ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 24 (2005) 224–234

Functional topography of working memory for face or voice identity

Pia Rämä[a,b,c,*] and Susan M. Courtney[a,d,e]

[a]*Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD 21218, USA*
[b]*Department of Psychology, Cognitive Brain Research Unit, University of Helsinki, Helsinki, Finland*
[c]*Helsinki Brain Research Center, Helsinki, Finland*
[d]*Department of Neuroscience, Johns Hopkins University, Baltimore, MD 21205, USA*
[e]*F.M. Kirby Research Center for Functional Brain Imaging, Kennedy Krieger Institute, Baltimore, MD 21205, USA*

Received 22 April 2004; revised 12 July 2004; accepted 19 August 2004

We used functional magnetic resonance imaging (fMRI) to investigate whether the neural systems for nonspatial visual and auditory working memory exhibits a functional dissociation. The subjects performed a delayed recognition task for previously unfamiliar faces and voices and an audiovisual sensorimotor control task. During the initial sample and subsequent test stimulus presentations, activation was greater for the face than for the voice identity task bilaterally in the occipitotemporal cortex and, conversely, greater for voices than for faces bilaterally in the superior temporal sulcus/gyrus (STS/STG). Ventral prefrontal regions were activated by both memory delays in comparison with the control delays, and there was no significant difference in direct voxelwise comparisons between the tasks. However, further analyses showed that there was a subtle difference in the functional topography for two delay types within the ventral prefrontal cortex. Face delays preferentially activate the dorsal part of the ventral prefrontal cortex (BA 44/45) while voice delays preferentially activate the inferior part (BA 45/47), indicating a ventral/dorsal auditory/visual topography within the ventral prefrontal cortex. The results confirm that there is a modality-specific attentional modulation of activity in visual and auditory sensory areas during stimulus presentation. Moreover, within the nonspatial information-type domain, there is a subtle across-modality dissociation within the ventral prefrontal cortex during working memory maintenance of faces and voices.

Keywords: Working memory; fMRI; Face; Voice; Nonspatial information; Visual; Auditory

Introduction

Studies on working memory have suggested that there are dissociable domain-dependent functional topographies within the prefrontal cortex for working memory processing of verbal, spatial, and nonspatial visual information (for reviews, see Courtney et al., 1998b; Levy and Goldman-Rakic, 2000; Smith and Jonides, 1999). Dorsal prefrontal regions including the superior frontal sulcus have been proposed to be specialized for working memory maintenance of visual locations (e.g., Courtney et al., 1996, 1998a; Glahn et al., 2003; Gruber and von Cramon, 2003; Munk et al., 2002; Sala et al., 2003), whereas ventral prefrontal regions are more active for maintenance of nonspatial visual and verbal information (e.g., Awh et al., 1996; Courtney et al., 1997, 1998a; Jonides et al., 1997; Nystrom et al., 2000; Rämä et al., 2001). It has been proposed that the auditory system is also organized into two domain-dependent, spatial and nonspatial, processing streams, similar to that seen in the visual system (Ungerleider and Haxby, 1994; Ungerleider and Mishkin, 1982). Different regions of the auditory cortex have been shown to be differentially responsive to spatial and spectral features of auditory stimulation and these regions in the auditory belt cortex have been shown to project to distinct regions of the temporal, parietal, and prefrontal cortices (Romanski et al., 1999; Tian et al., 2001; for review, see Rauschecker and Tian, 2000). Recently, we and others have shown that during auditory working memory, maintenance of spatial and nonspatial information modulates activity preferentially in dorsal and ventral auditory pathways, respectively (Alain et al., 2001; Arnott et al., 2002; Rämä et al., 2004), suggesting that dorsal/ventral, spatial/nonspatial functional topography of the frontal cortex is similar for auditory and visual working memory. However, without a direct comparison to a comparable visual task in the same group of subjects, we cannot determine whether these auditory activations are in exactly the same anatomical location as the analogous visual working memory regions.

Single neurons showing selective responses during perceptual processing of monkey vocalizations have been localized within the ventral prefrontal cortex, but locations of auditory neurons were found to be more anterior and lateral than those of visually responsive neurons, suggesting that auditory and visual working memory systems are dissociable within the ventral prefrontal cortex (Romanski and Goldman-Rakic, 2002). It appears, however, that only small regions in the prefrontal cortex are modality-specific for auditory and visual information (Poremba et al., 2003).

---

* Corresponding author. Department of Psychology, Cognitive Brain Research Unit, PO Box 9, 00014 University of Helsinki, Finland. Fax: +358 9 191 29450.

E-mail address: prama@cc.helsinki.fi (P. Rämä).

doi:10.1016/j.neuroimage.2004.08.024

In the present study, we aimed to compare the neural systems involved in working memory for face and voice identities. We used voices because prefrontal neurons in monkeys were shown to respond better to natural sounds or monkey vocalizations than to pure tones (Azuma and Suzuki, 1984; Romanski and Goldman-Rakic, 2002), and also because the anterior part of the superior temporal sulcus (STS) in humans has been shown to exhibit selective activation for voices, leading to the suggestion that this region may be analogous to the face-sensitive area in the fusiform gyrus, which is a part of the ventral visual pathway (Belin et al., 2000, 2002; Shah et al., 2001). The subjects performed a delayed recognition task for previously unfamiliar faces and voices. To emphasize attentional and working memory processes, rather then perceptual mechanisms that are known to be different for auditory and visual stimuli, the faces and voices were presented simultaneously and the subjects were instructed to keep in mind either the face or the voice. Voice samples were non-imageable two-word phrases (e.g., “further consideration”) and were obtained by asking several individuals to utter a single word with a neutral emotional intonation. To find out whether a functional dissociation between auditory and visual information might occur during specific phases of this working memory task, we performed separate analyses of task-related activations evoked during the sample, delay, and test periods of the two memory tasks. The aim of the study was to find out (1) whether the neural mechanisms underlying encoding, maintenance, and recognition of faces and voices in working memory are dissociated or (2) whether nonspatial working memory activates common neuronal structures independent of stimulus modality. It was hypothesized that while neural structures in posterior cortical areas are segregated for modality-specific processing of identity information (face vs. voice), the ventral prefrontal cortex would be activated similarly for both working memory for faces and for voices. Such a result would suggest that a common neural network in the prefrontal cortex handles nonspatial information independent of the stimulus modality.

Materials and methods

Subjects

Twelve right-handed subjects (seven females) between the ages of 18 and 34 (mean 24 years) participated in the study. The subjects were native English speakers and were screened for mental and physical health. They had no history of head injury, drug or alcohol abuse, and no current use of medications that affect central nervous system or cardiovascular function. The subjects gave written informed consent, and were paid 50 USD for participating in the experiment. Nine (six females, mean 19 years) right-handed subjects participated in the behavioral study, which was conducted before brain scanning to ascertain whether the memory tasks were equally difficult. The experimental protocol was approved by the Institutional Review Boards of the Johns Hopkins University and the Johns Hopkins Medical Institutions.

Visual and auditory stimuli

Visual stimuli were presented using an LCD projector located outside of the scanning room connected to a Power Macintosh G3 computer running SuperLab software. The stimuli were projected onto a rear projection screen mounted inside the bore of the magnet, behind the subject’s head. Subjects viewed the stimuli using a mirror mounted at the top of the head coil. The stimuli were gray-scale images of previously unfamiliar female (n = 8) and male (n = 8) faces from three different views (Fig. 1A). A fixation cross was superimposed on the images, and the subjects were instructed to maintain visual fixation on the cross throughout the trial. Faces remained on the screen for 1.5 s, approximately the same duration as the auditory stimuli.

