTG suggests that a diminished range of signal intensities in this region was not responsible for its lack of correlation with crossmodal integration performance. A second exploratory analysis examined the unimodal and crossmodal integration performance of patients with lesions restricted to the pSTS/MTG and also failed to support the hypothesized relationship. These results suggest that while the pSTS/MTG region may co-activate with PRc in healthy subjects during crossmodal integration tasks, it is not necessary to bind crossmodal sensory objects.

These findings are consistent with hierarchical models of object processing which give a central role to anteromedial temporal
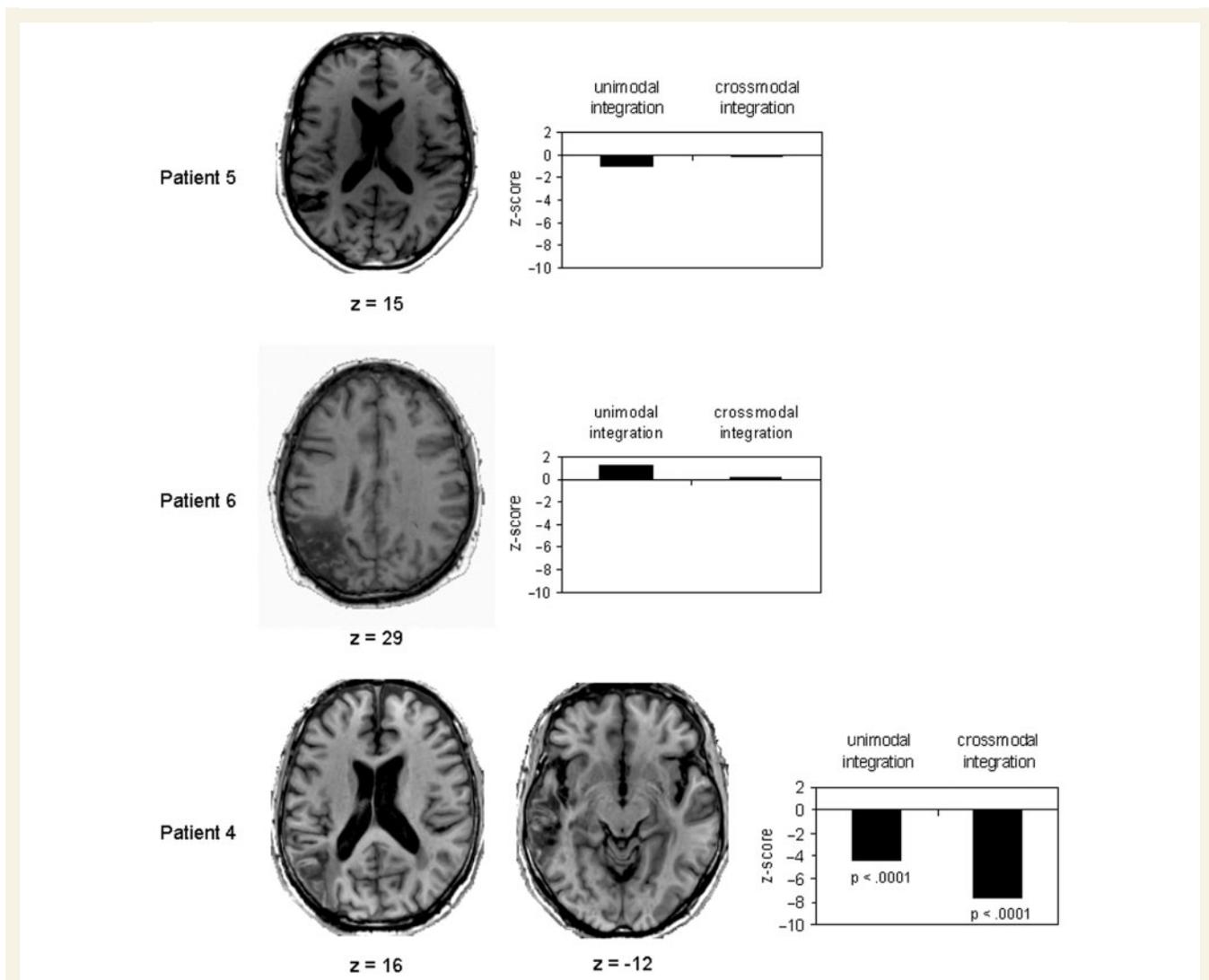
Figure 5 Patients with pSTS/MTG lesions do not show crossmodal integration deficits. The left panels shows the normalized T_1-weighted MRI scans of three patients whose lesions included the pSTS/MTG region previously associated with crossmodal integration performance and the right panel shows their corresponding performances on the unimodal and crossmodal integration tasks (z-scores, where negative values indicate relatively poorer performance compared with a group of control participants). MNI coordinates are reported, and L = L.

