THE COMMONALITY OF NEURAL NETWORKS FOR VERBAL AND VISUAL SHORT-TERM MEMORY

Steve Majerus[12], Arnaud D'Argembeau[12], Trecy Martinez Perez[12], Sanaâ Belayachi[1], Martial Van der Linden[13], Fabienne Collette[124], Eric Salmon[4], Ruth Seurinck[5], Wim Fias[5] & Pierre Maquet[24]

[1] Center for Cognitive and Behavioural Neuroscience, Université de Liège

[2] Fund for Scientific Research – FNRS, Belgium

[3] Cognitive Psychopathology and Neuropsychology Unit, Université de Genève

[4] Cyclotron Research Center, Université de Liège

[5] Faculty of Psychology and Educational Sciences, Universiteit Gent

Address for correspondence

Steve Majerus, PhD

Center for Cognitive and Behavioural Neuroscience

Université de Liège

Boulevard du Rectorat, B33, 4000 Liège, Belgium

Tel: 003243664656

Fax: 003243662808

Email: smajerus@ulg.ac.be

ABSTRACT

While many neuroimaging studies have considered verbal and visual short-term memory (STM) as relying on neurally segregated short-term buffer systems, the present study explored the existence of shared neural correlates supporting verbal and visual STM. We hypothesized that networks involved in attentional and executive processes as well as networks involved in serial order processing underlie STM for both verbal and visual list information, with neural specificity restricted to sensory areas involved in processing the specific items to be retained. Participants were presented sequences of nonwords or unfamiliar faces, and they had to maintain and recognize order or item information. For encoding and retrieval phases, null conjunction analysis revealed an identical fronto-parieto-cerebellar network comprising the left intraparietal sulcus, bilateral dorsolateral prefrontal cortex and the bilateral cerebellum, irrespective of information type and modality. A network centered around the right intraparietal sulcus supported STM for order information, in both verbal and visual modalities. Modality-specific effects were observed in left superior temporal and mid-fusiform areas associated with phonological and orthographic processing during the verbal STM tasks, and in right hippocampal and fusiform face processing areas during the visual STM tasks, these modality effects being most pronounced when storing item information. The present results suggest that STM emerges from the deployment of modality-independent attentional and serial ordering processes towards sensory networks underlying the processing and storage of modality specific item information.

226 words

INTRODUCTION

The tripartite working memory model developed by Baddeley and Hitch (1974) has been a major conceptual framework for short-term memory (STM) research over the past 25 years. Following this model, separate and specialized short-term buffers are responsible for the storage of verbal and visuo-spatial information, these buffers being independent from sensory and long-term memory processes. Common processes are also assumed to exist, however these processes are supposed to reflect domain independent executive processes such as involved in the coordination and manipulation of information held in STM. In other words, common executive processes are supposed to intervene mainly in STM tasks with high executive load such as working memory tasks, but not in simple passive short-term storage tasks. The results of early neuroimaging studies appeared to support this assumed specificity of verbal and visual STM systems, with activation in the left inferior frontal gyrus (Broca's area) and the supramarginal gyrus ascribed to verbal short-term articulatory rehearsal and buffer systems, and activation in right occipito-parietal and occipito-temporal junctions ascribed to spatial and visual short-term buffer systems, respectively (Frackowiak, 1994; Jonides et al., 1993; Paulesu, Frith, & Frackowiak, 1993; Salmon et al., 1996; Smith & Jonides, 1995).

Recent functional neuroimaging studies however have raised considerable doubts about the specificity of neural correlates for verbal and visual STM. With respect to verbal STM, the search for a dedicated verbal short-term buffer system has proven elusive. A number of studies have shown that the left supramarginal gyrus, initially proposed to subtend the verbal STM buffer (Paulesu et al., 1993), is sensitive to the phonological processing requirements of verbal information during a STM task rather than to verbal short-term storage *per se* (Becker, MacAndrew, & Fiez, 1999; Martin, Wu, Freedman, Jackson, & Lesch, 2003; Ravizza, Delgado, Chein, Becker, & Fiez, 2004; Zatorre, Evans, Meyer, & Gjedde, 1992; see also Buchsbaum & D'Esposito, 2008, for a recent review). More generally, short-term storage and sensory processing appear to be intimately related, and this for both visual and verbal STM domains. An increasing number of studies have shown that, for verbal STM tasks, language processing regions in superior, middle and inferior temporal areas are not only involved during

initial identification and encoding of the verbal stimuli, but they also show some continued activation during the retention delay, at least during early maintenance stages (Fiebach, Friederici, Smith, & Swinney, 2007; Ruchkin et al., 1999; Pa, Wilson, Pickell, Bellugi, & Hickok, 2008). The same has been observed for visual STM: in STM tasks entailing the presentation of sequences of unfamiliar faces, fusiform and inferior temporo-occipital areas involved in face identification and representation processes show continued activation during maintenance (Postle, Druzgal, & D'Esposito, 2003; Postle, 2005; see also Fuster, 1999). These data suggest that verbal and visual representational systems are actively contributing to verbal and visual STM. Furthermore, regions that have been shown to be sensitive to STM load and which are located outside sensory processing areas are in fact very similar in verbal and visual STM tasks. The left intraparietal sulcus has been shown to be sensitive to STM load in both verbal and visual STM tasks (e.g., Ravizza et al., 2004; Todd & Marois, 2004; Todd, Fougnie, & Marois, 2005). Studies having directly compared verbal and visual STM tasks also observed a common involvement of the intraparietal sulcus as well as dorsolateral prefrontal areas (Brahmbhatt, McAuley, & Barch, 2008; Hautzel et al., 2002; Lycke, Specht, Ersland, & Hugdahl, 2008; Nystrom et al., 2000; Rämä et al., 2001).

