Brain (2001), **124,** 804–815

The neural correlates of person familiarity
A functional magnetic resonance imaging study with clinical implications

Nadim J. Shah,[1] John C. Marshall,[2] Oliver Zafiris,[3] Anna Schwab,[4] Karl Zilles,[1,4] Hans J. Markowitsch[5] and Gereon R. Fink[1,6]

[1]*Institut für Medizin, Forschungszentrum Jülich, Jülich,* [2]*Neuropsychology Unit, University Department of Clinical Neurology, Radcliffe Infirmary, Oxford, UK,* [3]*Universität Koblenz/Landau, Abt. Landau, Landau,* [4]*C. & O. Vogt-Hirnforschungsinstitut, Heinrich-Heine-Universität Düsseldorf, Düsseldorf,* [5]*Physiologische Psychologie, Universität Bielefeld, Bielefeld and* [6]*Neurologische Klinik, RWTH Aachen, Aachen, Germany*

Correspondence to: Gereon R. Fink, MD, Institut für Medizin, Forschungszentrum Jülich, 52425 Jülich, Germany
E-mail: g.fink@fz-juelich.de

Summary
Neural activity was measured in 10 healthy volunteers by functional MRI while they viewed familiar and unfamiliar faces and listened to familiar and unfamiliar voices. The familiar faces and voices were those of people personally known to the subjects; they were not people who are more widely famous in the media. Changes in neural activity associated with stimulus modality irrespective of familiarity were observed in modules previously demonstrated to be activated by faces (fusiform gyrus bilaterally) and voices (superior temporal gyrus bilaterally). Irrespective of stimulus modality, familiarity of faces and voices (relative to unfamiliar faces and voices) was associated with increased neural activity in the posterior cingulate cortex, including the retrosplenial cortex. Our results suggest that recognizing a person involves information flow from modality-specific modules in the temporal cortex to the retrosplenial cortex. The latter area has recently been implicated in episodic memory and emotional salience, and now seems to be a key area involved in assessing the familiarity of a person. We propose that disturbances in the information flow described may underlie neurological and psychiatric disorders of the recognition of familiar faces, voices and persons (prosopagnosia, phonagnosia and Capgras delusion, respectively).

Keywords: voices, faces, prosopagnosia, phonagnosia, Capgras delusion

Abbreviations: BOLD = blood oxygenation level-dependent; FF = familiar faces; fMRI = functional MRI; FV = familiar voices; SPM = statistical parametric mapping; UF = unfamiliar faces; UV = unfamiliar voices

Introduction
Current models of how one recognizes a known person distinguish between unimodal and heteromodal perceptual mechanisms (Bruce and Young, 1986). Thus, domain-specific circuits for recognition by face or by voice are contrasted with domain-independent 'person identity nodes' that can be accessed from either facial or vocal information, or both (Bruce and Young, 1986). These identity-specific nodes (Bruce and Young, 1986; Ellis, 1986) provide access to biographical information that is stored in semantic and episodic memory (Markowitsch, 1995; Nyberg *et al.*, 1996; Schacter, 1996).

Functional imaging studies have shown that the recognition of famous faces (actors, politicians, sports figures and media celebrities) compared with unknown faces selectively activates the temporofrontal cortex (Sergent *et al.*, 1992; Gorno Tempini *et al.*, 1998; Leveroni *et al.*, 2000). It has been suggested that these regions are crucially involved in the 'stored knowledge of personal identity' (Gorno Tempini *et al.*, 1998) and 'long-term retrieval from the person identity system rather than face processing and face recognition in general' (Leveroni *et al.*, 2000). These findings are broadly consistent with the results of lesion studies (Ellis *et al.*, 1989; Kapur *et al.*, 1992). Anterior temporal lobe damage can be associated with impaired semantic memory for 'singular

objects' (Ellis *et al.*, 1989), including famous faces, voices, animals (e.g. Lassie), buildings (e.g. the Taj Mahal) and product names. Nonetheless, it is reported that these patients with anterior temporal cortex lesions have 'an ''emotional'' familiarity recognition' of faces they cannot identify (Kapur *et al.*, 1992) and can make relatively accurate familiarity ratings between well-known, lesser-known and unknown faces (Ellis *et al.*, 1989). 'Familiarity checking' (Ellis, 1983) and 'person identity' information (Bruce and Young, 1986) must accordingly draw, in part at least, upon different neuroanatomical substrates.

To our knowledge, no prior neuroimaging study has investigated face and voice recognition in the same experiment. This is essential if, in addition to isolating brain regions for face and voice recognition, we are to investigate the neural substrates of familiarity checking (Ellis, 1983) and person identity nodes (Bruce and Young, 1986), defined here as those brain regions that respond to familiarity or identity *per se*, irrespective of the modality of stimulus presentation (visual or vocal). We therefore designed a factorial functional imaging study in which healthy volunteers were presented with both familiar and unfamiliar faces and voices. For the familiar stimuli, we chose the faces and voices of people personally known to the subjects. It is most unlikely that famous but not personally known faces or voices (as used in many previous studies) will evoke the emotional reaction characteristic of seeing or hearing friends and relatives.

Accordingly, personal familiarity in the sense investigated here may evoke activation of the limbic system rather than the anterior temporal cortex (Henson *et al.*, 2000). The participants were instructed to try to recognize the faces shown or the voices heard, but they did not make an overt response to these stimuli. Rather, random trials of a subsidiary task were inserted. The subsidiary task involved responding to the onset of a chequerboard probe during the presentation of faces or the onset of white noise during the presentation of voices. These responses were required to ensure that the volunteers were attending to the series of stimuli that were of primary interest. Because stimuli were blocked into familiar and unfamiliar sets, we did not want the volunteers to cease attending to the blocks of unfamiliar stimuli (George *et al.*, 1999). The familiar stimuli were not drawn from famous people but rather from friends and relatives of the participating volunteers. The unfamiliar stimuli were drawn from a similar population in a different town. It may be that the combination of passive viewing and listening (rather than explicit discrimination or recognition) with the use of familiar (not famous) faces and voices will place more emphasis upon person familiarity rather than person identity (and the stored information about known people).

Material and methods

Subjects

Ten healthy, right-handed male volunteers (aged 25–33 years; mean age 28.5 years, SD 2.75) with no history of neurological or psychiatric illness were studied. We employed only male volunteers in order to avoid the normal variation in brain size and shape between the sexes and thus to improve image normalization (see Image processing). The study was approved by the ethics committee of the University Hospital of Heinrich Heine University, Düsseldorf. Informed consent was obtained from the participants before the study began.