lobe, in particular the PRc, in the crossmodal integration of meaningful object features (Murray *et al.*, 1998; Murray and Bussey, 1999; Simmons and Barsalou, 2003; Bussey *et al.*, 2005; Taylor *et al.*, 2007). According to these models, hierarchical information processing streams operate largely independently within each sensory modality, processing increasingly more complex combinations of unisensory object features. Thus, visual object processing is claimed to proceed from posterior (V1, V2, V3, V4/V8) through more anterior sites (e.g. lateral-occipital complex) to anterior ventral temporal and anteromedial temporal lobe regions (Felleman and Van Essen, 1991; Lerner *et al.*, 2001; Murray and Richmond, 2001). In the auditory modality, increasingly more complex information is processed from primary auditory cortex in Heschl's gyrus to the surrounding secondary auditory cortex in the superior temporal gyrus, including posterior superior

temporal gyrus (Wise *et al.*, 1991; Rauschecker and Tian, 2000; Semple and Scott, 2003; Lewis *et al.*, 2004). The crossmodal integration of object features is thought to take place where the different sensory streams converge (Murray and Bussey, 1999; Bussey *et al.*, 2005). Anatomical tracing studies in non-human primates demonstrate that the first site of convergence in the ventral object processing stream is the PRc (Suzuki and Amaral, 1994). This anatomically grounded hypothesis that PRc is a critical component in the network integrating crossmodal object features has received converging behavioural support from ablations studies in animals (Murray *et al.*, 1998; Parker and Gaffan, 1998), our previous fMRI study with healthy human participants (Taylor *et al.*, 2006) and now the results of the present study.

The discrepancy between previous fMRI findings of crossmodal integration effects in pSTS/MTG (Beauchamp *et al.*, 2004*a, b*;

Taylor et al., 2006) and the present results highlight the need to conduct studies with lesioned patients to complement fMRI findings. Taken together, the results suggest that pSTS/MTG plays a supportive but non-essential role in the crossmodal integration of audiovisual object features. The function of this supportive role may be to transform the temporal auditory and spatial visual information into a common code, thereby facilitating crossmodal integration (Beauchamp, 2005), or to generate multidirectional sensory predictions based on partially shifting receptive fields (Pouget et al., 2002; Avillac et al., 2005). Irrespective of the mechanism, the lack of semantic effects in the pSTS/MTG region during crossmodal integration in humans (Beauchamp et al., 2004a, b; Taylor et al., 2006) is consistent with this purely perceptual, translational hypothesis which facilitates, but is not required for successful crossmodal integration of object features. This object processing task must be differentiated from conceptual processes involving auditory and visual verbal material, which have been associated with the posterior temporal and inferior parietal lobes (Hart and Gordon, 1990).