Voice samples (Fig. 1B) consisted of pair of words. The first word was a two-syllable adjective, and the second word, a five-syllable noun. The samples were recorded in a sound-proof room using CSL software. The sampling rate was 44.1 kHz. The targeted words were situated within a sentence (“John says that [further consideration] is important”) to encourage natural speech, and the speakers were instructed to read the sentences in a neutral tone. The pair of words that was to be later used as a stimulus in the experiment was always situated in the same position within a sentence. The speakers read each sentence twice during the recording. Eight female and eight male voices of native English speakers were recorded. Ten pairs of words were recorded, and three of them (“further consideration”, “simple inauguration”, and “constant unreality”), excerpted from the recorded sentences, were chosen for use in the study. The mean durations of the word pairs were 1155 ms (SD 104 ms), 1243 ms (SD 119 ms), and 1234 ms (SD 130 ms), respectively. There were no significant differences in durations between the three voice samples [$F(2, 45) = 2.66, P = 0.081$]. The energy levels (db) of voice samples were normalized using CSL/ASPP software. Faces and voices were presented simultaneously but independently so that each face was equally likely to appear with any of the voices at both sample and test.

For the control task, the visual stimuli were phase-scrambled in the Fourier domain maintaining contrast, luminance, and frequency information equal to those in the memory task (Fig. 1A). The auditory stimuli were also phase-scrambled in the Fourier domain, maintaining frequency information and stimulus amplitude envelopes equal to those in the memory task (Fig. 1B).

Each face and voice was presented 12 times during the experiment. Before the experiment, the subjects saw each face and heard each voice once to gain familiarity with the stimuli, and once or twice more during the memory task training. To ensure that subjects could not perform the task using an old/new recognition judgment, the same set of faces and voices were used repeatedly across trials. Subjects were familiarized with the stimuli before the scanning session to reduce variance between the beginning and the end of the scanning session due to changes in stimulus familiarity or strategy.

Working memory and sensorimotor control tasks

Two working memory tasks and a control task (Fig. 1C) were presented in a delayed recognition paradigm in which subjects were instructed to remember either faces or voices. One second before each trial, the subjects were presented with an instruction image (for 0.5 s) consisting of the word “face” (for the face task), “voice” (for the voice task), or “none” (for the control task), indicating which task was to be performed. In the delayed recognition paradigm, face and voice samples were presented simultaneously but the subjects were instructed to pay attention to only one of the stimulus types. In the face task, the subjects were to memorize the face of the image independent of angle of view, and in the voice task, the identity of the speaker independent of words spoken. Different views of the faces and different words spoken from sample to test ensured that

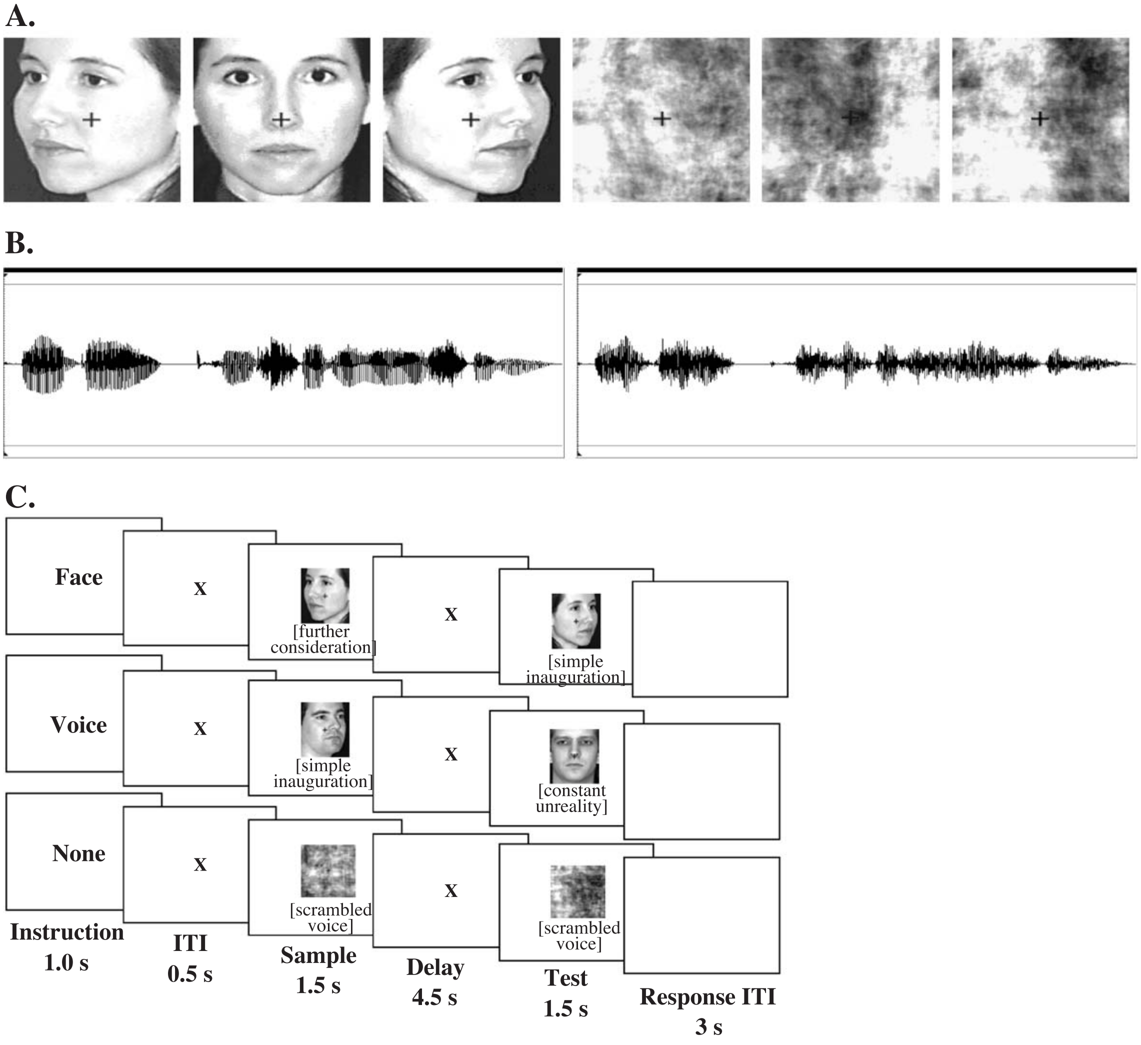

Fig. 1. (A) Examples of face and scrambled face stimuli. (B) Examples of voice and scrambled voice stimuli. (C) Illustration and timing of the delayed recognition and control tasks.

participants could not rely on surface stimulus features and had to encode and maintain those aspects of the stimuli that were constant to that person's identity across different photographs or voice recordings. The sample was presented for approximately 1.5 s followed by a memory delay of 4.5 s, during which the subjects saw a blank screen with a fixation cross. Then, a test stimulus was presented for approximately 1.5 s, during which time the subject indicated with a left or right button press whether the test stimulus was the same as the one in the sample period. Each subject was allowed to choose whether the right or left hand would correspond to the "match" or "nonmatch" response. Responses were made with left or right thumb presses of hand-held button boxes that were connected via a fiber optic cable to a Cedrus RB-6x0 Response Box. In the face task, the view of the face image presented during the test period never matched the view presented during the sample period. For the voice task, recorded words presented during the test period never matched the words presented during the sample period. Also, for the voice task, the face presented during the test period never matched the face presented during the sample period, and similarly for the face task, the voice presented during the test period never matched the voice presented during the sample period. Following each trial, there was an intertrial interval of 3.0 s. Subjects also performed a sensorimotor control (Ctrl) task with no mnemonic demand. For this task, the scrambled stimuli were presented with the same timing as in the memory tasks but the subjects were instructed that they need not remember the images or voices, but simply press both buttons when the test stimulus appeared.