Experimental design

The critical conditions of this factorial design were conditions 1 and 2 (factor 1: familiarity), in which subjects were presented with familiar faces (FF, condition 1) or with familiar voices (FV, condition 2). In conditions 3 and 4, subjects were presented with unfamiliar faces (UF, condition 3) or with unfamiliar voices (condition 4, UV). Comparing conditions 1 and 2 (familiar faces and voices) with conditions 3 and 4 (unfamiliar faces and voices) should reveal the neural correlates underlying person familiarity. Factor 2 (stimulus modality: visual in conditions 1 and 3, auditory in conditions 2 and 4) was introduced for the following reasons. First, the introduction of stimulus modality as a factor enables the identification of areas that respond to personally familiar stimuli regardless of modality. Secondly, factor 2 served as an internal quality control for the study. The direct comparison of all visual conditions versus all auditory conditions (and vice versa) should show activations in the occipital and temporal cortices, respectively. Finally, the interaction term should assess putative differential effects of stimulus modality on the neural activity resulting from person familiarity. Conditions 1–4 were presented in counterbalanced order for stimulus modality and familiarity within and between subjects in order to avoid potential confounds between stimulus conditions and the subjects' alertness.

Preparation of stimuli

For familiar faces, photographs of friends and relatives of each individual studied were taken with a single-lens reflex camera and then digitized with a scanner and PC. The digital images were manipulated to ensure that face size and luminance were approximately similar in each image. The backgrounds were edited out and a white fixation cross was inserted just above the bridge of the nose. In order to ensure that unfamiliar faces were indeed unfamiliar and had not been seen by the subjects by chance, pictures were taken of volunteers recruited from outside the university town where the subjects for this study were resident. The male : female ratio of familiar faces ranged from 3 : 9 to 8 : 4 depending on the volunteer being studied; the ratio for unfamiliar faces was held constant at 6 : 6 for all volunteers.

Voices were recorded using a DAT (digital audio tape) recorder. For familiar voices, voices of the same friends and relatives of each individual studied were recorded. For unfamiliar voices, voice recordings were taken of the volunteers recruited from outside the university town where

the subjects for this study were resident. The mouth-to-microphone distance and the recording room were identical for all stimuli. Stimuli were then transferred to a PC. Five sequences of German phrases, such as (in translation), 'also right here', 'and his interest' and 'up and down', were recorded. The onset, duration and intensity of the stimuli were edited and synchronized by means of a speech editor. The criterion for temporal alignment of the syllables was the onset of articulatory release. Individual phrases were faded in and out to remove clicks at the beginning and end of digital-to-analogue conversion. Sequence duration was 1–1.1 s. Auditory stimuli were presented binaurally using a digital playback system consisting of a PC, an audio amplifier and a magnetically shielded transducer system. The acoustic stimulation system terminated in tightly occlusive headphones that allowed unimpeded conduction of the stimulus with good suppression of ambient scanner noise by ~20 dB.

Tasks

In all conditions except the baseline (rest without stimulus presentation), subjects viewed faces or listened to voices. In order to ensure that subjects maintained attention, subsidiary tasks were interleaved between the stimuli of primary interest. During the presentation of faces, subjects were required to indicate the onset of a chequerboard pattern whilst they fixated on the central cross. During the voice conditions, subjects indicated the onset of a short period of white noise. The occurrence of both the chequerboard and the white noise was random with respect to the other stimuli being presented. Subjects were asked to press a button with their right index finger to indicate detection of either the chequerboard or the white noise. Reaction times of the button presses (measured from stimulus onset) and error rates were recorded as measures of task difficulty and task performance.

MRI acquisition

The entire brain was scanned in 3D to obtain anatomical images, using a high-resolution, T_1-weighted MP-RAGE (magnetization-prepared, rapid acquisition gradient echo) pulse sequence with the following parameters: TR (repetition time) = 11.4 ms; TE (echo time) = 4.4 ms; θ (flip angle) = 15°; one excitation; FOV (field of view) = 230 mm; matrix = 200×256; 128 sagittal slices with 1.41 mm slice thickness. Echo planar imaging was performed with a Magnetom Vision 1.5 tesla scanner (Siemens Medical Systems, Erlangen, Germany) equipped with a gradient booster system; the standard radiofrequency head coil was used for transmission of radiofrequency and signal reception. Pulse sequence parameters were as follows: gradient echo echo planar imaging; TE = 66 ms; TR = 7 s; FOV = 200×200 mm; θ = 90°; matrix size = 64×64; pixel size = 3.125×3.125 mm; slice thickness = 3.0 mm; interslice gap = 0.3 mm; 30 slices. The blocked design functional MRI (fMRI) paradigm, which was preceded by three dummy scans to allow the magnetic resonance signal to reach a steady state, comprised four repetitions of a 42 s baseline state of rest without stimulus presentation (6 TR) and an 84 s activation state (12 TR). Thus, during the activation state, 12 whole-brain images were acquired and during the baseline state six whole-brain data sets were acquired. A blocked design was adopted in preference to an event-related approach in order to maximize efficiency, i.e. to maximize the variance that the explanatory factors accounted for (Friston *et al.*, 1999). Each volunteer performed five repetitions of the experiment during the complete measurement procedure. In conditions 1 and 3, the stimuli (FF and UF, respectively) were presented in the centre of a screen (the whole screen represented a horizontal visual angle of ~60° and a vertical visual angle of ~31°). After image acquisition in each TR period, three visual stimuli were presented on a white background during the fMRI measurements for 1150 ms using MEL^{TM} (MEL Professional; Psychology Software Tools, Pittsburgh, Pa., USA). Using a mirror, subjects viewed the display from a distance of 25 cm (screen to mirror 14 cm, mirror to subject's eyes 11 cm). For conditions 2 and 4, the auditory stimuli were presented as described above.

Image processing

All calculations and image manipulations were performed on SPARC Ultra 1 workstations (Sun Microsystems Computers) using MATLAB (Mathworks, Natick, Mass, USA) and SPM97d/SPM99 [Statistical Parametric Mapping (SPM) software; Wellcome Department of Cognitive Neurology, London, UK; http:www.fil.ion.ucl.ac.uk]. SPM was used for image realignment, image normalization and smoothing and to create statistical maps of significant relative changes in regional blood oxygenation level-dependent (BOLD) responses (Friston *et al.*, 1995*a*, *b*).

The first three images of each time series, during which the MRI signal reached a steady state, were discarded. The remaining 64 volume images of each time series were realigned automatically to the first image (corresponding to the fourth acquired image of the time series) to correct for head movement between scans. Image sets of the four conditions and the baseline were then co-registered to the 3D anatomical data set using MPItool^{TM} (MPItool, MPIfnF; Max–Planck–Institute for Neurological Research, Cologne, Germany), their anterior and posterior commissure points were identified, and the sets were transformed into a standard stereotaxic space (Talairach and Tournoux, 1988), using the intercommissural line as the reference plane for the transformation. This spatial transformation uses linear proportions and a non-linear sampling algorithm (Friston *et al.*, 1995*a*). Data were expressed thereafter in terms of standard stereotaxic coordinates in the *x*-, *y*- and *z*-axes (as defined in Table 1). The resulting pixel size in standard stereotaxic space was 4×4 mm with an interplane distance of 4 mm. Transformed functional data sets from each subject were smoothed, with a Gaussian kernel of 6 mm for single-

Table 1 *Relative increases in brain activity during performance of the various conditions*