Clearly, the brain regions implicated in crossmodal integration extend beyond the borders of the PRc, although these appear to be underestimated in the standard atlases (Talairach and Tournoux, 1988; Insausti et al., 1998; Suzuki and Amaral, 2003; Taylor and Probst, 2008). These included upstream structures in anterior inferior temporal lobe which provide the primary source of input to the PRc (Suzuki and Amaral, 1994; see also Catani et al., 2003; Powell et al., 2004). Thus, anterior inferotemporal damage would disconnect PRc from the majority of its ventral stream input, impairing the capacity of this structure to integrate visual with other sensory information. On the other hand, anterior inferotemporal and temporopolar structures have been associated with naming both auditory and pictorial objects, suggesting that these structures perform the modality neutral function of supporting the retrieval of word forms from the conceptual system (Tranel et al., 2005). Thus, anterior inferotemporal and temporopolar structures may be involved in the automatic activation and explicit retrieval of word-form information associated with the audiovisual objects.

The structural integrity of other anteromedial structures situated downstream to the PRc was also associated with crossmodal integration ability, including the entorhinal cortex and hippocampus. Connectivity findings in non-human primates suggest that the role of these structures is to elaborate the crossmodal object representation processed in the PRc. Specifically, the majority of inputs to the PRc originate in unisensory cortices. The associative interconnections within the PRc suggest that unisensory information is integrated in this structure. The majority of PRc outputs is forwarded to the entorhinal cortex, which receives most of its inputs from higher-order, associative cortex. Associative interconnections in the entorhinal cortex likewise indicate that this structure may bind multimodal object information received from PRc with other, higher order associative information. Finally, entorhinal cortex outputs its information via the perforant path to the hippocampal system, itself characterized by a complex system of interconnectivity, which is presumably responsible for synthesizing the highly integrated information it receives (Lavenex and Amaral, 2000, see also Davachi, 2006). Ablation studies in animals offer partial support for this 'hierarchy of associativity' model: while

rhinal cortex ablations (i.e. those including both the entorhinal and PRc) result in crossmodal integration deficits, lesions of the hippocampus (and the amygdala) do not lead to persistent impairments in the retention or new learning of crossmodal associations in monkeys (Murray et al., 1998).

Future elaborations on hierarchical multimodal object processing models must be able to account for the results of recent anatomic tracing studies and temporally sensitive imaging techniques (e.g. EEG, MEG) which indicate early interactions between the sensory cortices when multimodal objects are processed (Schroeder et al., 2003; Foxe and Schroeder, 2005; Ghazanfar and Schroeder, 2006; Macaluso, 2006; Driver and Noesselt, 2008). For example, anatomical tracing studies in monkeys have demonstrated connections from the core and belt auditory regions to primary visual areas (Falchier et al., 2002; Rockland and Ojima, 2003). Similar connections in humans could account for early audiovisual interaction effects in visual cortex starting at 40 ms post-stimulus as measured by event-related potentials (ERPs; Giard and Peronnet, 1999) and between 35 and 65 ms post-stimulus as measured by MEG (Shams et al., 2005), as well as audiovisual interaction effects in auditory cortex between 90 and 110 ms post-stimulus (ERPs; Giard and Peronnet, 1999) (see also Molholm et al., 2004). Despite the growing evidence of early multisensory interactions, the present results suggest that they are not sufficient to independently support the crossmodal integration of object features into meaningful representations. Instead, early interactions may support low-level multimodal object processing by modulating neuronal oscillations in primary sensory cortices (Lakatos et al., 2007), thereby perhaps enhancing the sensory perception of the multimodal objects, especially in noisy environments (Macaluso, 2006). While these early effects may allow sensory stimuli to be attended, thereby entering the processing stream, the processing of meaningfully related crossmodal objects appears to depend on more downstream structures receiving more highly integrated and multimodal information, i.e. anteromedial temporal lobe.