During the scanning, six runs were conducted. In each run, both memory task conditions were presented in four alternating blocks of four trials each. Each block of four memory task trials was preceded and followed by one control trial. Thus, in each run, there were eight memory test trials for each information type and eight control trials. The order of tasks was counterbalanced across runs within each subject, and the order of runs was counterbalanced across subjects. The reaction times and match/no-match responses were recorded during the scanning. After the scanning, each subject was asked to fill out a questionnaire rating the difficulty of each task and mnemonic strategies used in his/her task performance.

In the behavioral study conducted in the laboratory environment before the brain scanning, the number of trials, number of blocks in each run, and the timings were equal to the memory tasks performed in the magnet. Control trials were not included in the behavioral study. During half of the runs, the subjects were exposed to the noise induced by an echo-planar imaging (EPI) sequence. The EPI sequence was recorded at the F.M. Kirby Research Center for Functional Brain Imaging, and played while the subject performed the tasks in the behavioral testing laboratory. Analysis of variance for repeated measures (BMDP2v, BMDP Statistical Software, Inc., Release 7.1) was used to test the main effects and interactions of task (Face vs. Voice) and experimental condition (silence vs. background EPI noise) on reaction times (RTs) and the accuracy of the task performance. For the data obtained during scanning, a pairwise t test was used to test the effect of task on the accuracy and RTs.

fMR imaging and data analysis

MR images were acquired with a 1.5-T Philips Gyroscan ACS-NT MR scanner (Philips Medical Systems). A T1-weighted structural image (70 axial slices, 2.5 mm, no gap, TR = 20 ms, TE = 4.6 ms, flip angle = 30°, matrix 256 × 256) was obtained before the functional scanning. During the performance of the tasks, subjects underwent T2*-weighted interleaved gradient-echo, echo-planar imaging (19 axial slices, 5-mm thickness, no gap, TR = 1500 ms, TE = 40 ms, flip angle = 70°, matrix 64 × 64). The images were phase-shifted using Fourier transformation to correct for slice acquisition time, then motion-corrected using automatic image registration (AIR) software (Woods et al., 1998), and analyzed separately for each subject using multiple regression (Friston et al., 1995) with Analysis of Functional NeuroImages (AFNI) software (Cox, 1996). Changes in neural activity were modeled as square-wave functions matching the time course of events of experimental tasks. These square waves were convolved with a gamma function model of the hemodynamic response using the following values: 2.0 s for lag, 3.0 s for rise time, and 5.0 s for fall time to create the regressors of interest in the multiple regression analysis. Additional regressors were included to model sources of variance not related to the experimental manipulations (mean intensity between and linear drift within time series). Both memory task conditions (face and voice) were separately contrasted to control tasks, and to each other, for each of the three main events of the tasks (encoding, delay, and recognition). For the direct comparisons between the memory tasks (face vs. voice), the analysis was restricted to only those voxels showing significantly greater activity for any of the memory versus control tasks (face vs. control task, voice vs. control task, both tasks vs. control task) in a given main event of the tasks. Each of these contrasts resulted in a Z-map for each subject.

Z-maps were registered into the Talairach coordinate system (Talairach and Tournoux, 1988) and resampled to 1 mm^3. Average Z-maps were computed by dividing the sum of individual subject Z-values by the square root of the sample size using AFNI software (Cox, 1996). This procedure is mathematically equivalent to doing a group (fixed-effect) multiple regression analysis. All tests of voxelwise significance, which were applied to the mean Z images, were held to a Z threshold of 2.33, corresponding to a $P < 0.01$, and corrected for multiple comparisons ($P < 0.05$) using a measure of probability that uses the individual voxel Z score threshold and the number of contiguous significant voxels. Based on a Monte Carlo simulation run via AFNI (Ward, 2000), it was estimated that a 387 mm^3 contiguous volume (six voxels, each measuring 3.59 × 3.59 × 5 mm) for the volume of the entire brain would meet the $P < 0.05$ threshold. For the direct comparison between memory tasks, the analysis was restricted to only those voxels showing significantly greater activity for any of the memory tasks versus control. Within this restricted number of voxels, a 258 mm^3 cluster size (4 voxels) satisfied a 0.05 experiment-wise probability. Activations were anatomically localized in the averaged maps using T_1-weighted images.

Region of interest (ROI) analysis

Based on anatomical hypotheses derived from previous studies of nonspatial working memory for visual stimuli (e.g., Rämä et al., 2001), ROIs encompassing the anterior inferior frontal gyrus and anterior insula (IFG/Insula) and posterior inferior frontal gyrus (IFG) were drawn in both hemispheres of a Talairach transformed brain according to Brodmann areas (BAs) and anatomical landmarks of the Talairach (Talairach and Tournoux, 1988) and Damasio (Damasio, 1995) brain atlases. The IFG/Insula ROI included BAs 45 and 47 of the IFG (z = −5.0–16.00 mm). The posterior border of the IFG/Insula ROI was the anterior bank of the sylvian fissure (z = −5.0–12.0 mm) and the anterior bank of the precentral sulcus (PreCS) (z = 12.0–16.00 mm). The anterior border of the IFG/Insula ROI was the inferior frontal sulcus (IFS). The IFG ROI included BAs 44 and 45 of the IFG (z = 17.0–34.0 mm). The posterior border of the IFG ROI was the anterior bank of PreCS, and the anterior border was the posterior bank of IFS.

For each ROI, the number of voxels significantly activated (not corrected for multiple comparisons) during memory delays relative to the control delays was computed for each subject. The number of significantly activated voxels was then normalized by dividing by the total number of voxels in each ROI. In addition, for each ROI, the mean signal intensities (β-coefficients) across all significantly activated voxels (corrected for multiple comparisons), determined as described above, were computed for each subject. Analysis of variance for repeated measures with subject as a random factor (BMDP2v, BMDP Statistical Software, Inc., Release 7.1) was used to test the main effects and interactions of task, event, hemisphere, and brain region on both the number of suprathreshold voxels and β-coefficients. A pairwise t test was then used to test the effect of task on the number of activated voxels and β-coefficients separately for each ROI.