Region	Side	Stereotaxic coordinates			*Z*-score
		x	*y*	*z*	
Familiar faces + unfamiliar faces > familiar voices + unfamiliar voices					
Fusiform gyrus	R	+42	−62	−22	9.3
	L	−38	−60	−22	9.1
Familiar voices + unfamiliar voices > familiar faces + unfamiliar faces					
Superior temporal gyrus	R	+64	−8	−6	9.8
	L	−62	−10	−4	9.7
Heschl's gyrus	R	+44	−28	+10	9.5
	L	−42	−32	+8	9.5
Inferior frontal gyrus	R	+60	+16	+16	7.0
	L	−50	+14	+16	7.4
Familiar voices + familiar faces > unfamiliar voices + unfamiliar faces					
Retrosplenial cortex	R	+2	−54	+16	4.9
	L	−4	−58	+4	4.5

Coordinates are in standard stereotaxic space and refer to maximally activated foci as indicated by the highest *Z*-score within an area of activation associated with the contrasts described. *x* is the distance in millimetres to the right (+) or left (–) of the midsagittal (interhemispheric) line; *y* is the distance anterior (+) or posterior (–) to the vertical plane through the anterior commissure; and *z* is the distance above (+) or below (–) the intercommissural line. This is based on the stereotaxic atlas and the group mean MRI. R = right, L = left.

subject analysis and 10 mm (full-width half-maximum) for the group analysis, to meet the statistical requirements of the general linear model and to compensate for normal variation in brain size, shape and sulcal/gyral anatomy among subjects. Voxels that had values greater than 0.8 of the volume mean in all the images were selected in order to restrict analysis to intracranial regions.

Statistical analysis

Statistical analysis was performed after stereotaxic normalization and smoothing of the data. Low-frequency cosine waves were used to model and remove subject-specific low-frequency drifts in signal.

For the comparison of global (whole brain) signals across all experimental conditions, global BOLD signals were estimated. This allowed us to check for differences in general arousal (overall alertness) associated with the four experimental conditions.

Thereafter, global means were normalized by proportional scaling. Data were analysed by modelling the different conditions (visual: familiar, unfamiliar; auditory: familiar, unfamiliar) as reference waveforms (i.e. boxcar functions convolved with a haemodynamic response function in the context of the general linear model employed by SPM97d). We accordingly defined a design matrix comprising contrasts modelling the alternating periods of 'baseline' and 'activation' by the use of a delayed boxcar reference vector that accounted for the delayed cerebral haemodynamic response function after stimulus presentation. Specific effects were tested by applying appropriate linear contrasts to the parameter estimates for each condition, resulting in a *t*-statistic for each voxel. These *t*-statistics (transformed to *Z*-statistics) constituted a statistical parametric map (SPM{*Z*}). The SPM{*Z*} were then interpreted by referring to the probabilistic behaviour of Gaussian random fields. Voxels were identified as significant only if they passed a height threshold of $Z = 4.5$ ($P < 0.05$, corrected for multiple comparisons) and belonged to a cluster of at least 30 activated voxels (Friston *et al.*, 1995*b*).

The data were first analysed for the main effects of stimulus modality (i.e. FF + UF > FV + UV and FV + UV > FF + UF) and familiarity (i.e. FF + FV > UF + UV and UF + UV > FF + FV). Secondly, the data were analysed for the interaction of stimulus modality with familiarity.

For the fMRI data group analysis, repeated measures (scans) were collapsed within subject after having adjusted for both global blood flow, by using proportional scaling, and for low-frequency physiological drifts, by using a high-pass filter; this gave one scan per condition per subject. These conditions were then compared between subjects, thereby effecting a random effects model, allowing inference to the general population. In addition, individual fMRI data were analysed to aid the anatomical localization of significant activations detected in the group analysis.

For individual subject analysis, fMRI data were analysed in an identical way to the group data but without collapsing the individual scans. The SPM was now thresholded at $P < 0.05$, uncorrected, as this analysis was restricted to areas showing significant activations in the group analysis.

Localization of activations

The stereotaxic coordinates of the pixels of the local maximum significant activation were determined within areas of significant relative activity change associated with the

tasks. The anatomical localization of these local maxima was assessed with reference to the standard stereotaxic atlas (Talairach and Tournoux, 1988) and validation of this method of localization was obtained by superimposition of the SPM maps on the group mean MRI calculated after each individual's MRI had been stereotaxically transformed into the same standard stereotaxic space (Friston *et al.*, 1995*a*).

Results

Global brain activity associated with each of the four experimental conditions

There were no significant differences in the global BOLD signal (in arbitrary units), either in terms of the mean or the standard error of the mean, associated with the four experimental conditions: familiar faces (80.1 ± 0.2), familiar voices (80.1 ± 0.1), unfamiliar faces (80.0 ± 0.2) and unfamiliar voices (80.1 ± 0.2). This suggests that there were no differences in general arousal (overall alertness) across conditions.

Brain activity associated with faces and voices

As expected, faces relative to voices (i.e. FF + UF > FV + UV) led to significant ($P < 0.05$, corrected for multiple comparisons) increases in neural activation centred on the fusiform gyrus bilaterally (Fig. 1A), extending into the striate and extrastriate cortex bilaterally. The magnitude of these effects is depicted in Fig. 2A and B for the local maximum of activity increase in the left and right fusiform gyrus, respectively. By contrast, voices relative to faces (i.e. FV + UV > FF + UF) led to increased ($P < 0.05$, corrected) neural activity centred on the superior temporal gyrus bilaterally, extending into the primary (Heschl's gyrus) and secondary auditory cortex bilaterally (Fig. 1B). The magnitude of these effects is depicted in Fig. 2C and D for the local maximum in the left and right superior temporal gyrus, respectively. In addition, the inferior frontal gyrus was activated bilaterally, probably corresponding to Broca's area and its homologous region in the right hemisphere.

Brain activity associated with personal familiarity

Increased ($P < 0.05$, corrected) neural activity associated with personal familiarity, irrespective of the stimulus modality (i.e. FF + FV > UF + UV), was observed bilaterally in the posterior cingulate cortex (Braak, 1979; Vogt *et al.*, 1995), including only the retrosplenial cortex (Brodmann areas 26, 29 and 30) (Fig. 3). The magnitude of these effects is depicted in Fig. 2E and F for the local maximum in the left and right retrosplenial cortex, respectively.

Activations associated with personal familiarity that did not meet our strict criterion were nonetheless observed at the lower threshold of $P < 0.01$, uncorrected, in the following regions: left inferior frontal gyrus, right premotor cortex, left and right posterior superior temporal gyrus, anterior cingulate and supplementary motor area bilaterally, left insula, right orbitofrontal cortex and right amygdala.

The reverse comparison (i.e. UF + UV > FF + FV) showed no significant activation ($P < 0.05$, corrected) associated with unfamiliarity. No significant interactions between stimulus type (faces, voices) and familiarity (familiar, unfamiliar) were observed. Despite the lack of a significant interaction, there was a trend for the retrosplenial cortex to be more activated by familiar faces than by familiar voices (Fig. 2E and F).