The successful combination of multi-sensory object features is not sufficient to know what an object means, which is the goal of the object processing system. To determine which anatomical regions are implicated in processing meaningful aspects of the crossmodal stimuli, we performed correlations between voxel-based signal intensities across whole brains and performance on meaningfully unrelated versus related crossmodal integration trials (e.g. a cat picture and the sound 'woof' versus a dog picture and the sound 'chirp'), and with performances on crossmodal integration trials with stimuli representing living versus nonliving things. The neural integrity of bilateral temporopolar and anteromedial temporal regions was significantly associated with both semantic measures. These findings are consistent with those from our fMRI study, where PRc activity was greater during the crossmodal integration of incongruent compared with congruent and living compared with nonliving things (Taylor et al., 2006).

Hierarchical models of perceptual object processing were not developed to account for such semantic effects. Thus, to understand how object meanings are represented and processed additionally requires a cognitive model of semantic object memory. One set of cognitive theories shares key architectural assumptions with hierarchical object processing models, i.e. that objects are

represented and processed in a distributed, feature-based system. The Conceptual Structure Account (Tyler and Moss, 2001; Tyler et al., 2000) represents one such distributed cognitive model (see also McRae et al., 1997; Rogers and McClelland, 2004). It proposes that objects in different categories can be characterized by the number and statistical properties of their constituent features. For example, living things are composed of large clusters of features, most of which are shared by other category members (e.g. many animals have four legs, eyes and a nose). Nonliving things, on the other hand, typically have smaller clusters of features and relatively more features that are distinctive to the particular object (i.e. not shared by other category members). While shared features are ambiguous with respect to object identity, distinctive features are informative about the identity of an object [e.g. it is more useful to know that the object has an udder (a distinctive feature) than that the object has eyes (shared feature)]. By integrating this cognitive account of object representations with the hierarchical object processing account, one would predict that those objects with many shared features (i.e. living things) which are more difficult to differentiate from one another would require more complex conjunctions of features to be identified than objects with relatively more distinctive features (i.e. nonliving things). This hypothesis was confirmed in our recent fMRI experiment, in which participants were presented with pictures of living and nonliving things, and instructed to name (i.e. identify) each picture. Functional activity in anteromedial temporal lobe regions was significantly greater during the naming of living compared with matched sets of nonliving things (Moss et al., 2005). Thus, the significant correlation between the structural integrity of anterior ventral temporal and anteromedial temporal lobe regions and performance on the crossmodal integration trials with living compared with nonliving things provides additional support an integrated 'neurocognitive' account which combines the neural model of hierarchical object processing with a cognitive account of conceptual representations (Moss et al., 2002; Tyler et al., 2004; Taylor et al., 2007).

References

Amedi A, von Kriegstein K, van Atteveldt NM, Beauchamp MS, Naumer MJ. Functional imaging of human crossmodal identification and object recognition. Exp Brain Res 2005; 166: 559–71.

Amici S, Brambati SM, Wilkins DP, Ogar J, Dronkers NL, Miller BL, et al. Anatomical corelates of sentence comprehension and verbal working memory in neurodegenerative disease. J Neurosci 2007; 27: 6282–90.

Avillac M, Deneve S, Olivier E, Pouget A, Duhamel J-R. Reference frames for representing visual and tactile locations in parietal cortex. Nat Neurosci 2005; 8: 941–9.

Baayen R, Piepenbrock R, Gulikers L. The CELEX Lexical Database (Release 2) (CD-ROM). Philadelphia: Linguistic Data Consortium, University of Pennsylvania; 1995.

Barraclough NE, Xiao D, Baker CI, Oram MW, Perrett DI. Integration of visual and auditory information by superior tmeporal sulcus neurons responeive to the sight of actions. J Cogn Neurosci 2005; 17: 377–91.

Barsalou LW. Perceptual symbol systems. Behav Brain Sci 1999; 22: 577–660.

Beauchamp MS. See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. Curr Opin Neurobiol 2005; 15: 1–9.

Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A. Unraveling multisensory integration: patchy organization within human STS multisensory cortex. Nat Neurosci 2004a; 7: 1190–2.