Response topography correlation analysis

The multiple regression on the original data was re-run with separate regressors for either odd or even numbered blocks of each task. Then, these four sets of coefficients (even-faces, odd-faces, even-voices, odd-voices) for all significantly activated voxels within each subject were used in the correlation analysis. Correlations were calculated between the response topography on odd blocks and the response topography on even blocks of the same WM task, for a measure of within-task consistency of the topography. Correlations were also calculated between the response topography on odd (even) blocks of one WM task and the response topography on odd (even) blocks of the other WM task, for a measure of between-task consistency of the topography. In the group average Z-maps, clusters of voxels that were activated in any of the planned contrasts—face versus control, voice versus control, or both tasks combined versus control—were assigned to five broadly defined anatomical regions (each including both hemispheres except posterior parietal cortex in which activity was observed only in the left hemisphere): ventral prefrontal cortex (IFG/Insula, IFG, and MFG), dorsal prefrontal cortex (SFS, PreCG, and MFG), posterior parietal cortex, temporal cortex, and occipital cortex. Correlation coefficients were converted to Z scores and t tests were performed to test whether the response topography within each region was more highly correlated within

Table 1
Sample activity for faces and voices versus control

Area	Face > Control				Voice > Control			
	x, y, z	Peak Z	Mean Z	Spatial extent (mm^3)	x, y, z	Peak Z	Mean Z	Spatial extent (mm^3)
Occipital								
FG	−38, −67, −16	5.40[a]	3.12	7971				
	33, −80, −21	5.72[b]	3.09	11,528				
LG	12, −38, −2	3.82	2.78	968				
LOG	−29, −88, −1	4.98[a]						
	38, −82, −6	5.99[b]						
Cuneus	−6, −79, 19	4.17	2.82	913				
Temporal								
STS/STG	−59, −16, −2	4.28	2.90	1188	−55, −4, 0	5.16	2.94	930
	59, −36, 6	4.11	2.82	435	51, −18, −2	6.74	3.21	2165
					45, −45, 7	5.25	2.95	1250
Parietal								
IPS/AG	34, −57, 37	3.93	2.85	413	33, −63, 31	4.46	2.80	677
IPS/SPL	−27, −69, 25	4.01	2.92	622				
SPL	−20, −84, 25	5.40	2.98	667				
Frontal								
IFG/Insula					51, 5, 4	4.43	2.85	432
IFG					−44, 19, 28	4.34	2.84	388

Note. Areas of significant activity, the peak and mean Z values, the spatial extent of a given activity, and the Talairach coordinates of maximum Z value within each region during the sample period of the memory tasks relative to control tasks.

[a] The cluster included two separate activation loci with local maxima in both the left FG and left LOG.

[b] The cluster included two separate activation loci with local maxima in the right FG and right LOG.

Table 2
Delay activity for faces and voices versus control

Area	Face > Control				Voice > Control			
	x, y, z	Peak Z	Mean Z	Spatial extent (mm^3)	x, y, z	Peak Z	Mean Z	Spatial extent (mm^3)
Occipital								
FG	35, −41, −19	4.41	2.94	1087				
	−41, −51, −16	3.54	3.19	699				
Temporal								
STS/STG	−55, −29, 1	3.65	2.85	541	−60, −31, 5	5.24	3.06	2445
					54, −3, 1	5.53	2.95	636
STS/MTG					44, −47, 0	5.27	2.98	1285
Parietal								
IPS/SMG	−33, −69, 36	6.39	3.25	4229	−35, −64, 35	4.70	3.00	754
Frontal								
IFG/Insula	−27, 18, 9	6.84	3.41	1301	−35, 11, 8	5.82	3.11	1677
	30, 16, 8	4.63	3.15	664	29, 14, 8	5.14	3.18	408
					40, 14, 14	5.33	3.02	980
IFG	−45, 26, 20	4.53	3.24	387	−41, 3, 29	5.14	3.06	2571
	−35, 10, 26	6.67	3.36	4695				
	42, 11, 26	4.70	3.07	3109	37, 0, 30	4.53	2.95	1357
MFG					−37, −4, 54	5.30	3.08	412
SFS/PreCG	−24, −15, 50	4.93	3.16	1472				
SFGm	4, 16, 45	7.33	3.41	5582	−1, −5, 56	9.26	3.59	4463
Others								
Caudate nucleus	−14, 11, 5	3.31	2.83	415				

Note. Areas of significant activity, the peak and mean Z values, the spatial extent of a given activity, and the Talairach coordinates of maximum Z value within each region during the delay period of the memory tasks relative to control tasks.

task than between tasks. ANOVAs and correlation analyses that were performed on the β-coefficients derived from the initial multiple regression analysis all included subjects as a random factor. Also, within each of these regions, for each subject, the correlation analysis was limited to those voxels that were significantly positively activated for that subject individually, in any of the same contrasts. Within each subject, within each of the five regions (ventral prefrontal, dorsal prefrontal, parietal, temporal, and occipital), these voxels were ordered hierarchically in two ways. These two sorting procedures each gave a one dimensional "voxel index" that included all values for each of the three spatial directions. In one sorting procedure, voxels were sorted first by posterior to anterior voxel value. Within each group of voxels with the same anterior–posterior value, voxels were sorted again by ventral–dorsal, and then again within the group of voxels that had the same anterior–posterior and ventral–dorsal values, voxels were sorted again by left–right voxel value. The second sorting procedure was identical, but with ventral–dorsal first, then anterior–posterior, and finally left–right. The beta coefficients as a function of voxel index for each WM task thus became a metric for the response topography within each region.

Results

Behavioral data

A separate group of participants was tested on both the auditory and visual working memory tasks before the fMRI portion of the study to test for difficulty differences between the tasks and the possible interference of the background noise of MRI scanning on task performance. These results showed that there was no significant main effect of memory task (face vs. voice) on the accuracy of the task performance [$F(1,8) = 1.33$, $P = 0.28$]. The main effect of condition (silence vs. background recorded EPI noise) was significant [$F(1,8) = 7.11$, $P < 0.05$]. The subjects made more errors during magnet noise (mean accuracy for face task was 84%, for voice task 86%) than without magnet noise (for face task 90%, for voice task 91%) during both tasks. However, there was no interaction between task and condition [$F(1,8) = 0.00$, $P = 0.99$]. The main effects of task [$F(1,8) = 5.40$, $P < 0.05$] and condition [$F(1,8) = 8.89$, $P < 0.05$] on reaction time (RT) were significant. RTs were longer for the voice than for the face task, and longer during magnet noise condition (1603 ms for face task, 1675 ms for voice task) than during silence (1485 ms for face task, 1578 ms for voice task). Again, there was no interaction between task and condition [$F(1,8) = 0.35$, $P = 0.57$]. During fMRI scanning, however, the subjects made more incorrect responses during the voice (mean accuracy 82%) than the face (mean accuracy 89%) task [$t(12) = 2.60$, $P < 0.05$]. The RTs were also longer for the voice (1595 ms) than for the face (1376 ms) task but the difference was not statistically significant.

fMRI data

Voxelwise multiple regression

Face and voice task activations relative to control (Tables 1–3)

Sample period. For the face samples, activation was detected bilaterally in the fusiform gyrus (FG) and lateral occipital gyrus (LOG), in the right lingual gyrus (LG), and in the left cuneus. The superior temporal sulcus/gyrus (STS/STG) and the inferior parietal cortex were bilaterally activated, and the superior parietal cortex was activated on the left hemisphere. For the voice samples, there was bilateral activation of the STS/STG and inferior frontal gyrus, whereas the inferior parietal cortex was activated in the right hemisphere (Table 1).

Delay period. Several occipital, temporal, parietal, and frontal regions were activated during the delay period of the tasks. The FG was bilaterally activated during face delays. For voice delays, there was bilateral activation of the STS/STG, whereas the left STS/STG was activated during face delays. The left intraparietal sulcus/supramarginal gyrus (IPS/SMG) was activated during both delays. Finally, in the frontal lobe, the left middle frontal gyrus was activated for voice delays and the left superior frontal sulcus/precentral gyrus (SFS/PreCG) for face delays. The inferior frontal gyrus/insula (IFG/insula), IFG, and medial part of the superior frontal gyrus (SFGm) were activated during both delays. Caudate nucleus was also activated during face delays (Table 2; Fig. 2B).