Single-subject analysis

All 10 subjects showed a pattern of neural activations consistent with the group results for faces and voices, i.e. all subjects showed activations in the relevant fusiform and superior temporal gyrus/Heschl's gyrus associated with faces and voices, respectively ($P < 0.05$, uncorrected or better). For familiarity, eight of the 10 subjects showed increased neural activity in the posterior cingulate cortex, including the retrosplenial cortex ($P < 0.05$, uncorrected or better). Figure 4 illustrates the activations in the posterior cingulate cortex for the eight subjects who did show the effect and the two subjects who did not show the effect.

Behavioural results

There were no significant differences in reaction times to the chequerboard or noise probes inserted across the face and voice conditions (UF, 506 ± 100 ms; FF, 496 ± 101 ms; UV, 483 ± 140 ms; FV, 511 ± 148 ms). This again suggests that there were no significant arousal or attentional differences across the experimental conditions.

When the subjects were debriefed after the fMRI study with the same stimuli, the subjects reported that they could identify 100% of the familiar faces and 93% of the familiar voices. Subjects reported 97% of the supposedly unfamiliar faces as indeed unfamiliar and 96% of the supposedly unfamiliar voices as unfamiliar.

Discussion

Functional imaging of the normal brain has reliably implicated the fusiform gyrus, more strongly perhaps in the right than the left hemisphere (Sergent and Signoret, 1992; Dolan *et al.*, 1997; Kanwisher *et al.*, 1997), as a crucial node in the neural circuit that underlies face recognition. Lesion of these areas reliably results in prosopagnosia, the inability to recognize familiar and famous faces (Damasio *et al.*, 1982). The association is sufficiently robust that the designation 'fusiform face area' has become established within the neurosciences. Not all imaging studies, however, have distinguished between the role that the fusiform gyrus plays in recognizing that a face is a face as opposed to identifying whose face it is. Functional imaging, according to a recent paper (George

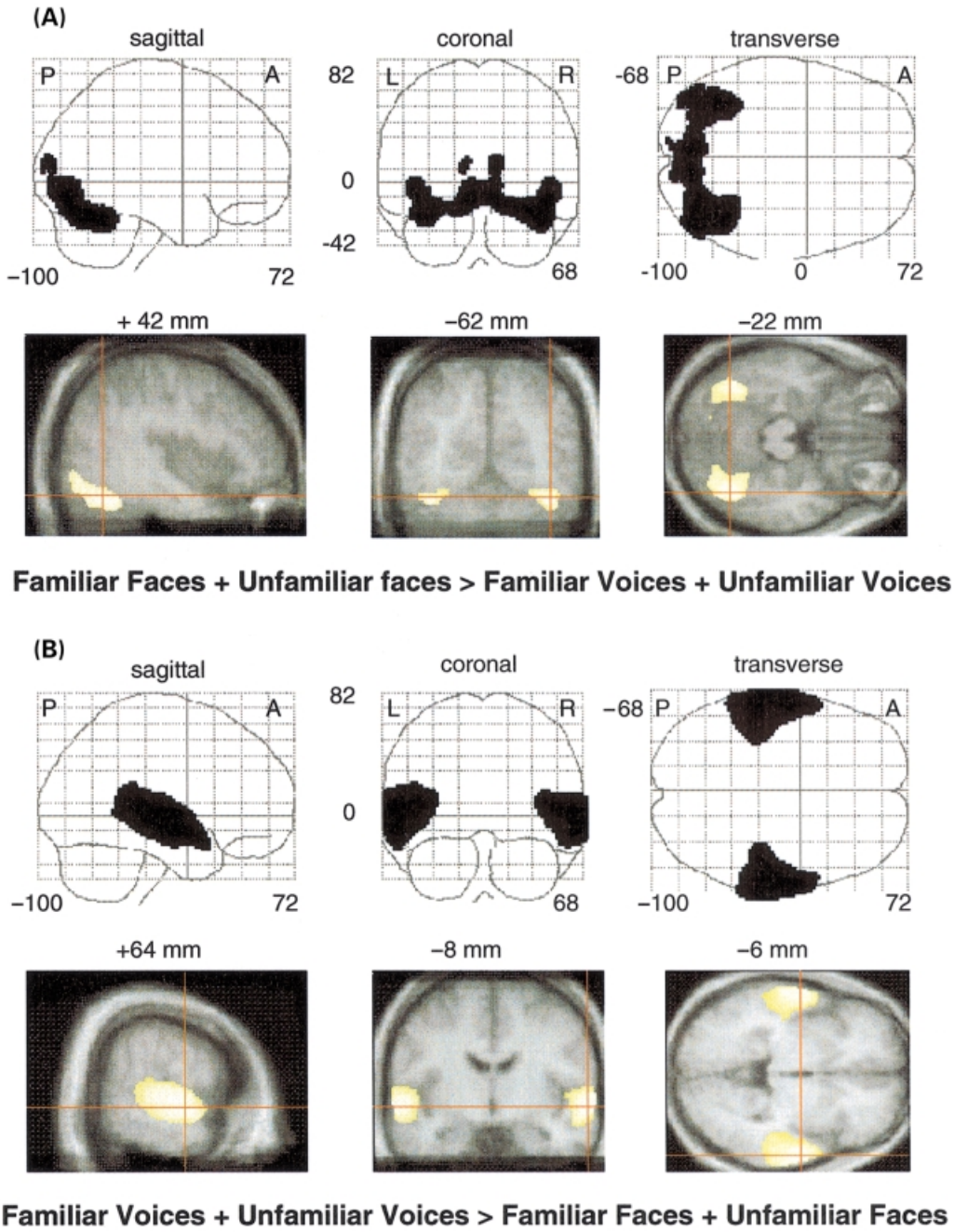


Fig. 1 Relative increases in neural activity (for the 10 subjects) associated with (**A**) faces and (**B**) voices, irrespective of familiarity. Areas of significant relative increase in neural activity are shown as through-projections onto representations of standard stereotaxic space. Sagittal, side view; coronal, view from the back; transverse, view from above. The local maxima of the areas of significant relative increase ($P < 0.05$, corrected for multiple comparisons) in neural activity are displayed superimposed on transverse MRI sections to detail the functional anatomy of the activations and their relationship to underlying structural anatomy. For these illustrative purposes, only activations with a *Z*-score of >8.5 are displayed. There is bilateral fusiform activation (**A**) associated with faces and bilateral superior temporal gyrus activation (**B**) associated with voices. The exact coordinates of the local maxima within the areas of activation and their *Z* statistics are given in Table 1. R = right; L = left; A = anterior; P = posterior.

et al., 1999), 'has revealed face-responsive visual areas in the human fusiform gyrus, but their role in recognizing familiar individuals remains controversial'. Recently, however, it has been shown that famous faces do activate the right fusiform gyrus to a greater extent than do newly learned faces (Leveroni *et al.*, 2000), and Gauthier and colleagues (Gauthier *et al.*, 2000) have suggested that this region is concerned with distinguishing between faces.