In contrast to the tripartite working memory model, these recent studies suggest that (1) possible specificities in neural networks for verbal and visual STM are restricted to the specific sensory processing areas recruited when processing the specific information to be stored rather than to the existence of modality-specific STM buffers distinct from sensory processing, and (2) common neural networks in extra-sensory regions support verbal and visual STM. These assumptions do not fit the predictions of the standard working memory model. However, many recent STM models, including a revised form of the working memory model, assume that long-term and sensory knowledge are actually a major determinant of STM performance (e.g., Baddeley, 2000; Burgess & Hitch, 2005; Gupta, 2003; Martin & Saffran, 1992). This is also supported by behavioural studies showing that recall performance in STM tasks is strongly dependent upon the availability of long-term memory knowledge: in the verbal STM domain, recall performance is consistently higher for word list recall as compared to nonword list recall or for phonologically familiar but meaningless verbal sequences as compared to phonologically unfamiliar verbal sequences (Hulme, Maughan & Brown, 1991; Majerus & Van der Linden,

2003; Majerus, Van der Linden, Mulder, Meulemans, & Peters, 2004). Other recent theoretical models argue that common attentional factors determine verbal and visual STM tasks, and this even for tasks having very low demands on working memory type executive processes (e.g., Cowan et al., 2005; Fuster, 1999). In these views, STM is the product of temporarily activated long-term memory representations, which, for Cowan, are held within the focus of attention, and, for Fuster, are updated and reorganized as a function of ongoing task requirements. In line with these assumptions, recent neuroimaging studies showed that the left intraparietal sulcus is sensitive to domain general attentional factors during a STM task (Todd & Marois, 2004; Todd et al., 2005). Todd and colleagues observed that activation in the left intraparietal sulcus increased as a function of STM load, and increase of STM load was associated with a decreased ability to detect irrelevant visual stimuli briefly presented during the retention delay, suggesting a competition of attentional resources between the STM task and visual perception processes.

Although these newer theoretical proposals can account for the impact of sensory processing areas in STM tasks and predict the recruitment of common fronto-parietal attentional / executive networks during verbal and visual STM, some of these theoretical frameworks contain important additional specifications. These specifications, as we will show, lead to new predictions concerning the neural networks associated with verbal and visual STM, and allow us to refine the current assumptions about the respective impact of long-term memory and attentional processes during STM tasks. Although these specifications have been developed most explicitly in the context of verbal STM models, they can be readily transposed to the visual STM domain. Many recent STM models make indeed a critical distinction between item and serial order STM processes, item STM referring to the retention of the items and their linguistic and visual features, while order STM refers to the coding and retention of the serial positions in which the items have occurred (e.g., Burgess & Hitch, 1999; Gupta, 2003). Importantly, behavioural studies have shown that the impact of long-term knowledge on STM recall is strongest for recall of item information while recall of order information is relatively insensitive to the impact of long-term knowledge (e.g., Nairne & Kelley, 2004). For example, when comparing STM list recall for highly frequent versus less frequent words, or for semantically similar versus dissimilar words, a consistent

recall advantage is observed for item recall (as estimated by the proportion of item errors), but less so for order recall (as estimated by the proportion of order errors) (Nairne & Kelley, 2004; Saint-Aubin & Poirier, 2000). In other words, recall of item information is strongly dependent upon the availability of long-term and sensory knowledge. On the other hand, order STM is supposed to be supported by a specialized system for maintaining order information. This system interacts with the language system and monitors the order of activation events in the language network (e.g., Gupta, 2003). This system is also proposed to underlie the learning of new verbal sequences, a prediction which is borne out by a number of recent studies linking order but not item STM measures to the ability to acquire new verbal sequences (Majerus et al., 2006a, 2008). This distinction between long-term memory-dependent item STM processes and specific order STM processes leads to a number of new hypotheses concerning the neural architecture of verbal and visual STM.

First, if, as assumed, modality-specific effects in verbal and visual STM tasks are due to the recruitment of specific sensory and long-term knowledge bases, these modality-specific effects should be most pronounced for the retention of item information, since item recall has been shown to be dependent upon access to sensory and long-term memory representations, but less order recall. No study has yet compared item and order retention processes when contrasting verbal and visual STM tasks. A second question relates to the neural correlates of recall for serial order information: given its independence from long-term memory and sensory representations and given its supposed specificity for a dedicated STM system, order recall should most strongly highlight STM specific networks. Recent neuroimaging studies comparing item and order short-term recognition for verbal sequences observed a specific recruitment of a network including the anterior right intraparietal sulcus and the superior frontal gyrus for order retention, as compared to item retention (Henson, Burgess, & Frith, 2000; Majerus et al., 2006b, 2008; Marshuetz et al., 2000). These regions have been associated with temporal processing, temporal grouping and order judgment, these processes being highly relevant for serial position coding and recognition (Rao, Mayer, & Harington, 2001; Turconi, Jemel, Rossion, & Seron, 2004). With respect to visual STM, the important question that arises here is to which extent visual STM may rely on a similar network for storing serial position information. Behavioural studies suggest that identical serial position coding

effects are observed in verbal and visual serial order reconstruction tasks, suggesting the existence of common mechanisms for storing verbal and visual sequence information (e.g., Smyth, Dennis, & Hitch, 2005). Majerus et al. (2007), exploring item and order short-term recognition for visual sequences (sequences of unfamiliar faces) also observed a specific recruitment of the right intraparietal sulcus for the order retention condition, echoing the earlier results observed during a verbal order retention task (Majerus et al., 2006). However, although these studies suggest the possible existence of common mechanisms and neural networks for storing serial order information in the verbal and the visual domain, a direct comparison of networks involved in serial order processing in verbal and visual STM tasks still has to be conducted. Finally, with respect to theoretical accounts of STM considering that domain general focused attention and associated executive control capacities determine any kind of STM task, we expected an identical involvement of the left intraparietal sulcus and the dorsolateral prefrontal cortex in verbal and visual STM tasks, and this independently of item or order information to be retained. As shown before, increased activation of the intraparietal sulcus is associated with a decreased availability of attention towards task-irrelevant stimuli (Todd et al., 2005). Majerus et al. (2006b, 2007, 2008) also observed a similar involvement of the left (but not right) intraparietal sulcus and bilateral dorsolateral prefrontal cortex in verbal and visual short-term recognition experiments for item or order information, suggesting that this area has a domain general function in STM task. More generally, the superior parietal lobule, including the anterior intraparietal sulcus and the dorsolateral prefrontal cortex, have been shown to be consistently involved in oriented and executive attention networks (see Raz & Buhle, 2006, for a review).