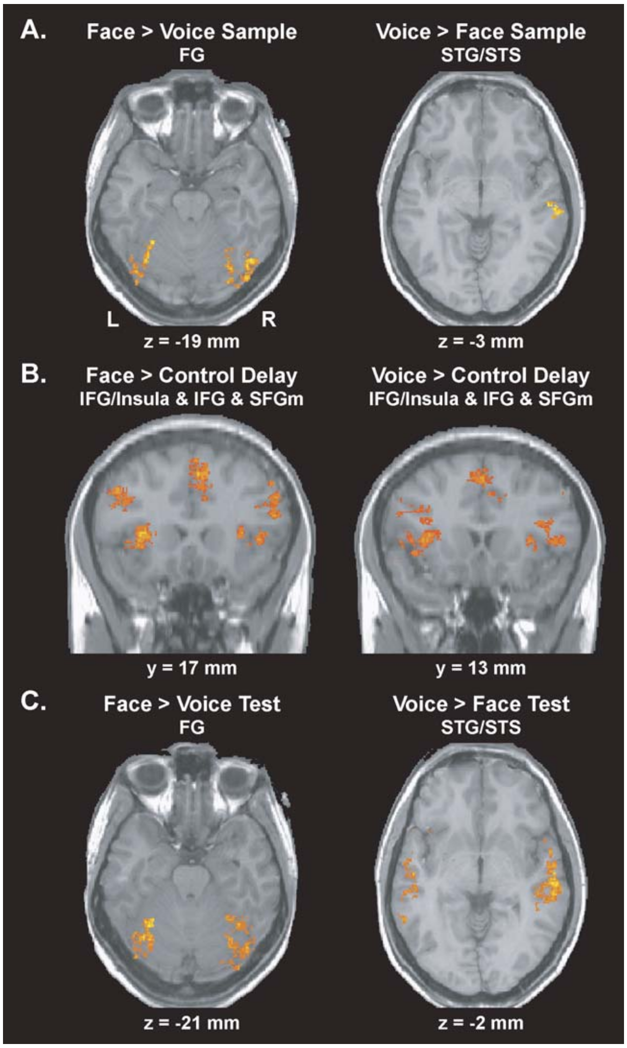

Fig. 2. Cross-subject average statistical maps of activation (A) during direct comparisons between activations during the sample period of the face and voice tasks overlaid on a Talairach normalized anatomical image, (B) during the delay period of the face and voice tasks (relative to control task), and (C) during direct comparisons between activations during the test period of the face and voice tasks.

Test period. Several occipital, temporal, parietal, and frontal regions were also activated during the test period of the tasks. There was bilateral activation of the FG, STS/STG, and IPS/SMG during both tasks. In the frontal lobe, bilateral IFG/Insula was activated during both tasks. The left IFG was activated during the face task, whereas the right IFG during both tasks. The right MFG and SFGm were activated during both tasks. Parahippocampal gyrus/hippocampus (ParaHC/HC) and posterior cingulate gyrus were activated during the face task (Table 3).

Direct voxelwise comparisons: face > voice and voice > face. During the sample period, the FG and LOG were activated more for the face than for the voice task, whereas the right STS/STG was activated more for voice than for face samples. Similarly, during the test period, the FG was more active for the face task and the STS/STG was more active for the voice task. In addition, several frontal areas (IFG/Insula, IFG, MFG, and SFGm) were activated more for the test period of the voice than for the face task. During the delay period, the left FG and the left IPS/SMG were activated more for the face than for the voice task, whereas the SFGm was activated more for voice than for face delays (Table 4; Figs. 2A,C).

ROI analyses in the ventral prefrontal areas

During the stimulus presentation, the voxelwise multiple regression analysis showed a dissociation between occipital (FG/LG) and temporal (STS/STG) cortical areas for face versus voice tasks, whereas during the delays, the ventral prefrontal cortex was active for both tasks and there were no significant differences between the tasks in the direct voxelwise comparisons within this region. Such a result, however, does not prove absence of a functional dissociation, and therefore the data for delay periods were further analyzed using two other methods. First, an ROI analysis was performed within the ventral prefrontal cortex (comparing the IFG to the IFG/Insula as described in Materials and methods) on both the number of activated voxels and the signal intensity (β-coefficients), and then, functional topography correlation analysis was performed on the spatial pattern of activation magnitude across voxels within activated clusters in the occipital, temporal, parietal, and prefrontal cortices.

The ROI analysis during the memory delays demonstrated that there was a significant main effect of brain region on the number of significantly activated voxels [$F(1,11) = 7.13$, $P < 0.05$] but no main effect of task nor interaction between the task and brain region. For signal intensities, however, an interaction between task

Table 3
Test activity for faces and voices versus control

Area	Face > Control				Voice > Control			
	x, y, z	Peak Z	Mean Z	Spatial extent (mm^3)	x, y, z	Peak Z	Mean Z	Spatial extent (mm^3)
Occipital								
FG	−36, −71, −19	6.68	3.23	8348	−37, −68, −21	3.94	2.89	689
	35, −55, −21	6.56	3.54	9237	38, −68, −19	4.10	3.07	968
Temporal								
STS/STG	−62, −32, 0	6.47	3.18	5574	−56, −27, −1	8.26^{a}	3.38	17 398
	53, −23, −2	6.57^{b}	3.23	23 256	54, −23, −2	10.61^{c}	3.50	33 010
Parietal								
IPS/SMG	−36, −72, 41	5.16	3.12	1168	−28, −63, 35	4.45	2.92	811
	35, −71, 43	5.89	3.17	4716	30, −69, 37	5.70	3.20	5982
Frontal								
IFG/Insula	−26, 22, 3	6.66	3.16	2747	−27, 22, 3	8.15^{a}		
	46, 9, 0	7.76^{b}			42, 13, 1	6.20^{c}		
IFG	−46, 24, 30	5.35	2.91	690				
	−42, 7, 33	4.40	2.88	879				
	44, 32, 23	6.86^{b}			44, 31, 18	6.53^{c}		
					45, 17, 31	5.94^{c}		
MFG	37, −5, 57	5.42	3.00	474	44, 11, 41	6.39^{c}		
	39, 20, 32	6.25^{b}						
SFGm	3, 2, 55	6.04	3.13	4152	1, 7, 52	7.17	3.23	5767
Others								
ParaHC/HC	−20, −14, −10	3.88	2.86	637				
	13, −15, −10	6.05	2.95	1946				
Post. Cingulate G	2, −27, 25	6.10	3.01	2399				

Note. Areas of significant activity, the peak and mean Z values, the spatial extent of a given activity, and the Talairach coordinates of maximum Z value within each region during the test period of the memory tasks relative to control tasks.

[a] The cluster included two separate activation loci with local maxima in the left STS/STG and left IFG/Insula.

[b] The cluster included four separate activation loci with local maxima in the right STS/STG, left IFG/Insula, left IFG, and left MFG.

[c] The cluster included four separate activation loci with local maxima in the right STS/STG, right IFG/Insula, right IFG and right MFG.