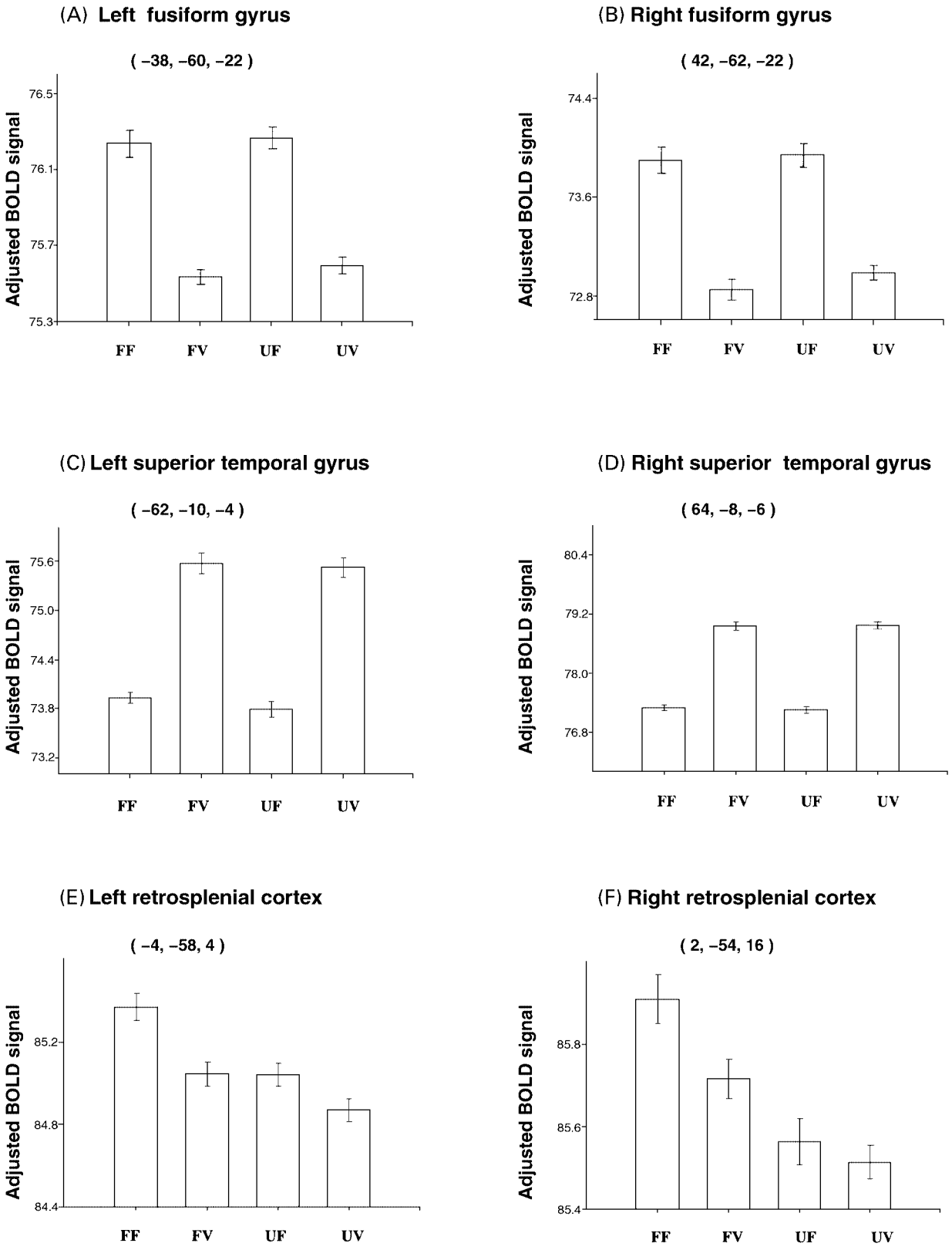


Fig. 2 Local brain activity (for the 10 subjects) as a function of the experimental conditions (FF = familiar faces; FV = familiar voices; UF = unfamiliar faces; UV = unfamiliar voices) in the left (**A**) and right (**B**) fusiform gyrus, the left (**C**) and right (**D**) superior temporal gyrus, and the left (**E**) and right (**F**) retrosplenial cortex. The coordinates for the local maxima within the regions of significant increase associated with the relevant comparisons (Table 1) are given for each of the regions. The histograms show the adjusted local BOLD signal (in arbitrary units, mean ± standard error of the mean) as a function of the experimental conditions.

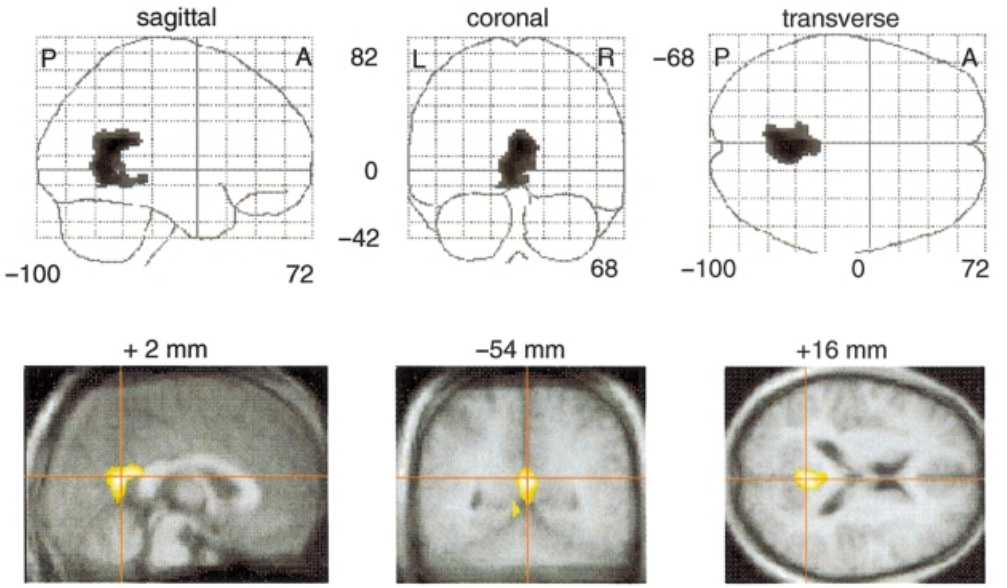
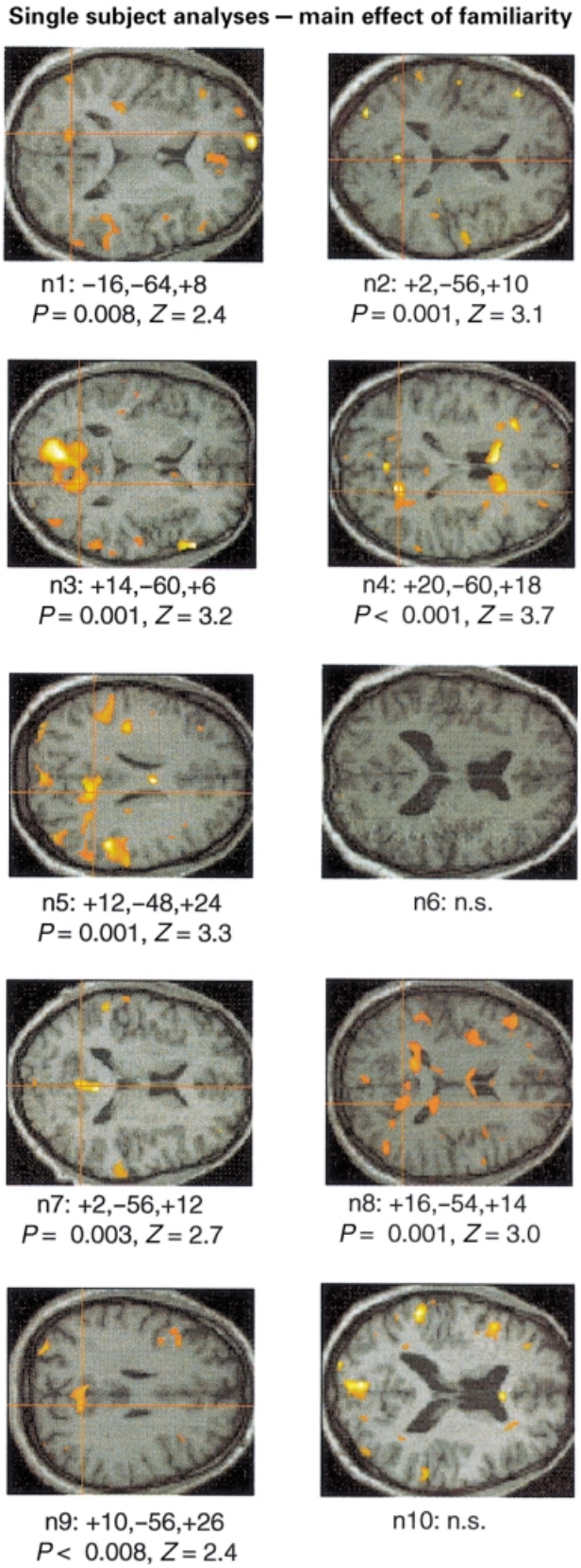