In sum, the present study aimed at studying the neural correlates of verbal and visual STM by hypothesizing that (1) a fronto-parietal network including the left intraparietal sulcus and the dorsolateral prefrontal cortex underlies executive and attentional processes during verbal and visual STM, for the retention of order and item information, (2) a fronto-parietal network centered around the right intraparietal sulcus underlies the processing of order information, for verbal and visual sequences, and (3) modality-dependent activations are most pronounced for item STM processes, involving the recruitment of the respective sensory networks necessary for processing the specific item information. We assessed visual STM via the presentation of a

spatially organized sequence of four unfamiliar face stimuli, followed by two probe stimuli after a variable and long retention delay. A spatial sequence rather than a sequence of temporally discrete events was chosen in order to allow comparison with previous experiments that explored order STM processes in the visual domain and which also used spatially organized sequences (Majerus et al., 2007; see also Marshuetz et al., 2000, for a similar design in the verbal domain). Furthermore, pilot testing had established that behavioural performance was most likely to be matched between the different STM conditions when using spatially organized sequences; with temporally organized sequences, and for the type of unfamiliar stimuli used in the present experiment (see below), behavioural performance had been systematically lower in the visual *versus* verbal conditions. For visual STM, we used unfamiliar face stimuli given the well defined sensory networks, located in right fusiform, hippocampal and temporal areas, which are associated with the processing of novel face stimuli (e.g., Henson et al., 2003). Verbal STM was assessed using nonword stimuli; these stimuli were pronounceable but unfamiliar like the faces used in the visual STM experiment. Here also we had strong a priori knowledge about the target sensory networks, including mainly left-sided regions associated with phonological and orthographic processing, situated in the superior, middle and inferior temporal gyri as well as the inferior frontal gyrus (Broca's area) (Binder et al., 2000; Majerus et al., 2002; Price & Devlin, 2003; Scott et al., 2000). For each STM modality, the participants were instructed to focus on and recognize either sequence information (the serial position of each stimulus) or item information (item identity). A first model investigated activation across the four STM conditions as a function of STM phase by modeling encoding, maintenance and retrieval phases as separate regressors, assuming statistical independence between the three events. However, some studies suggest that encoding and maintenance activity may not be entirely independent, maintenance activity being characterized by ongoing activity of regions initially activated during encoding, or by activation starting shortly before retrieval and then continuing over the retrieval phase (Pa et al., 2008; Rowe et al., 2000). In order to account for this possibility, a second model explored sustained activity over the entire STM trials, and allowed us to analyze the time course of activity during STM trials considered as blocks.

METHODS

Participants

Twenty-three right-handed native French-speaking young adults (10 female), with no diagnosed psychological or neurological disorders, were recruited from the university community. The study was approved by the Ethics Committee of the Faculty of Medicine of the University of Liège, and was performed in accordance with the ethical standards described in the Declaration of Helsinki (1964). All participants gave their written informed consent prior to their inclusion in the study. Age ranged from 19 to 30 years, with a mean of 23.04 years. Minimal number of years of education was 14.

Task description

The structure of the task used in the present study was based on serial order and item STM probe recognition tasks that have been previously shown to reliably differentiate neural networks involved in serial order and item STM processes (Majerus et al., 2006, 2007). For each trial, the encoding phase consisted of the presentation of a list of four faces or four nonwords ordered horizontally (fixed duration: 4000 ms), followed by a maintenance phase indicated by the display of a fixation cross (variable duration: random Gaussian distribution centered on a mean duration of 7250 ± 2000 ms). The retrieval phase consisted of an array of two probe stimuli ordered vertically. Participants indicated within 3000 ms if item or order information for the two probe stimuli matched information in the memory list (by pressing the button under the third finger) or not (by pressing the button under the index) (see Figure 1 for further details on stimulus duration and timing). More specifically, in the order condition, the participants judged whether the probe stimulus presented on the top of the screen had occurred in a more leftward position (relative to the position of the two stimuli in the memory list) than the probe face presented on the bottom of the screen. In the item condition, the probe stimuli were twice the same stimulus (in order to match the amount of information displayed in the order and item retrieval phases) and the participants judged whether the probe stimuli were identical to one of the stimuli in the memory list.

The nonwords and faces used in the different trials were pseudo-randomly sampled from a pool of 60 nonwords or unfamiliar faces. The nonwords were constructed by selecting first a set of 30 pairs of mono- or disyllabic words differing by a single consonant, and by replacing this consonant by a different consonant resulting in a nonword; this procedure ensured that we obtained 30 minimal pairs of unfamiliar yet easily pronounceable verbal stimuli. The digram and diphone frequencies of the nonword segments differing between the two nonword sets were matched (4431.56 *versus* 4336.53, t(59)<1, n.s., for digram frequencies following the Lexique database, New, Pallier, Brysbaert, and Ferrand, 2006; 877.99 *versus* 985.43, for diphone frequencies following the phonetic database by Tubach & Boë, 1990). The unfamiliar face stimuli were constructed following a similar procedure. The faces were selected from a database of faces of American background (FERET database, Phillips, Wechsler, Huang, & Rauss, 1998). By means of the software MorphEditor (SoftKey Corporation, Cambridge, MA), pairs of morphed faces were obtained by incorporating the facial features of one "master" face into two other faces, so that the two faces had the features taken for this "master" face in common. 30 pairs of faces having 55 % of features in common were obtained. This was done in order to obtain pairs of faces that differed very minimally, further reducing the likelihood of verbal encoding. This also enabled us to construct negative probes that differed only very minimally from the target stimulus, requiring the participants to maintain very detailed item representations: negative probe trials consisted in the presentation of one member of the nonword or face pair in the memory list and the other member in the probe array. In the order conditions, the probe trials always contained two adjacent stimuli of the target stimulus list, but they were presented either in the same or a reversed ordering (see Figure 1). By probing adjacent but not distant positions, we were able to maximize the difficulty and sensitivity of the order STM condition as very precise order representations are needed when probing two adjacent items (see also Henson et al., 2000; Marshuetz et al., 2006, 2007). There were an equal number of positive and negative probe trials, probing equally all item positions.

A baseline condition, controlling for nonword identification and perceptual face analysis processes as well as for motor response and stimulus decision processes not of interest in this study, consisted of the presentation of a sequence containing two identical nonwords and two identical face stimuli ordered

horizontally, followed by a delay interval (a fixation cross of variable duration) and a response display showing twice the same face or nonword stimuli, with the two stimuli oriented in the standard way or one stimulus showing upside down (see Figure 1); the participants had to decide whether the stimuli showed typical or reversed orientation by pressing the buttons under the third finger (for 'yes' responses) or under the index (for 'no' responses).

< INSERT FIGURE 1 ABOUT HERE >

The four STM conditions and the baseline condition were presented in a single session, using an event-related design. There were 30 trials for each STM condition and 20 trials for the baseline condition. The different trials were presented in pseudo-random order, with the restriction that two successive trials of the same condition could not be separated by more than 5 trials of a different condition (i.e., by more than 82 seconds on average), in order to keep BOLD signals for same condition epochs away from the lowest frequencies in the time series (see below). The variable maintenance delay between the encoding and retrieval phases ensured minimal temporal autocorrelation between the encoding and retrieval phases, assuming these phases are independent (Cairo, Liddle, Woodward, & Ngan, 2004; see also below for further technical details). Before the start of a new trial, a brief instruction appeared on the centre of the screen informing the participant what type of information s/he had to retain (order trials: "remember the order"; item trials: "remember the identity"; control trials: "look"). In order to avoid confusion between the different STM conditions, the instruction remained visible during the encoding and maintenance phases where it was displayed at the highest position of the screen. The duration of the inter-trial interval was also variable (random Gaussian distribution centered on a mean duration of 2000 ± 200 ms) and further varied as a function of the participants' response times: the probe array disappeared immediately after pressing the response button, followed by the presentation of the next trial. If the participant did not respond within 3000 ms, a 'no response' was recorded and the next trial began. A practice session outside the MR environment, prior to starting the experiment, familiarized the participants with the specific task requirements and presentation rate, by administering at least 8 practice trials for each item and order STM conditions.