Table 4
Direct comparisons between the tasks

Area	Face > Voice				Voice > Face			
	x, *y*, *z*	Peak *Z*	Mean *Z*	Spatial extent (mm^3)	*x*, *y*, *z*	Peak *Z*	Mean *Z*	Spatial extent (mm^3)
Sample								
FG	−32, −61, −20	4.62	3.12	1503				
	42, −62, −19	5.23	3.03	1501				
LOG	−33, −88, 5	4.82	3.14	605				
STS/STG					59, −26, −3	4.17	2.91	260
Delay								
FG	−39, −53, −18	4.34	3.22	321				
IPS/SMG	−26, −66, 42	3.96	2.93	420				
SFGm					−4, −4, 56	4.40	2.90	377
Test								
FG	−34, −61, −21	5.93	3.23	1170				
	33, −51, −21	5.05	3.11	2753				
STS/STG					−54, −27, −1	5.21	3.07	2380
					56, −22, −3	6.12	3.12	3686
IFG/Insula					−47, 15, 4	5.49	3.06	551
					51, −1, 3	5.20	3.10	636
IFG					47, 15, 25	5.07	2.97	948
MFG					37, 3, 45	4.73	2.96	796
SFGm					3, 24, 45	4.05	2.80	377
					3, 1, 57	4.17	3.13	437

Note. Areas of significant activity, the peak and mean *Z* values, the spatial extent of a given activity, and the Talairach coordinates of maximum *Z* value within each region during each period of the memory tasks for faces versus voices.

and brain region was significant [$F(1,11) = 11.36$, $P < 0.01$]. Signal intensity of activated voxels was greater for face than for voice delays in the left IFG ($P < 0.05$), whereas signal intensity was greater for voice than for face delays in the right IFG/Insula but this difference did not quite reach statistical significance ($P = 0.059$).

Functional topography correlation analysis

To test for functional topographies of activation during memory delays, we performed a correlation analysis on the pattern of activation magnitude within activated clusters in the prefrontal, parietal, temporal, and occipital cortices. The analysis is described in detail in Materials and methods, and the results in two individual subjects are shown in Fig. 3.

Functional topographies were significantly more similar within task than between tasks for ventral frontal ($r = 0.68$ and 0.63, respectively, $P < 0.05$), dorsal frontal ($r = 0.71$ and 0.59, respectively, $P < 0.05$), and temporal ($r = 0.73$ and 0.69, respectively, $P < 0.05$) cortices. In the ventral prefrontal cortex, activation for the face task increased more than for the voice task with increasing ventral–dorsal voxel index (slopes = 8.84^{-05} vs. 0.49^{-05}, $P < 0.05$), indicating a ventral–dorsal auditory–visual topography in the ventral prefrontal cortex. This is consistent with the results of the ROI analysis, which showed more activation in the IFG/Insula (BA 45/47) for the voice task and more activation in the IFG (BA 44/45, dorsal to the IFG/insula) for the face task). Dorsal frontal cortex activation for the face identity task decreased significantly more than for the voice task with increasing posterior–anterior voxel index (slopes = -3.58^{-04} vs. -0.73^{-04}, $P < 0.005$). This is consistent with the voxelwise analysis, which indicated that the SFS/PreGC was more active for face delays relative to control delays, whereas the posterior part of the MFG, located anterior to the SFS/PreCG, was more active for voice delays relative to control delays (Table 2). In the temporal cortex, there were no statistically significant differences in slopes for any of the voxel index orders.

In summary, voxelwise multiple regression analysis showed that during the stimulus presentation, activation was greater for the face than for the voice identity task in the FG and LOG and, conversely, greater for voices than for faces in the STS/STG. Ventral prefrontal regions (IFG/Insula and IFG) were activated by both memory delays in comparison with the control delays and there was no difference in direct individual voxel comparisons between the tasks. Further analyses (ROI and functional topography correlation analyses) for memory delays, however, provided evidence for a difference in topography for face and voice tasks, suggesting that separable neural systems in the ventral prefrontal cortex are recruited by nonspatial visual and auditory working memory processing.

Discussion

The results suggest that there is a modality-specific attentional modulation of activity in visual and auditory sensory areas during the stimulus presentation in a working memory task. The FG and LOG were more active for faces than for voices whereas the STG/STS was more active for voices than for faces. Earlier, the FG has been shown to be more active during viewing or matching of faces than that of other objects (Kanwisher et al., 1997; McCarthy et al., 1997; Sergent et al., 1992), locations (Courtney et al., 1996, 1997; Haxby et al., 1994), or scrambled faces (Clark et al., 1996, 1998; Courtney et al., 1996, 1997; Kanwisher et al., 1997; Puce et al.,

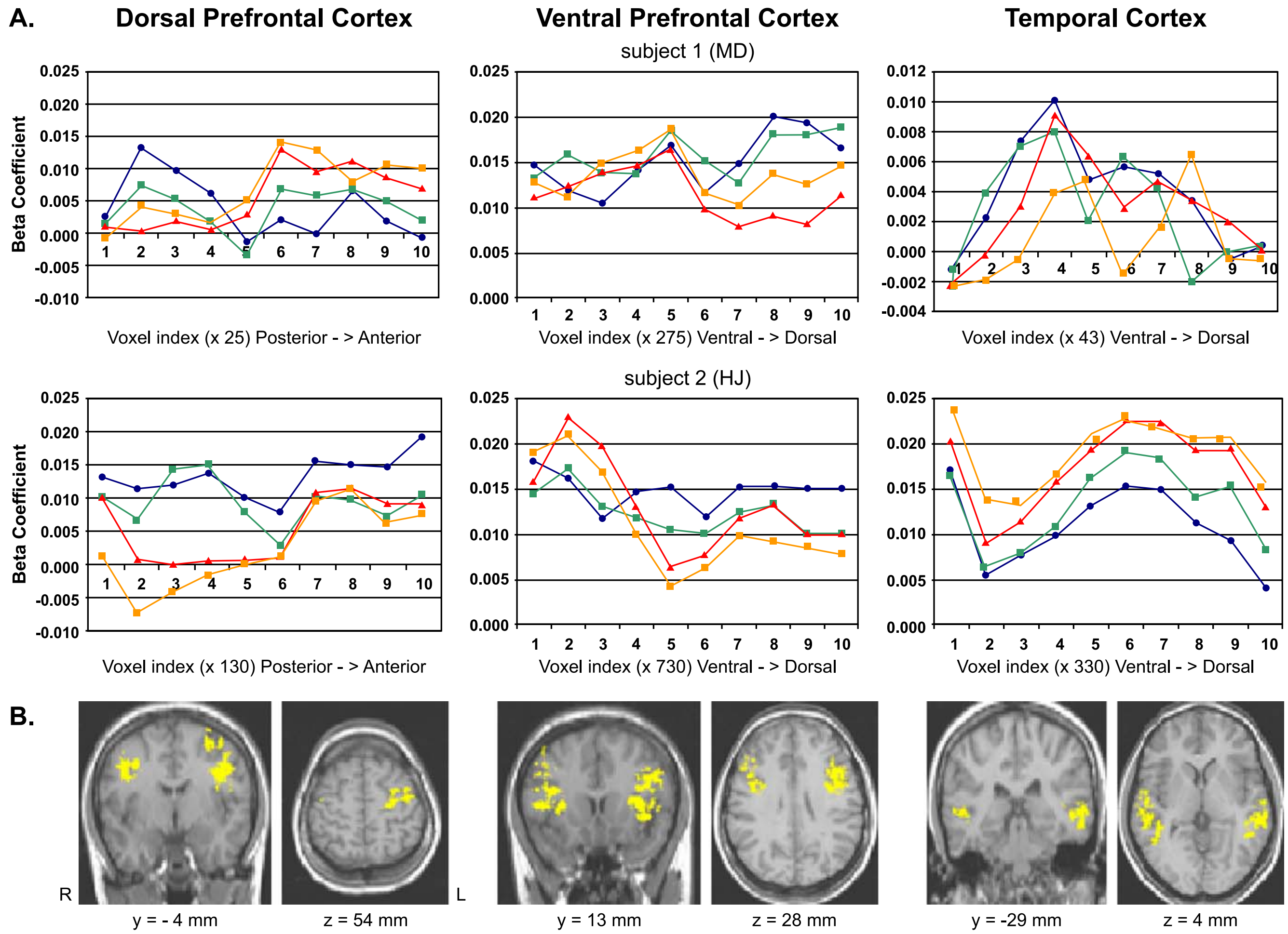

Fig. 3. Results of the functional topography correlation analysis for two individual subjects. (A) The topographies are plotted with response magnitude (beta coefficient) as a function of voxel index. Blue and green lines illustrate odd and even trials for face task, and red and yellow lines odd and even trials for voice task, respectively. For the dorsal frontal region, the voxels are ordered first from posterior to anterior. For the ventral prefrontal and temporal regions, the voxels are ordered first from ventral to dorsal. For illustrative purposes, each data point shown is the average of the beta coefficients from 25 to 730 voxels depending on the total number of significantly activated voxels for that subject within each region. (B) The horizontal and coronal brain images show the three regions (defined from the group multiple regression analysis) within which the response topographies were analyzed.