Beauchamp MS, Lee KE, Argall BD, Martin A. Integration of auditory and visual information about objects in superior temporal sulcus. Neuron 2004b; 41: 809–23.

Benevento LA, Fallon J, Davis BJ, Rezak M. Auditory-visual interaction in single cells in the cortex of the superior temporal sulcus and the orbital frontal cortex of the macaque monkey. Exp Neurol 1977; 57: 849–72.

Bright P, Moss HE, Longe O, Stamatakis EA, Tyler LK. Conceptual structure modulates anteromedial temporal involvement in processing verbally presented object properties. Cereb Cortex 2007; 17: 1066–73.

Bruce C, Desimone R, Gross CG. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. J Neurophysiol 1981; 46: 369–84.

Bussey TJ, Saksida LM, Murray EA. The perceptual-mnemonic/feature conjunction model of perirhinal cortex function. Q J Exp Psychol 2005; 58B: 269–82.

Calvert GA. Crossmodal processing in the human brain: insights from functional neuroimaging studies. Cereb Cortex 2001; 11: 1110–23.

Caplan D, Hildebrant N. Disorders of syntactic comprehension. Cambridge, MA: MIT Press; 1988.

Caplan D, Hildebrant N, Makris N. Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. Brain 1996; 119: 933–49.

Catani M, Jones DK, Donato R, ffytche DH. Occipito-temporal connections in the human brain. Brain 2003; 126: 2093–107.

Cree GS, McRae K. Analyzing the factors underlying the structure and computation of the meaning of Chipmunk, Cherry, Chisel, Cheese, and Cello (and many other such concrete nouns). J Exp Psychol Gen 2003; 132: 163–201.

Damasio AR. Time-locked multiregional retroactivation: a systems-level proposal for the neural substrates of recall and recognition. Cognition 1989; 33: 25–62.

Damasio AR. Category-related recognition defects as a clue to the neural substrates of knowledge. Trends Neurosci 1990; 13: 95–98.

Davachi L. Item, context and relational episodic encoding in humans. Curr Opin Neurobiol 2006; 16: 693–700.

Desimone R, Ungerleider LG. Neural mechanisms of visual perception in monkeys. In: Boller F, Grafman J, editors. Handbook of neuropsychology. Vol. 2. Amsterdam: Elsevier; 1989. p. 267–99.

Driver J, Noesselt T. Multisensory interplay reveals crossmodal influences in 'sensory-specific' brain regions, neural responses and judgments. Neuron 2008; 57: 11–23.

Ettlinger G, Wilson WA. Cross-modal performance: behavioural processes, phylogenetic conserations and neural mechanisms. Behav Brain Res 1990; 40: 169–92.

Falchier A, Clavagnier S, Barone P, Kennedy H. Anatomical evidence of multimodal integration in primate striate cortex. J Neurosci 2002; 22: 5749–59.

Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex 1991; 1: 1–47.

Forster KI, Forster JC. A windows display program with millisecond accuracy. Behav Res Meth Instrum Comput 2003; 35: 116–24.

Foxe JJ, Schroeder CE. The case for feedforward multisensory convergence during early cortical processing. Neuroreport 2005; 16: 419–23.

Friston KJ, Jezzard P, Turner R. Analysis of functional MRI time series. Hum Brain Mapp 1994; 1: 153–71.

Friston KJ, Holmes AP, Worsley KJ, Poline J-B, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: A general linear approach. Hum Brain Mapp 1995; 2: 189–210.

Ghazanfar AA, Schroeder CE. Is neocortex essentially multisensory? Trends Cogn Sci 2006; 10: 278–85.

Giard MH, Peronnet F. Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. J Cogn Neurosci 1999; 11: 473–90.

Hart J, Gordon B. Delineation of single-word semantic comprehension deficits in aphasia, with anatomical correlation. Ann Neurol 1990; 27: 226–31.