MRI acquisition

Data were acquired on a 3Tesla scanner (Siemens, Allegra, Erlangen, Germany) using a T2* sensitive gradient echo EPI sequence (TR = 2130 ms, TE = 40 ms, FA 90°, matrix size 64 X 64 X 32, voxel size 3.4 X 3.4 X 3.4 mm³). Thirty-two 3-mm thick transverse slices (FOV 22 X 22 cm²) were acquired, with a distance factor of 30%, covering nearly the whole brain. Structural images were obtained using a T1-weighted 3D MP-RAGE sequence (TR = 1960 ms, TE = 4.4 ms, FOV 23 X 23 cm², matrix size 256 X 256 X 176, voxel size 0.9 X 0.9 X 0.9 mm). In each session, between 1098[1] and 1158 functional volumes were obtained. Head movement was minimized by restraining the subject's head using a vacuum cushion. Stimuli were displayed on a screen positioned at the rear of the scanner, which the subject could comfortably see through a mirror mounted on the standard head coil.

fMRI analyses

Data were preprocessed and analyzed using SPM5 software (Wellcome Department of Imaging Neuroscience, http//www.fil.ion.ucl.ac.uk/spm) implemented in MATLAB (Mathworks Inc., Sherbom, MA). Functional scans were realigned using iterative rigid body transformations that minimize the residual sum of square between the first and subsequent images. They were normalized to the MNI EPI template (voxel size: 2x2x2mm) and spatially smoothed with a Gaussian kernel with full-width at half maximum (FWHM) of 8 mm (in order to minimize noise and to assure that the residual images conform to a lattice approximation of Gaussian random fields).

For each subject, brain responses were estimated at each voxel, using a general linear model with epoch regressors. In a first model, separate epoch durations were defined to cover encoding, maintenance and retrieval phases, permitting the modeling of phase specific STM-related brain activity for each of the four STM conditions. The encoding regressor ranged from the time of the onset of each trial until the onset of the fixation

---

[1] For one participant, 908 functional volumes were obtained due to premature ending of the session by the participant. Discarding this participant from analyses did not change the outcome of results. The results presented here include this participant.

cross of the maintenance interval; the maintenance regressor ranged from the onset of the fixation cross until the onset of the probe display; the retrieval regressor ranged from the onset of the probe display to the participant's response. In order to ensure minimal autocorrelation between the three phase-specific regressors, the maintenance regressor was further orthogonalized relative to the other two regressors using the same procedure as in Majerus et al. (2006, 2007): shared variance between the encoding and early maintenance phases was attributed to the encoding regressor, and shared variance between retrieval and late maintenance phases was attributed to the retrieval regressor. The fact that this model includes three orthogonal regressors assumes that the three STM phases are independent, and that activation during the maintenance phase is independent from activation in the encoding and retrieval phases. As we have noted, this might however not be entirely accurate given that maintenance processes are likely to start during encoding and that activation in sensory and language processing areas during encoding has been shown to continue during maintenance, at least during early maintenance stages (e.g., Postle et al., 2003; Pa et al., 2008; Ruchkin et al., 2003). Furthermore, general task-related attentional processes might show sustained activation during the entire STM trial. For these reasons, a second model was constructed measuring sustained activity during the four STM conditions, by defining an epoch regressor covering the entire trial, from the onset of display of the memory list until the participants' response to the probe stimuli. Time course analyses were then used to determine hemodynamic response as function of encoding, maintenance and retrieval stages. For both models, the baseline condition was modeled implicitly.

For each model, boxcar functions representative for each regressor and each STM condition were convolved with the canonical hemodynamic response. The design matrix also included the realignment parameters to account for any residual movement-related effect. A high pass filter was implemented using a cut-off period of 128s in order to remove the low frequency drifts from the time series. Serial autocorrelations were estimated with a restricted maximum likelihood algorithm with an autoregressive model of order 1 (+ white noise). For the phase-specific model, twelve linear contrasts corresponding to each cell of the design were defined (four STM condition x three STM phases). For the sustained effect model, four linear contrasts

corresponding to each STM condition were defined. The resulting set of voxel values constituted a map of t statistics [SPM{T}]. These contrast images were then smoothed again (6-mm FWHM gaussian kernel) in order to reduce remaining noise due to inter-subject differences in anatomical variability in the individual contrast images. They were then entered in second-level analyses, corresponding to ANOVA random effects models. A first 4 (STM conditions) by 3 (STM phases) ANOVA assessed main effects, main differential effects and the interactions between conditions and phases. Null conjunction analyses assessed the commonality of activation profiles across conditions and phases (Friston et al., 2005). This type of conjunction analyses is a conservative method to estimate effects that are present in all conditions of interest, and, contrary to global conjunction analyses, limits the risk of observing of false positives (regions appearing to be activated in common over different conditions but being driven mainly by one condition where these regions show a particularly high level of activation). A second ANOVA (4 STM conditions) assessed sustained effects in the four STM conditions, with null conjunction analyses assessing the commonality of activation profiles across conditions. As a rule, statistical inferences were performed at the voxel level at $p < 0.05$ corrected for multiple comparisons across the entire brain volume. When *a priori* knowledge was available about the potential response of a given area in our different STM conditions, a small volume correction (Worsley, Marrett, Neelin, & Evans, 1996) was computed on a 10-mm radius sphere around the averaged coordinates published for the corresponding location of interest (see below).