1995). The STS has been earlier shown to be selectively activated by voices (Belin et al., 2000, 2002; Shah et al., 2001). Our results support the suggestion that the STS may be analogous to the face sensitive area in the FG (Belin et al., 2000, 2002; Shah et al., 2001). The ventral prefrontal cortex was active for both face and voice delay types and there was no difference in a direct individual voxel comparison between the tasks in the present study. Such a result, however, does not prove the absence of a functional dissociation, and therefore the data for delay periods were further analyzed using two other methods (ROI and functional topography correlation analyses). Both of these analyses showed that there was a within-domain, across modality dissociation within the ventral prefrontal cortex for face and voice memory delays.

Within the ventral prefrontal cortex, there was a ventral/dorsal auditory/visual functional topography for working memory maintenance. The inferior part of the ventral prefrontal cortex was more active for voice delays whereas the more dorsal part of the ventral prefrontal cortex was more active for face delays. The present results are partly consistent with those of an earlier positron emission tomography study that compared working memory for aurally and visually presented phonological information (Schumacher et al., 1996). That study showed that Broca's area in the ventral prefrontal cortex was active for both tasks relative to control tasks but the inferior part of the region was more active for the auditory than for the visual task. Consistent with that study and the current study, although only small regions in the prefrontal cortex are shown to be modality-specific for auditory and visual information in monkeys, the auditory modality-specific regions in the ventral prefrontal cortex were found inferiorly and anteriorly to those of visual modality-specific regions (Poremba et al., 2003). In a more recent fMRI study, the modality-specific activity was not found in the ventral, but in the dorsolateral prefrontal cortex, which was shown to be more active for auditory than for visual working memory (Crottaz-Herbette et al., 2004). That study did not find a double dissociation between auditory and visual information within prefrontal cortex. Both of these earlier imaging studies, however, used phonological (e.g., letters or digits) material to be remembered in their memory tasks, and auditory and visual stimuli were presented separately in the different task blocks. To our knowledge, the present study is a first attempt to study nonverbal auditory and visual working memory.

In our memory tasks, face and voice samples were presented simultaneously but the subjects were instructed to pay attention to one or the other of the stimulus types. Such an experimental design ensures that the differences that are observed are due to differences in the neural systems for active working memory maintenance of auditory versus visual information and not simply to differences in sensory stimulation. Thus, the relative rather than absolute dissociation observed in this study is not surprising. There are several differences between the memory and control tasks and activity common for both tasks relative to the control task is nonspecific and could relate to several different cognitive processes. The simultaneous presentation of face and voice stimuli may have additionally diminished differences in activation between the tasks, because even though the subjects in the present study were not asked to maintain the conjunction of faces and voices in memory, they may have done so automatically due to nature of stimuli. Working memory maintenance of conjunction of spatial and nonspatial visual information has been shown to recruit the same neural system as the maintenance of either information type alone but the neural responses are less within the dorsal and ventral areas compared to the responses obtained during processing of only their preferred information (Sala and Courtney, 2002). On the other hand, it has been shown that cross-modal binding of auditory and visual information is associated with greater activity in the multimodal areas, such as insula/frontal operculum, dorsolateral frontal, posterior parietal, and superior temporal gyri, but lower activity in the unimodal areas (Bushara et al., 2003; Calvert et al., 2000). In the present study, however, we did not compare activity during separate presentation of face or voices to simultaneous presentation of both stimulus types.

In addition to prefrontal areas, several other regions were activated during the delay period of the tasks. Superior temporal gyrus/sulcus was activated during both face and voice delays relative to control delays whereas the fusiform gyrus was activated only during face delays. Superior temporal sulcus, in addition to exhibiting voice-selective activity, has been shown to be part of the ventral visual pathway involved in processing of facial expressions, gaze direction and lip movements (for review, see Haxby et al., 2000). The left posterior parietal cortex was activated during both face and voice delays relative to control delays but in the direct comparison between the tasks; activation was greater during the face than voice delays. In our face task, the subjects were instructed to maintain the identity of the face in working memory independently of the direction of view. It is possible that the activity in the IPS/SMG is related to engagement of the spatial processing system.

Earlier, single neurons sensitive to monkey vocalization in the ventral prefrontal cortex were found to be anatomically located more anteriorly and laterally than those of visually responsive neurons (Romanski and Goldman-Rakic, 2002). The present results suggest that there might also be a within-domain, across-modality dissociation with the human ventral prefrontal cortex for face and voice memory delays. However, this within-information-type (nonspatial) and across-modality topography appears to be much less pronounced than the within-modality and across-information-type (nonspatial and spatial) topography previously reported within visual (e.g., Courtney et al., 1998a; Glahn et al., 2003; Munk et al., 2002; Sala et al., 2003) and auditory (e.g., Alain et al., 2001; Arnott et al., 2002; Rämä et al., 2004) systems. Comparison across these previous studies suggests that spatial information is preferentially maintained by the superior frontal cortex for both visual and auditory information. Further research is needed to ascertain whether there may also be a functional topography within the superior frontal cortex for visual versus auditory spatial information maintenance.

Acknowledgments

Pia Rämä is supported by the Academy of Finland (75790). This research was supported by the National Institute of Mental Health (R01 MH61625). The authors thank the entire staff of F. M. Kirby Research Center for Functional Brain Imaging, Kennedy Krieger Institute, where the data were acquired. We wish to thank Dr. James Haxby for providing face stimuli for the study. We thank Dr. Jusczyk in Department of Psychological and Brain Sciences, at Johns Hopkins University, for providing facilities for recording and editing the voice samples, and Dr. Elliott Moreton in Cognitive Science Department, at Johns Hopkins University, for programming software to scramble the sounds.

References

Alain, C., Arnott, S.R., Hevenor, S., Graham, S., Grady, C.L., 2001. "What" and "where" in the human auditory system. Proc. Natl. Acad. Sci. U. S. A. 98, 12301–12306.

Arnott, S.R., Alain, C., Hevenor, S., Graham, S., Dade, L.A., Grady, C., 2002. What, where, and how in the human prefrontal cortex. Program no. 181.1. Abstract Viewer/Itinerary Planner. Society for Neuroscience, Washington, DC (Online).

Awh, E., Jonides, J., Smith, E.E., Schumacher, E.H., Koeppe, R.A., Katz, S., 1996. Dissociation of storage and rehearsal in verbal working memory: Evidence from positron emission tomography. Psychol. Sci. 7, 25–31.

Azuma, M., Suzuki, H., 1984. Properties and distribution of auditory neurons in the dorsolateral prefrontal cortex of the alert monkey. Brain Res. 298, 343–346.

Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. Nature 403, 309–312.