Fig. 3 Relative increases in neural activity (for the 10 subjects) associated with familiarity (irrespective of whether faces or voices were used as stimuli). Areas of significant relative increase ($P < 0.05$, corrected) in neural activity are shown as through-projections onto representations of standard stereotaxic space. The local maximum of the area of significant relative increase in neural activity is displayed superimposed on MRI sections to detail the functional anatomy of the activation and its relationship to underlying structural anatomy. There is bilateral activation of the retrosplenial cortex. The exact coordinates of the local maxima within the areas of activation and their *Z* scores are given in Table 1. All abbreviations as in Fig. 1.

With respect to voice recognition, lesion studies have suggested that discriminating between unfamiliar voices is impaired by damage to the temporal lobes of either hemisphere, while the recognition of famous voices (e.g. politicians and entertainers whose voices are known from radio and television) is impaired by right parietal damage (van Lancker *et al.*, 1989). Both of these impairments have been referred to as phonagnosia (van Lancker *et al.*, 1989), but it is clear that they are two separable functions, each of which can be selectively impaired (van Lancker and Kreiman, 1987). Neuroimaging studies in healthy volunteers have demonstrated that distinguishing voices from other auditory stimuli induces bilateral activation along the upper bank of the superior temporal sulcus (Belin *et al.*, 2000). By contrast, recognizing and identifying particular speakers induces bilateral activation of the temporal poles (Imaizumi *et al.*, 1997), a result that is not immediately compatible with those of lesion studies (van Lancker and Canter, 1982; van Lancker *et al.*, 1989). It should be remembered, however, that in the experiment of Imaizumi and colleagues (Imaizumi *et al.*, 1997), subjects became familiar with the stimulus voices during the course of the study, while in the experiments of van Lancker and her colleagues the familiar voices were those of people well-known in politics and entertainment (van Lancker *et al.*, 1988).

Our present results are compatible with previous neuroimaging studies. We have confirmed the role of the fusiform gyrus in face perception (Damasio *et al.*, 1982; Sergent and Signoret, 1992; Dolan *et al.*, 1997; Kanwisher *et al.*, 1997) and of the superior temporal gyri and Heschl's gyri in voice perception (Imaizumi *et al.*, 1997; Belin *et al.*, 2000; Binder *et al.*, 2000). Activation of Broca's area (and its right hemisphere homologue) suggests that listeners may have been 'shadowing' the voices as an aid to assessing speaker familiarity (or identity). We did not find the right parietal activations that would be predicted from lesion studies of impaired recognition of familiar voices (phonagnosia) (van Lancker *et al.*, 1989), which is consistent with an earlier experiment on voice perception using PET (Imaizumi *et al.*, 1997).

More importantly, we have shown for the first time a multimodal region of activation in the limbic association cortex that responds to the familiarity of both faces and voices. We were careful to counterbalance all four conditions of the experiment (see Experimental design). It is therefore unlikely that the retrosplenial cortex was activated as a result of priming effects that could have occurred if all familiar faces had been presented before their equivalent voices, or vice versa (George *et al.*, 1999; Henson *et al.*, 2000). The present experiment gives no evidence about whether this region is specific to familiar faces and voices. It may be that a similar activation would be seen if familiar scenes or familiar (non-verbal) sounds were contrasted with unfamiliar sounds (Nakamura *et al.*, 2000). Nonetheless, this region—



the retrosplenial cortex—is, we suggest, part of the neural circuit concerned with familiarity checking (and perhaps with identifying people) originally postulated in psychological studies of face recognition (Ellis, 1983; Young *et al.*, 1986). That our volunteers judged that they could identify 100 and 93% of the familiar faces and voices, respectively, suggests that the retrosplenial cortex may be involved in the recognition of both familiarity and identity. Our results are also compatible with the fact that the amnesic syndromes seen after retrosplenial lesions (Gainotti *et al.*, 1998) have been reported to include impaired recognition memory for faces (Valenstein *et al.*, 1987; Rudge and Warrington, 1991).

Our experiment cannot distinguish conclusively between familiarity checking and person identification. Indeed, it would be very difficult to dissociate the two in normal individuals. One can, of course, have familiarity without identification, but one cannot, without pathology, have identification without familiarity. Nevertheless, the fact that the retrosplenial cortex, rather than the anterior temporal cortex (which showed no increase in neural activity even at a lowered statistical threshold of $P < 0.01$, uncorrected) is significantly activated suggests that the checking of familiarity played a larger role in determining our results than did person identification and the associated 'retrieval of information from the person identity semantic system' (Leveroni *et al.*, 2000). The area activated in our study in association with personal familiarity, i.e. the posterior cingulate/retrosplenial cortex, has anatomical connections with the perirhinal, parahippocampal and entorhinal cortex (Vogt and Gabriel, 1993; Kobayashi and Amaral, 2000). Thus, the hippocampal formation and the medial temporal lobe memory systems are closely associated with the retrosplenial cortex. These dense connections to areas implicated in memory and emotion may be relevant to activations previously observed with familiar stimuli and to the activations observed in the present study that failed to reach our strict criteria for significance. Prior studies of famous faces have found activation of the posterior cingulate cortex (Gorno Tempini *et al.*, 1998; Leveroni *et al.*, 2000), but have not explicitly linked this region with the familiarity checking operation (Ellis, 1983), which plays a crucial role in cognitive models of face recognition. The observed trend for an interaction of stimulus type with familiarity (augmentation of retrosplenial cortex activation due to familiarity when faces were shown) might indicate that the sight of familiar faces leads to a greater emotional reaction than the sound of familiar voices.