A priori locations of interest

A small number of *a priori* locations of interest were used for small volume corrections, based on published coordinates in the literature for STM recognition tasks similar to that used in the present study, as well as on the results obtained in previous studies contrasting item and order STM. These regions of interest included the bilateral IPS, but also bilateral premotor, dorsolateral prefrontal, subcortical and cerebellar regions which are consistently activated in STM recognition tasks. Other regions of interest concerned more specifically activation in areas in the ventral occipital, hippocampal and fusiform cortex associated with face perception and

14

recognition and which we hypothesized to be specifically recruited in the face item STM condition. Similarly, for the verbal item STM condition, we hypothesized the recruitment of left superior temporal and inferior temporal / mid-fusiform areas associated with phonological and orthographic processing. We only report here the coordinates for those regions where significance thresholds did not resist corrections for whole brain volume and for which small volume corrections were actually performed. All stereotactic coordinates refer to the MNI space. The *a priori* locations of interest were the following:

Order STM: superior frontal gyrus [24, 10, 56] (Majerus et al., 2006, 2007); right IPS [48, -40, 44] (Majerus et al., 2006, 2007, 2008);

Verbal item STM : inferior frontal gyrus (phonological and articulatory processing [-54, 6, 18; Majerus et al., 2002], superior temporal gyrus and superior temporal sulcus [-53, -26, 9; -59, -30, -1] (phonological processing: Binder et al., 2000, Scott, Blank, Rosen, & Wise, 2000, Majerus et al., 2006) ; mid-fusiform and inferior temporal [-42, -57, -6] (orthographic processing: Price & Devlin, 2003)

Visual item STM; hippocampus [24, -12, -18] (Henson et al., 2003)

STM (general): SMA [0, 18, 54] (Majerus et al., 2006, 2007), middle frontal gyrus [-50, 26, 32 ; 46, 36, 22] (Cairo et al., 2004 ; Majerus et al., 2006, 2007 ; Ravizza et al., 2004); left IPS [-41, -46, 47] (Majerus et al., 2006, 2007); caudate [-10, -4, 24; -12, 20, -8; -26, -31, 22; 24, -32, 12 ; -20, -42, 14 ; 8, 4, 22] (Cairo et al., 2004; Majerus et al., 2006, 2007; Ravizza et al., 2004).

RESULTS

Behavioral data

Mean accuracy levels for the four STM conditions were very close (see Table 1); the only significant difference was a slightly higher accuracy for the verbal order STM condition relative to the visual order STM condition, $t(22)=2.24$, $p=.04$, all other paired t-tests being non-significant at $p<.05$. At the level of response latencies, the participants took longer to respond in the order STM conditions relative to the item conditions, as we did expect, due to the recruitment of serial order scanning processes during retrieval in the order conditions (verbal order vs.

verbal item, t(22)=9.88, p<.001, visual order vs. visual item, t(22)=11.98, p<.001). The only other difference were even higher response times in the visual order as compared to the verbal order conditions, t(22)=2.97, p<.01.

< INSERT TABLE 1 ABOUT HERE>

Imaging data

Transient STM effects

First, we report the main effect for each of the four STM conditions (relative to the implicitly modeled baseline condition), as a function of STM phase. As shown in Figure 2, overall similar fronto-parieto-cerebellar networks were observed across the four STM conditions, during encoding and retrieval. We also observed the involvement of left inferior frontal cortex (Broca's area) and left superior, middle and inferior temporal areas during the verbal STM conditions, and the involvement of right occipital and fusiform regions during the visual STM conditions. During maintenance however, activation patterns appeared to be much less pronounced. Next, we assessed the commonality of these activation patterns across the four STM conditions more formally, using null conjunction analysis, as a function of encoding, maintenance and retrieval stages. During encoding, a fronto-parieto-cerebellar network was observed to be activated across the four conditions, including, bilaterally, the supplementary motor area, the anterior middle prefrontal gyrus, the caudate nuclei, the superior lateral cerebellum, and in the left hemisphere, the anterior intraparietal sulcus and the anterior inferior frontal gyrus and insula (see Table 2). During the maintenance phase, only a small area within the left precentral gyrus appeared to be activated in common across the four conditions (at p<.001, uncorrected); this general absence of common activation during maintenance was due to an absence of specific activation patterns during the maintenance phase, as already noticed. For the retrieval phase, an extensive fronto-parieto-cerebellar network was again observed to be activated across all four STM conditions, including, bilaterally, the supplementary motor area, the inferior frontal gyrus, the anterior middle frontal gyrus, the anterior intraparietal sulcus, the superior cerebellum. In sum, an identical fronto-parieto-cerebellar network was observed to be activated across the four STM conditions, during encoding and retrieval.

< INSERT TABLE 2 AND FIGURE 2 ABOUT HERE >

Next, we assessed condition specific effects. First, we investigated main differential effects, by contrasting order and item conditions, as well as verbal and visual conditions, independently of STM phase. We also computed null conjunction analyses on these comparisons, in order to determine differential brain activity that is common to both contrasts underlying the respective differential main effects (see Table 3 for details). When comparing the order to the item STM conditions, greater activation was observed in the bilateral intraparietal sulci, the bilateral precuneus, the superior occipital gyri and the superior and middle frontal gyri. The conjunction analyses on the two contrasts underlying this differential effect revealed a more restricted network, involving the right but not the left intraparietal sulcus, the right precuneus, the right superior occipital gyrus and the bilateral superior frontal gyri. The reverse comparison, i.e., item *versus* order STM conditions, failed to reach statistical significance. When investigating the main differential effects for verbal *versus* visual conditions, an exclusively left-hemisphere network involving Broca's area, adjacent premotor and motor cortex, the superior and inferior temporal gyri as well as the temporo-parietal junction was observed. Conjunction analyses highlighted a very similar set of brain areas, except for the superior temporal areas and the temporo-parietal junction. Finally, the reverse effect, i.e., visual *versus* verbal STM conditions, highlighted a network involving the right fusiform gyrus, the bilateral posterior middle temporal area close to the angular gyrus, the bilateral inferior temporal gyrus, the right hippocampus, the medial and superior frontal gyri and the precuneus. Conjunction analyses restricted these activations to the right fusiform gyrus, the right inferior temporal gyrus and the bilateral posterior middle temporal area. As hypothesized, these results show that the comparison of verbal *versus* visual STM conditions generates effects in left fronto-temporal areas specialized in language processsing, and the reverse comparison generates effects in fusiform, hippocampal and inferior temporal areas specialized in face processing. Furthermore, these analyses show that order processing, as compared to item processing, is subtended by an identical network involving the right intraparietal sulcus, precuneus and superior frontal gyrus. However, these analyses do not preclude the possibility of further condition specific effects. We had hypothesized that the activation of language and face processing areas, although involved in both item and

17

order STM conditions, should be strongest in item STM conditions. These effects will be specific to each item STM condition, and thus cannot be properly revealed by main differential effects analyses reflecting global differential activation due to either the nonword item condition, the face item condition or both. Furthermore, the recruitment of language and face processing as well as attentional and order processing networks is likely to interact with STM phase, as suggested by the results of the main effects analyses, revealing contrasted activation profiles between encoding, maintenance and retrieval stages.