Hikosaka K, Iwai E, Saito H-A, Tanaka K. Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. J Neurophysiol 1988; 60: 1615–37.

Insausti R, Juottonen K, Soininen H, Insausti AM, Partanen K, Vainio P, et al. MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. Am J Neuroradiol 1998; 19: 659–71.

Lakatos P, Chen C-M, O'Connell MN, Mills A, Schroeder CE. Neuronal oscillations and multisensory interaction in primary auditory cortex. Neuron 2007; 53: 279–92.

Lavenex P, Amaral DG. Hippocampal-neocortical interaction: a hierarchy of associativity. Hippocampus 2000; 10: 420–30.

Lerner Y, Hendler T, Ben-Bashat D, Harel M, Malach R. A hierarchical axis of object processing stages in the human visual cortex. Cereb Cortex 2001; 11: 287–97.

Lewis JW, Wightman FL, Brefczynski JA, Phinney RE, Binder JR, DeYoe EA. Human brain regions involved in recognizing environmental sounds. Cereb Cortex 2004; 14: 1008–21.

Macaluso E. Multisensory processing in sensory-specific cortical areas. Neuroscientist 2006; 12: 327–38.

McRae K, de Sa VR, Seidenberg MS. On the nature and scope of featural representations of word meaning. J Exp Psych Gen 1997; 126: 99–130.

McRae K, Cree GS. Factors underlying category-specific semantic deficits. In: Forde EME, Humphreys G, editors. Category-specificity in mind and brain. Sussex: Psychology Press; 2002. p. 211–50.

Molholm S, Ritter W, Javitt DC, Foxe JJ. Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. Cereb Cortex 2004; 14: 452–65.

Moss HE, Rodd JM, Stamatakis EA, Bright P, Tyler LK. Anteromedial temporal cortex supports fine-grained differentiation among objects. Cereb Cortex 2005; 15: 616–27.

Moss HE, Tyler LK, Devlin J. The emergence of category specific deficits in a distributed semantic system. In: Forde E, Humphreys GW, editors. Category-specificity in brain and mind. Sussex: Psychology Press; 2002. p. 115–48.

Murray EA, Bussey TJ. Perceptual-mnemonic functions of the perirhinal cortex. Trends Cogn Sci 1999; 3: 142–51.

Murray EA, Malkova L, Goulet S. Crossmodal associations, intramodal associations, and object identification in macaque monkeys. In: Milner AD, editor. Comparative neuropsychology. Oxford: Oxford University Press; 1998. p. 51–69.

Murray EA, Richmond BJ. Role of perirhinal cortex in object perception, memory, and associations. Curr Opin Neurobiol 2001; 11: 188–93.

Parker A, Gaffan D. Lesions of the primate rhinal cortex cause deficits in flavour-visual associative memory. Beh Brain Res 1998; 93: 99–105.

Pouget A, Deneve S, Duhamel J-R. A computational perspective on the neural basis of multisensory spatial representations. Nat Rev Neurosci 2002; 3: 741–7.

Powell HWR, Guye M, Parker GJM, Symms MR, Boulby P, Koepp MJ, et al. Noninvasive in vivo demonstration of the connections of the human parahippocampal gyrus. Neuroimage 2004; 22: 740–7.

Rauschecker JP, Tian B. Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc Natl Acad Sci USA 2000; 97: 11800–6.

Randall B, Moss HE, Rodd JM, Greer M, Tyler LK. Distinctiveness and correlation in conceptual structure: Behavioral and computational studies. J Exp Psychol Learn Mem Cogn 2004; 30: 393–406.

Rockland KS, Ojima H. Multisensory convergence in calcarine visual areas in macaque monkey. Int J Psychophysiol 2003; 50: 19–26.

Rogers TT, McClelland JL. Semantic cognition: a parallel distributed processing approach. Cambridge, MA: MIT Press; 2004.