Fig. 4 Relative individual increases in neural activity associated with familiarity (irrespective of whether faces or voices were used as stimuli) for the eight of the 10 subjects who showed activation in the retrospinal cortex and for the remaining two subjects, who did not show significant activation in this area. The relevant areas of significant relative increase ($P < 0.05$, uncorrected) in neural activity are shown superimposed on the corresponding transverse MRI section of the same subject after co-registration of the images and normalization into the same stereotaxic space.

The evolutionary significance of assessing facial and vocal familiarity (without necessarily accessing biographic information about the person) makes it reasonable to expect that an allocortical area (such as the retrosplenial cortex) might be associated with this function in an organism that relies heavily on sight and sound. That both the global BOLD signal measures and the reaction time data (see Results) showed no significant differences between the four experimental conditions argues strongly against the possibility that this effect in the retrosplenial cortex might reflect changes in general alertness or arousal. Nonetheless, it must be stressed that the faces and voices we used were those of friends and relatives of the subjects. That is, the stimuli involved individuals who were closely associated with the subjects in the study and hence likely to evoke emotional responses, as indicated, for example, by autonomic reactivity. Damasio has proposed recently that the posterior cingulate is implicated in what he calls second-order mapping of bodily states (Damasio, 1999). These mappings are conscious 're-representations' of primary perceptual information and may be associated with the changes in autonomic state induced by the familiar faces and voices of friends and relatives.

An intact familiarity check (Ellis, 1983) of this nature may be responsible for the dissociation observed between covert and overt face recognition in neurological patients with prosopagnosia. Some patients who cannot explicitly identify the faces of familiar people nonetheless give indirect evidence of recognition. They show, for example, larger electrodermal skin conductance responses, characteristic of emotional arousal, to familiar than to unknown faces (Tranel and Damasio, 1985). Likewise, the patients may learn correct pairings of names and faces faster than incorrect pairings, despite not knowing explicitly the identity of the faces (de Haan *et al.*, 1987). Distinct neurological routes are known to mediate the overt recognition of familiar faces and emotional reactions to these faces (Bauer, 1984; Tranel and Damasio, 1985). The fusiform lesions that result in prosopagnosia (failure of overt identification of familiar or famous faces) can thus leave intact the recognition of a face's emotional significance to the viewer. A dorsal route connecting the visual association cortex, the inferior parietal cortex and the cingulate gyrus (Bauer, 1984; Ellis and Young, 1990) has been conjectured to be responsible for the 'autonomic-discriminatory responses' (Bauer, 1984) that are implicated in this preserved ability. Although the latter claim is controversial (Tranel *et al.*, 1995; Breen *et al.*, 2000), the fact that the posterior retrosplenial cingulate cortex is reliably activated by 'emotionally salient stimuli' (Maddock, 1999) is consistent with the role we propose it plays in judging familiarity and perhaps in accessing information about familiar people irrespective of stimulus modality (Fink *et al.*, 1996; Leveroni *et al.*, 2000).

The retrosplenial cortex may also be implicated in the Capgras delusion, a condition in which the patient believes that a close relative, usually a spouse, has been replaced by an imposter. The delusion is thought to result (in part) from damage to the neuroanatomical pathways concerned with appropriate emotional responses to the face [or sometimes the voice (Rojo *et al.*, 1991; Reid *et al.*, 1993)] of familiar people. Patients with the Capgras delusion do not show enhanced electrodermal skin conductance responses to familiar faces (Ellis *et al.*, 1997; Hirstein and Ramachandran, 1997), although they may show covert face recognition, as assessed by priming tasks (Ellis *et al.*, 2000). Despite the fact that the Capgras delusion is most frequently seen in psychiatric patients without obvious structural pathology, there is some evidence that the condition is also associated with lesions that disrupt information flow between the inferior temporal lobes and the medial limbic circuit (Signer, 1994), a disconnection that had previously been associated with visual hypoemotionality (Bauer, 1982). Lesions of a neuroanatomical circuit that includes the retrosplenial cortex may thus underlie the failure of patients with the Capgras delusion to respond correctly to familiar people by disabling appropriate interactions among face or voice recognition, memory and emotion (Fink *et al.*, 1996; Maddock, 1999).

Acknowledgements

We wish to thank M. L. Grosse-Ruyken and B. Elghawagi for technical assistance. J.C.M. is supported by the Medical Research Council, UK. K.Z., H.J.M. and G.R.F. are supported by the Deutsche Forschungsgemeinschaft, Germany. Further support from the Volkswagen-Stiftung is gratefully acknowledged.

References

Bauer RM. Visual hypoemotionality as a symptom of visual–limbic disconnection in man. Arch Neurol 1982; 39: 702–8.

Bauer RM. Autonomic recognition of names and faces in prosopagnosia: a neuropsychological application of the Guilty Knowledge Test. Neuropsychologia 1984; 22: 457–69.

Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex. Nature 2000; 403: 309–12.

Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Springer JA, Kaufman JN, et al. Human temporal lobe activation by speech and non-speech sounds. Cereb Cortex 2000; 10: 512–28.

Braak H. Pigment architecture of the human telencephalic cortex. IV. Regio retrosplenialis. Cell Tissue Res 1979; 204: 431–40.

Breen N, Caine D, Coltheart M. Models of face recognition and delusional misidentification: a critical review. Cogn Neuropsychol 2000; 17: 55–71.

Bruce V, Young A. Understanding face recognition. Br J Psychol 1986; 77: 305–27.

Damasio AR. The feeling of what happens: body and emotion in the making of consciousness. New York: Harcourt Brace; 1999.

Damasio AR, Damasio H, van Hoesen GW. Prosopagnosia: anatomic basis and behavioral mechanisms. Neurology 1982; 32: 331–41.

de Haan EHF, Young AW, Newcombe F. Face recognition without awareness. Cogn Neuropsychol 1987; 4: 385–415.

Dolan RJ, Fink GR, Rolls E, Booth M, Holmes A, Frackowiak RS, et al. How the brain learns to see objects and faces in an impoverished context. Nature 1997; 389: 596–9.

Ellis HD. The role of the right hemisphere in face perception. In: Young AW, editor. Functions of the right cerebral hemisphere. London: Academic Press; 1983. p. 33–44.

Ellis HD. Processes underlying face recognition. In: Bruyer R, editor. The neuropsychology of face perception and facial expression. Hillsdale (NJ): Lawrence Erlbaum; 1986. p. 1–27.