<center>< INSERT TABLE 3 ABOUT HERE ></center>

Further condition-specific effects were investigated by an interaction analysis involving the four STM conditions and the three STM phases. As shown in Table 4, a large set of regions showed condition-by-STM phase interactions. A first set of regions showed greater activation during the two verbal STM conditions, and this specifically during retrieval and/or encoding but not maintenance; these regions were situated in Broca's area and adjacent sensori-motor cortex, in superior and inferior temporal regions typically associated with articulatory, phonological and semantic language processes, in the left inferior occipital gyrus and in the right lateral superior cerebellum (see Figure 3 and Table 4). Importantly, there was a bilateral posterior middle temporal region which showed particularly high deactivation during encoding and retrieval of verbal item information, relative to the verbal order STM condition ($F(1,22)=21.82$, $p<.0001$, and $F(1,22)=15.03$, $p<.0001$, for the left and right middle temporal regions, respectively [2]), as well as relative to the visual item and visual order STM conditions ($F(1,22)=65.28$, $p<.0001$, and, $F(1,22)=106.14$, $p<.0001$, for the left and right middle temporal regions, respectively). A second set of regions showed specifically greater activation during the visual STM encoding and retrieval conditions, including the anterior superior frontal and orbito-frontal regions, the left and right fusiform gyrus, the right hippocampal gyrus, the right inferior occipital gyrus, the bilateral precuneus and the bilateral lateral cerebellum. Importantly, among these regions, the hippocampal activation was significantly stronger in the visual item STM condition than in the visual order condition ($F(1,22)=5.45$, $p<.05$)

---

[2] The region specific F-tests reported here compare condition specific estimates of beta coefficients extracted for the regions of interest, based on the coordinates shown in Table 4.

as well as relative to the two verbal STM conditions (F(1,22)=52.86, p<.0001). A third set of regions concerned areas recruited to a higher extent in verbal and visual order STM conditions, but again only during encoding and retrieval phases; like in the main differential effect analyses, these effects included the right anterior intraparietal sulcus and the bilateral superior posterior parietal cortex extending to the right superior occipital cortex. Finally, an anterior left intraparietal sulcus area showed activation during encoding and retrieval of all STM conditions, except for the visual item STM condition; this area was more anterior than the left intraparietal sulcus area observed to be activated in common across all four STM conditions, as shown by the previous conjunction analyses.

< INSERT FIGURE 3 AND TABLE 4 ABOUT HERE >

Overall, the results from differential main effect and interaction analyses show that the main differences between verbal and visual STM conditions are related to the recruitment of a fronto-temporal language processing network during verbal STM conditions, and a fronto-fusiform-hippocampal-occipital face processing network during visual STM conditions. As predicted, parts of these information-specific networks were involved to a greater extent in the item conditions, relative to the respective order conditions. With respect to order STM, a network including the right intraparietal sulcus and precuneus, the right superior occipital gyrus and bilateral superior frontal cortices appeared to underlie order STM in both verbal and visual conditions. However, these common activation patterns could nevertheless hide magnitude differences in activation profiles between the verbal and visual order STM conditions. Although the findings from the interaction analyses do not support this possibility, we computed further direct contrasts between the verbal order and visual order STM conditions, by focusing our analyses specifically on the regions revealed to show common activation in the conjunction analyses (i.e., right anterior intraparietal sulcus, precuneus, bilateral superior frontal gyri, superior occipital cortex, superior cerebellum). When comparing the verbal order to the visual order STM condition, no differences in activation magnitude emerged. When performing the reverse contrast, the only difference that emerged was greater activation in the right superior occipital gyrus during the visual order STM condition (x=44, y=-68, z=30, Z=4.12, k=233); no activation magnitude differences were observed in right intraparietal

sulcus, superior frontal cortex, precuneus and cerebellum. More generally, no interaction was observed in regions presumably associated with executive STM processes in dorsolateral prefrontal structures and subcortical structures, as well as with attentional processes in the left anterior intraparietal sulcus, except for a very anterior left intraparietal sulcus region more active in all conditions except for visual item STM.

Finally, the interaction analyses also highlighted the general lack of activation during the maintenance delay in all STM conditions. This finding could however be related to the fact that we modeled the maintenance phase as a completely independent event, relative to encoding and retrieval phases. As we have noted, maintenance might not be independent relative to encoding, and some studies suggest that maintenance is the continuation of encoding processes and decreases over longer maintenance delays (Pa et al., 2008; Ruchkin et al., 1999). In order to account for this possibility, a second set of analyses was conducted modeling each STM trial as a single event in order to capture sustained activity over the entire STM duration, more likely to characterize continuous STM processes.

Sustained STM effects

As shown in Table 5 and Figure 4, analysis of main effects for each STM condition showed again a very similar fronto-parietal-cerebellar network for the visual and verbal order STM conditions, while for the item STM networks, activation patterns appeared to be more modality dependent, encompassing, for the verbal item condition, Broca's area and the left superior temporal sulcus, and, for the face item condition, the right hippocampus. As in the preceding analyses, the commonality of these activations was more formally assessed via null conjunction analyses, and condition specific effects were investigated via differential effect analyses. A first null conjunction analysis over the four STM conditions revealed a fronto-subcortical network including the left middle frontal gyrus and the bilateral caudate nuclei (head and tail) as underlying common sustained activity during the entire trial for each STM condition (see Table 6). A next set of conjunction analyses assessed common sustained activity as a function of item and order STM conditions. For the order STM conditions, a common fronto-parieto-cerebellar network was observed for both visual and verbal conditions, including,

20

bilaterally, the supplementary motor area, the middle frontal gyrus, the anterior intraparietal sulci, the precuneus and, in the right hemisphere, the superior cerebellum. This network of sustained activity is virtually identical to the network underlying transient effects during encoding and retrieval, suggesting that this network is not only involved during order encoding and retrieval, but shows continued activation during maintenance. On the other hand, the network underlying common sustained activation in the two item conditions was much more restricted, including only the left middle frontal gyrus and bilateral caudate nuclei.