Schroeder CE, Smiley J, Fu KG, McGinnis T, O'Connell MN, Hackett TA. Anatomical mechanisms and functional implications of multisensory convergence in early cortical processing. Int J Psychophysiol 2003; 50: 5–17.

Semple MN, Scott BH. Cortical mechanisms in hearing. Curr Opin Neurobiol 2003; 13: 167–73.

Shams L, Iwaki S, CHawla A, Bhattacharya J. Early modulation of visual cortex by sound: an MEG study. Neurosci Lett 2005; 378: 76–81.

Simmons WK, Barsalou LW. The similarity-in-topography principle: reconciling theories of conceptual deficits. Cognit Neuropsychol 2003; 20: 451–86.

Stamatakis EA, Tyler LK. Identifying lesions on structural brain images: validation of the method and application to neuropsychological patients. Brain Lang 2005; 94: 167–77.

Suzuki WA, Amaral DG. Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferents. J Comp Neurol 1994; 350: 497–533.

Suzuki WA, Amaral DG. Where are the perirhinal and parahippocampal cortices? A historical overview of the nomenclature and boundaries applied to the primate medial temporal lobe. Neuroscience 2003; 120: 893–906.

Talairach J, Tournoux P. Coplanar stereotactic atlas of the human brain. Stuttgart: George Thieme Verlag; 1988.

Taylor KI, Moss HE, Stamatakis EA, Tyler LK. Binding crossmodal object features in perirhinal cortex. Proc Natl Acad Sci USA 2006; 103: 8239–44.

Taylor KI, Moss HE, Tyler LK. The conceptual structure account: a cognitive model of semantic memory and its neural instantiation. In: Hart J, Kraut M, editors. The neural basis of semantic memory. Cambridge, UK: Cambridge University Press; 2007. p. 265–301.

Taylor KI, Probst A. Anatomic localization of the transentorhinal region of the perirhinal cortex. Neurobiol Aging 2008; 29: 1591–6.

Tranel D, Grabowski TJ, Lyon J, Damasio H. Naming the same entities from visual or from auditory stimulation engages similar regions of left inferotemporal cortices. J Cogn Neurosci 2005; 17: 1293–1305.

Tyler LK, deMornay-Davies P, Anokhina R, Longworth C, Randall B, Marslen-Wilson WD. Dissociations in processing past tense morphology: neuropathology and behavioral studies. J Cogn Neurosci 2002; 14: 79–94.

Tyler LK, Marslen-Wilson W. Fronto-temporal brain systems supporting spoken language comprehension. Philos Transact R Soc 2008; 363: 1037–54.

Tyler LK, Marslen-Wilson W, Stamatakis EA. Dissociating neuro-cognitive component processes: voxel based correlational methodology. Neuropsychologia 2005a; 43: 771–8.

Tyler LK, Marslen-Wilson WD, Stamatakis EA. Differentiating lexical form, meaning and structure in the neural language system. Proc Natl Acad Sci USA 2005b; 102: 8375–80.

Tyler LK, Moss HE. Towards a distributed account of conceptual knowledge. Trends Cogn Sci 2001; 5: 244–52.

Tyler LK, Moss HE, Durrant-Peatfield MR, Levy JP. Conceptual structure and the structure of concepts: a distributed account of category-specific deficits. Brain Lang 2000; 75: 195–231.

Tyler LK, Stamatakis EA, Bright P, Acres K, Abdallah S, Rodd JM, et al. Processing objects at different levels of specificity. J Cogn Neurosci 2004; 16: 351–62.

von Monakow C. Die Lokalisation im Grosshirn und der Abbau der Funktion durch kortikale Herde. Wiesbaden, Germany: JF Bergmann; 1914.

Wise R, Chollet F, Hadar U, Friston K, Hoffner E, Frackowiak R. Distribution of cortical neural networks involved in word comprehension and word retrieval. Brain 1991; 114: 1803–17.