Ellis HD, Young AW. Accounting for delusional misidentifications. [Review]. Br J Psychiatry 1990; 157: 239–48.

Ellis AW, Young AW, Critchley EM. Loss of memory for people following temporal lobe damage. Brain 1989; 112: 1469–83.

Ellis HD, Young AW, Quayle AH, de Pauw KW. Reduced autonomic responses to faces in Capgras delusion. Proc R Soc Lond B Biol Sci 1997; 264: 1085–92.

Ellis HD, Lewis MB, Moselhy HF, Young AW. Automatic without autonomic responses to familiar faces: differential components of covert recognition in a case of Capgras delusion. Cogn Neuropsychiatry 2000; 5: 255–69.

Fink GR, Markowitsch HJ, Reinkemeier M, Bruckbauer T, Kessler J, Heiss W-D. Cerebral representation of one's own past: neural networks involved in autobiographical memory. J Neurosci 1996; 16: 4275–82.

Friston KJ, Ashburner J, Frith CD, Poline J-B, Heather JD, Frackowiak RSJ. Spatial registration and normalization of images. Hum Brain Mapp 1995a; 3: 165–89.

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 1995b; 2: 189–210.

Friston KJ, Zarahn E, Josephs O, Henson RN, Dale AM. Stochastic designs in event-related fMRI. Neuroimage 1999; 10: 607–19.

Gainotti G, Almonti S, DiBetta AM, Silveri MC. Retrograde amnesia in a patient with retrosplenial tumour. Neurocase 1998; 4: 519–26.

Gauthier I, Tarr MJ, Moylan J, Skudlarski P, Gore JC, Anderson AW. The fusiform 'face area' is part of a network that processes faces at the individual level. J Cogn Neurosci 2000; 12: 495–504.

George N, Dolan RJ, Fink GR, Baylis GC, Russell C, Driver J. Contrast polarity and face recognition in the human fusiform gyrus. Nat Neurosci 1999; 2: 574–80.

Gorno Tempini ML, Price CJ, Josephs O, Vandenberghe R, Cappa SF, Kapur N, et al. The neural systems sustaining face and proper-name processing. Brain 1998; 121: 2103–18.

Henson R, Shallice T, Dolan R. Neuroimaging evidence for dissociable forms of repetition priming. Science 2000; 287: 1269–72.

Hirstein W, Ramachandran VS. Capgras syndrome: a novel probe for understanding the neural representation of the identity and familiarity of persons. Proc R Soc Lond B Biol Sci 1997; 264: 437–44.

Imaizumi S, Mori K, Kiritani S, Kawashima R, Sugiura M, Fukuda H, et al. Vocal identification of speaker and emotion activates different brain regions. Neuroreport 1997; 8: 2809–12.

Kanwisher N, McDermott J, Chun MM. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci 1997; 17: 4302–11.

Kapur N, Ellison D, Smith MP, McLellan DL, Burrows EH. Focal retrograde amnesia following bilateral temporal lobe pathology. A neuropsychological and magnetic resonance study. Brain 1992; 115: 73–85.

Kobayashi Y, Amaral DG. Macaque monkey retrosplenial cortex: I. Three-dimensional and cytoarchitectonic organization. J Comp Neurol 2000; 426: 339–65.

Leveroni CL, Seidenberg M, Mayer AR, Mead LA, Binder JR, Rao SM. Neural systems underlying the recognition of familiar and newly learned faces. J Neurosci 2000; 20: 878–86.

Maddock RJ. The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. [Review]. Trends Neurosci 1999; 22: 310–6.

Markowitsch HJ. Which brain regions are critically involved in the retrieval of old episodic memory? [Review]. Brain Res Brain Res Rev 1995; 21: 117–27.

Nakamura K, Kawashima R, Sato N, Nakamura A, Sugiura M, Kato T, et al. Functional delineation of the human occipitotemporal areas related to face and scene processing. A PET study. Brain 2000; 123: 1903–12.

Nyberg L, McIntosh AR, Cabeza R, Nilsson LG, Houle S, Habib R, et al. Network analysis of positron emission tomography regional cerebral blood flow data: ensemble inhibition during episodic memory retrieval. J Neurosci 1996; 16: 3753–9.

Reid I, Young AW, Hellawell DJ. Voice recognition impairment in a blind Capgras patient. Behav Neurol 1993; 6: 225–8.

Rojo VJ, Caballero L, Iruela LM, Baca E. Capgras' syndrome in a blind patient [letter]. Am J Psychiatry 1991; 148: 1271–2.

Rudge P, Warrington EK. Selective impairment of memory and visual perception in splenial tumours. Brain 1991; 114: 349–60.

Schacter DL. Searching for memory: the brain, the mind, and the past. New York: Basic Books; 1996.

Sergent J, Signoret J-L. Functional and anatomical decomposition of face processing: evidence from prosopagnosia and PET study of normal subjects. Philos Trans R Soc Lond B Biol Sci 1992; 335: 55–62.

Sergent J, Ohta S, MacDonald B. Functional neuroanatomy of face and object processing. A positron emission tomography study. Brain 1992; 115: 15–36.

Signer SF. Localization and lateralization in the delusion of substitution. [Review]. Psychopathology 1994; 27: 168–76.

Talairach J, Tournoux P. Co-planar stereotaxic atlas of the human brain. Stuttgart: Thieme; 1988.

Tranel D, Damasio AR. Knowledge without awareness: an autonomic index of facial recognition by prosopagnosics. Science 1985; 228: 1453–4.

Tranel D, Damasio H, Damasio AR. Double dissociation between overt and covert face recognition. J Cogn Neurosci 1995; 7: 425–32.

Valenstein E, Bowers D, Verfaellie M, Heilman KM, Day AL, Watson RT. Retrosplenial amnesia. Brain 1987; 110: 1631–46.

van Lancker DR, Canter GJ. Impairment of voice and face recognition in patients with hemispheric damage. Brain Cogn 1982; 1: 185–95.

van Lancker DR, Kreiman J. Voice discrimination and recognition are separate abilities. Neuropsychologia 1987; 25: 829–34.

van Lancker DR, Cummings JL, Kreiman J, Dobkin BH. Phonagnosia: a dissociation between familiar and unfamiliar voices. Cortex 1988; 24: 195–209.

van Lancker DR, Kreiman J, Cummings J. Voice perception deficits: neuroanatomical correlates of phonagnosia. J Clin Exp Neuropsychol 1989; 11: 665–74.

Vogt BA, Gabriel M. Neurobiology of cingulate cortex and limbic thalamus. Boston: Birkhäuser; 1993.

Vogt BA, Nimchinsky EA, Vogt LJ, Hof PR. Human cingulate cortex: surface features, flat maps, and cytoarchitecture. J Comp Neurol 1995; 359: 490–506.

Young AW, McWeeny KH, Hay DC, Ellis AW. Access to identity-specific semantic codes from familiar faces. Q J Exp Psychol 1986; 38A: 271–95.

Received August 30, 2000. Revised October 26, 2000.
Accepted December 11, 2000