< INSERT TABLES 5 AND 6 AND FIGURE 4 ABOUT HERE >

Condition specific effects were assessed via differential main effect analyses for sustained activity during the verbal and visual STM conditions, as a function of item and order conditions (see Table 7). When comparing the verbal to the visual STM conditions, greater sustained activity was observed in a fronto-temporal network, including the inferior frontal cortex (Broca's area), the superior temporal gyrus and the mid-fusiform gyrus, this differential activation appearing furthermore to be slightly more extended in the item STM conditions (based on voxel counts). When computing the reverse contrasts, greater sustained activity was observed in the right fusiform gyrus, this differential activation being much more extended in the item STM conditions, additionally including the left fusiform gyrus, a posterior middle temporal region close to the angular gyrus, the right ventral inferior temporal gyrus, the bilateral hippocampi, as well as the anterior superior frontal area. Furthermore, we also assessed the possibility of modality dependent magnitude differences in activation patterns for regions identified to be activated in common between the verbal and visual order STM conditions. In order to do this, we recomputed the preceding differential effect analyses for the order conditions by restricting the analyses specifically to the fronto-parieto-cerebellar network identified in the conjunction analysis. No voxel reached significance in these analyses, confirming that an identical fronto-parieto-cerebellar network underlies sustained effects for verbal and visual order STM conditions.

< INSERT TABLE 7 ABOUT HERE >

Finally, in order to further characterize the observed sustained activation patterns, the time course for each activation reported in Table 6 was extracted for each condition and each participant, using the peri-stimulus

hemodynamic response function implemented in SPM 5. This function models the hemodynamic response in 16 time points, over a 35 second period starting at the onset of the stimulus. The analyses reported here focused on the 10 first time points (= 21.3 seconds), later activation being contaminated by the presentation of subsequent blocks. As shown in Figure 5a, for most regions of interest and for all four STM conditions, activation increased during the encoding phase, continued during the initial maintenance phase, but rapidly decreased over later maintenance periods, and then increased again during the retrieval phase. We further determined, for each time point, hemodynamic response differences between the different regions of interest, in order to assess whether all regions showed the same hemodynamic response pattern. As shown in Table 8, for the verbal order STM condition, we observed that the cerebellum was more activated compared to all other regions during the very early encoding stage, while the caudate nuclei presented a later onset and more prolonged activity during the maintenance phase; the head of caudate nuclei also showed higher activation than other regions during the late retrieval stage (see also Figure 5b). Similar results were obtained for the visual order STM condition, activation in caudate nuclei peaking later relative to the other regions, and showing somewhat more prolonged activity during the maintenance delay, together with the left precuneus; the right superior frontal gyrus showed higher activation during retrieval, but only relative to the caudate nuclei. For the verbal item STM condition, caudate nuclei again showed later peaks and more prolonged activation during maintenance, especially for the tail of caudate nuclei. Finally, for the visual item STM condition, caudate nuclei also peaked later during maintenance; the left middle frontal gyrus also showed increased activity relative to all other regions during late retrieval.

< INSERT TABLE 8 AND FIGURE 5 ABOUT HERE>

Overall, these time course analyses reveal very similar profiles across the four STM conditions, with activity during maintenance being characterized by ongoing activation starting during the encoding phase, this activity decreasing rapidly at longer maintenance delays. This pattern of time-dependent activation can explain the absence of maintenance specific activity observed when considering transient effects: activity in the maintenance phase decreased when it was supposed to reach its peak according to the transient maintenance

regressor. Only caudate nuclei showed a somewhat more prolonged activity during the maintenance delay, and this equally during the four STM conditions.

DISCUSSION

In contrast to the standard working memory model, predicting the existence of neurally segregated systems for verbal and visual STM, this study considered that common neural systems underlie verbal and visual STM, this commonality however depending upon the type of information, item *versus* serial order, to be maintained. We predicted that a network centered around the left intraparietal sulcus, involved in attentional and executive processes, intervenes in both verbal and visual STM, irrespective of order and item retention processes; a network centered around the right intraparietal sulcus was assumed to be involved in serial order retention processes for verbal and visual order information; modality specific effects were assumed to be restricted to the specific sensory processing areas necessary for representing the items of the memory list. We observed: (1) the activation of a fronto-parieto-cerebellar network centered around the left intraparietal sulcus during encoding and retrieval of item and order information for verbal and visual lists, (2) the activation of a fronto-parieto-cerebellar network centered around the right intraparietal sulcus during encoding, early maintenance and retrieval of order information, in verbal and visual STM conditions, (3) modality specific effects in areas specialized in language and face processing, with a greater involvement of left inferior frontal, superior temporal, inferior temporal and mid-fusiform areas during verbal STM, and a greater involvement of right fusiform and hippocampal areas during visual STM; these modality-specific activations were most pronounced during the item STM conditions, and were observed during encoding, early maintenance and retrieval.

The commonality of fronto-parietal networks centered around the left intraparietal sulcus during verbal and visual STM

A number of previous studies revealed the activation of similar fronto-parietal networks during verbal and visual working memory tasks; the interpretation of these findings, however, remained uncertain given the

use of working memory tasks which are typically very demanding at the level of strategic and executive processes (Baddeley et al., 1998; Brahmbhatt et al., 2008; Hautzel et al., 2002; Lycke et al., 2008; Nystrom et al., 2000). Although the standard working memory model distinguishes verbal and visual short-term storage systems, it assumes the existence of identical executive processes intervening in verbal and visual working memory tasks. Hence, at a theoretical level, the observation of common fronto-parietal networks during verbal and visual working memory tasks is not very informative. The present study, using simple delayed probe recognition tasks, shows that identical fronto-parietal networks are observed to support verbal and visual STM tasks with very minimal working memory requirements. This shows that common fronto-parietal networks centered around the left intraparietal sulcus underlie not only verbal and visual working memory processes, but also verbal and visual short-term retention processes. These results are much more in line with attentional accounts of STM, assuming that common executive and attentional processes underlie the short-term retention of verbal and visual information, but not modality-specific short-term buffers (e.g., Cowan, 1999; Fuster, 1999; Ravizza et al., 2004; Todd et al., 2004, 2005). Postle (2005) also showed, with respect to dorsolateral prefrontal cortex activation, that delay-period activation in dorsolateral prefrontal cortex is involved in sensory gating, helping the cognitive system to focus on the target information to be maintained in a STM task while guarding this information against interference from incoming irrelevant information. This function, very close to the definition of selective attention processes, has also been proposed to subtend left anterior intraparietal sulcus activation in verbal and visual STM tasks (Majerus et al., 2006, 2007; Todd et al., 2004).