Belin, P., Zatorre, R.J., Ahad, P., 2002. Human temporal-lobe response to vocal sounds. Cogn. Brain Res. 13, 17–26.

Bushara, K.O., Hanakawa, T., Immisch, I., Toma, K., Kansaku, K., Hallett, M., 2003. Neural correlates of cross-modal binding. Nat. Neurosci. 6, 190–195.

Calvert, G.A., Campbell, R., Brammer, M.J., 2000. Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Curr. Biol. 10, 649–657.

Clark, V.P., Keil, K., Maisog, J.M., Courtney, S., Ungerleider, L.G., Haxby, J.V., 1996. Functional magnetic resonance imaging of human visual cortex during face matching: a comparison with positron emission tomography. Neuroimage 4, 1–15.

Clark, V.P., Maisog, J.M., Haxby, J.V., 1998. fMRI study of face perception and memory using random stimulus sequences. Neurophysiology 79, 3257–3265.

Courtney, S.M., Ungerleider, L.G., Keil, K., Haxby, J.V., 1996. Object and spatial visual working memory activate separate neural systems in human cortex. Cereb. Cortex 6, 39–49.

Courtney, S.M., Ungerleider, L.G., Keil, K., Haxby, J.V., 1997. Transient and sustained activity in a distributed neural system for human working memory. Nature 386, 608–611.

Courtney, S.M., Petit, L., Maisog, J.M., Ungerleider, L.G., Haxby, J.V., 1998a. An area specialized for spatial working memory in human frontal cortex. Science 279, 1347–1351.

Courtney, S.M., Petit, L., Haxby, J.V., Ungerleider, L.G., 1998b. The role of prefrontal cortex in working memory: examining the contents of consciousness. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 353, 1819–1828.

Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29, 162–173.

Crottaz-Herbette, S., Anagnoson, R.T., Menon, V., 2004. Modality effects in verbal working memory: differential prefrontal and parietal responses to auditory and visual stimuli. NeuroImage 21, 340–351.

Damasio, H., 1995. Human Brain Anatomy in Computerized Images. Oxford Univ. Press, New York.

Friston, K.J., Holmes, A.P., Poline, J.-B., Grasby, P.J., Williams, C.R., Frackowiak, R.S.J., 1995. Analysis of fMRI time-series revisited. NeuroImage 2, 45–53.

Glahn, D.C., Kim, J., Cohen, M.S., Poutanen, V.P., Therman, S., Bava, S., Van Erp, T.G., Manninen, M., Huttunen, M., Lonnqvist, J., Standertskjold-Nordenstam, C.G., Cannon, T.D., 2003. Maintenance and manipulation in spatial working memory: dissociations in the prefrontal cortex. NeuroImage 17, 201–213.

Gruber, O., von Cramon, D.Y., 2003. The functional neuroanatomy of human working memory revisited. Evidence from 3-T fMRI studies using classical domain-specific interference tasks. NeuroImage 19, 797–809.

Haxby, J.V., Horwitz, B., Ungerleider, L.G., Maisog, J.M., Pietrini, P., Grady, C.L., 1994. The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. J. Neurosci. 14, 6336–6353.

Haxby, J.V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. Trends Cogn. Sci. 4, 223–233.

Jonides, J., Schumacher, E.H., Smith, E.E., Lauber, E.J., Awh, E., Minoshima, S., Koeppe, R.A., 1997. Verbal working memory load affects regional brain activation as measured by PET. J. Cogn. Neurosci. 9, 462–475.

Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17, 4302–4311.

Levy, R., Goldman-Rakic, P.S., 2000. Segregation of working memory functions within the dorsolateral prefrontal cortex. Exp. Brain Res. 133, 23–32.

McCarthy, G., Puce, A., Gore, J.C., Allison, T., 1997. Face-specific processing in the human fusiform gyrus. J. Cogn. Neurosci. 9, 605–610.

Munk, M.H., Linden, D.E., Muckli, L., Lanfermann, H., Zanella, F.E., Singer, W., Goebel, R., 2002. Distributed cortical systems in visual short-term memory revealed by event-related functional magnetic resonance imaging. Cereb. Cortex 12, 866–876.

Nystrom, L.E., Braver, T.S., Sabb, F.W., Delgado, M.R., Noll, D.C., Cohen, J.D., 2000. Working memory for letters, shapes, and locations: fMRI evidence against stimulus-based regional organization in human prefrontal cortex. NeuroImage 11, 424–446.

Poremba, A., Saunders, R.C., Crane, A.M., Cook, M., Sokoloff, L., Mishkin, M., 2003. Functional mapping of the primate auditory system. Science 299, 568–572.

Puce, A., Allison, T., Gore, J.C., McCarthy, G., 1995. Face-sensitive regions in human extrastriate cortex studied by functional MRI. J. Neurophysiol. 74, 1192–1199.

Rämä, P., Sala, J.B., Gillen, J.S., Pekar, J.J., Courtney, S.M., 2001. Dissociation of the neural systems for working memory maintenance of verbal and nonspatial visual information. Cogn. Affect. Behav. Neurosci. 1, 161–171.

Rämä, P., Poremba, A., Yee, L., Malloy, M., Mishkin, M., Courtney, S.M., 2004. Dissociable functional cortical topographies for working memory maintenance of voice identity and location. Cereb. Cortex 14, 768–780.

Rauschecker, J.P., Tian, B., 2000. Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc. Natl. Acad. Sci. U. S. A. 97, 11800–11806.

Romanski, L.M., Goldman-Rakic, P.S., 2002. An auditory domain in primate prefrontal cortex. Nat. Neurosci. 5, 15–16.

Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S., Rauschecker, J.P., 1999. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat. Neurosci. 12, 1131–1136.

Sala, J.B., Courtney, S.M., 2002. Working memory for the conjunction of pattern identity and location. NeuroImage 16(2), 724–725.

Sala, J.B., Rämä, P., Courtney, S.M., 2003. Functional topography of a distributed neural system for spatial and nonspatial information maintenance in working memory. Neuropsychologia 41, 341–356.

Schumacher, E.H., Lauber, E., Awh, E., Jonides, J., Smith, E.E., Koeppe, R.A., 1996. PET evidence for an amodal verbal working memory system. NeuroImage 3, 79–88.

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

Shah, N.J., Marshall, J.C., Zafiris, O., Schwab, A., Zilles, K., Markowitsch, H.J., Fink, G.R., 2001. The neural correlates of person familiarity. A functional magnetic resonance imaging study with clinical implications. Brain 124, 804–815.

Smith, E.E., Jonides, J., 1999. Storage and executive processes in the frontal lobes. Science 283, 1657–1661.

Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain. Thieme, New York.

Tian, B., Reser, D., Durham, A., Kustov, A., Rauschecker, J.P., 2001. Functional specialization in rhesus monkey auditory cortex. Science 292, 290–293.

Ungerleider, L.G., Haxby, J.V., 1994. What and where in the human brain. Curr. Opin. Neurobiol. 4, 157–165.

Ungerleider, L.G., Mishkin, M., 1982. Two cortical visual systems. In: Ingle, D.J., Goodale, M.A., Mansfield, R.J.W. (Eds.), Analysis of Visual Behavior. MIT Press, Cambridge.

Ward, B.D., 2000. Simultaneous Inference for fMRI Data. http://afni.nimh.nih.gov/pub/dist/doc.

Woods, R.P., Grafton, S.T., Holmes, C.J., Cherry, S.R., Mazziotta, J.C., 1998. Automated image registration: I. General methods and intrasubject, intramodality validation. J. Comput. Assist. Tomogr. 22, 139–152.