It should however be noted here that there was a very anterior portion of the intraparietal sulcus which interacted with STM conditions, being active during encoding across all verbal and visual STM conditions explored in our study, except for the visual item STM condition where only a small portion of the left intraparietal sulcus was shown to be active, suggesting that activation in the left IPS was more extensive in the verbal and visual order conditions as well as in the verbal item condition. A related finding was observed for sustained effects, showing that the fronto-parietal network was also active over the early maintenance phase for the verbal and visual order STM conditions, but less so for the verbal and visual item conditions, where common

24

activation was reduced to a small portion of left dorsolateral prefrontal cortex and caudate nuclei. This finding is in line with a study exploring item STM for faces and showing that face processing areas in the fusiform gyrus but not fronto-parietal regions were active during the retention delay (Postle et al., 2003). These results suggest that attentional control processes may be maintained to a larger extent during list-level retention processes for verbal and visual information, possibly to maintain an ordered and abstract list-level representation and guard it against interference, while during item maintenance processes, these attentional control processes may be involved to a lesser extent while activation is maintained in sensory processing areas necessary for representing item information.

The commonality of fronto-parietal networks centered around the right intraparietal sulcus during verbal and visual order STM

Another novel finding of the present study concerns the neural substrates for order processing in verbal and visual STM tasks. We had hypothesized that a network centered around the right intraparietal sulcus should show specific activation during order STM conditions, and this equally during verbal and visual order STM conditions, based on previous results showing activation in the anterior right intraparietal sulcus, adjacent superior parietal cortex, the superior frontal and lateral orbito-frontal cortex as well as the superior cerebellum during verbal or visual order STM conditions (Henson et al., 2000; Majerus et al., 2006b, 2007, 2008; Marshuetz et al., 2000). We indeed observed a common activation of the right intraparietal sulcus, the superior parietal lobule, the superior frontal gyrus and the superior cerebellum during verbal and visual order STM conditions, during encoding, early maintenance and retrieval. Most importantly, the main differential effect and interaction analyses for transient effects showed that the right intraparietal sulcus was specifically activated during the verbal and visual order STM conditions, but not at all during the verbal and visual item STM conditions. This is in contrast to the left intraparietal sulcus, which showed activity across both item and order STM conditions, although to a lesser extent for the visual item condition, as discussed earlier. This overlap of neural correlates for serial order processing of verbal and visual sequences is also in line with recent behavioural studies showing

identical serial position coding effects during serial order reconstruction tasks for face and word sequences, suggesting the operation of similar mechanisms for representing serial order in the verbal and the visual domain (Smyth et al., 2004). More generally, the right anterior intraparietal sulcus has been shown to be involved in a number of cognitive processes such as magnitude processing, spatial processing and temporal processing, which are all relevant for representing serially ordered information (Corbetta et al., 1995; Marshuetz et al., 2000; Rao et al., 2001; Turconi et al., 2004). Competing theoretical models currently exist, representing serial order information along magnitude, spatial/ positional or temporal dimensions (e.g., Burgess & Hitch, 2005; Gupta, 2003; Henson, 2000). The aim of the present study was not to decide between these different accounts of serial order coding, but rather to determine to what extent verbal and visual STM rely on identical neural networks when processing serial order information. Future studies will need to address how precisely serial order information is represented and maintained within the right intraparietal sulcus.

It could be argued that the current design, using the simultaneous presentation of a spatially organized sequence may have emphasized the use of spatial positional frames for encoding and maintaining serial order information in the visual and verbal STM conditions; the activation we observed in right superior parietal and superior occipital areas is indeed consistently reported in studies involving spatial processing (e.g., Baker et al., 1996; Hautzel et al., 2002; Ungerleider & Haxby, 1994). On the other hand, we think that a simple interpretation in terms of spatial processing is not sufficient to explain the activation pattern in the verbal and visual order STM conditions. First, the spatial organization of target and probe stimuli differed in our study, target stimuli being organized horizontally, and probe stimuli being organized vertically. Hence, the construction of a higher level representation of serial position information was necessary to compare the serial position information contained in probe and target stimuli; a direct match between the visual appearance of target and probe sequences did not enable the participants to make a correct yes/no recognition. This is also illustrated by the reaction time analyses, which were significantly longer for order recognition, as opposed to item recognition, in line with the use of serial scanning and rehearsal strategies (see also Marshuetz et al., 2000). Furthermore, given that attentional focus is limited to a very few items at the time, participants had to serially scan across the four

items despite their simultaneous presentation, in order to switch attentional focus from one item to the other. The right intraparietal sulcus has indeed been involved in these types of serial search and serial attention processes (e.g., Bricolo et al., 2002; Corbetta et al., 1995, 2000; Donner et al., 2000). Finally, in an earlier study using temporally ordered sequences (i.e., one word every 1250 ms) for verbal STM lists, we showed the recruitment of the same right anterior intraparietal sulcus area during order recognition, suggesting that the right intraparietal sulcus plays a specific role during order processing, whether order information is presented sequentially or simultaneously (Majerus et al., 2006b). We should however note that in the present and previous experiments, both verbal and visual stimuli were presented visually, potentially enhancing similarities between order networks involved in verbal and visual STM conditions. It thus remains to be shown whether order STM for auditory-verbal stimuli also shares the same neural substrates as those identified in the studies conducted as so far.

Modality-specific effects are restricted to sensory networks involved in item processing

The distinction between item and order information also allowed us to explore modality specific effects. Item STM is supposed to be directly related to the recruitment of representations in sensory processing networks which also provide the representational basis for encoding and maintenance of the items to-be-stored (e.g., Cowan, 1999; Fuster, 1999; Burgess & Hitch, 1999; Gupta, 2003). Hence, we expected modality specific effects to be restricted to activation in sensory processing networks, and to be strongest for item STM, as opposed to order STM conditions. The data obtained in the present study clearly support these hypotheses since for the verbal condition, modality specific effects were restricted to a left-sided fronto-temporal language processing network; for the visual condition, modality specific effects were restricted to a mainly right-sided fronto-temporo-hippocampal face processing network. As predicted, these modality specific effects were stronger during the item STM conditions as opposed to the order STM conditions. For the verbal STM condition, a bilateral posterior middle temporal region showed a particularly strong item-specific response, this response being characterized by a stronger deactivation during verbal item encoding as opposed to verbal order encoding. Given that this region has been implicated in lexico-semantic processes (Prabhakaran, 2006), a possible

interpretation of this result is that lexico-semantic processes were actively inhibited in this condition, in order to maintain an adequate representation of target nonword items while guarding them against interference coming from phonologically and orthographically similar word forms. The likelihood of this explanation is further supported by fact that the nonwords used in the present study differed from existing word forms by a single sound or letter, and hence the probability that similar sounding words could be activated was very high (Vitevitch & Luce, 1999). In the visual STM condition, fusiform, hippocampal and ventral temporal areas showed a strong item-specific response, and hippocampal areas showed sustained activity during the early maintenance phase, with an especially strong involvement during the item condition as compared to the order condition. Right fusiform, ventral temporal and hippocampal areas have been consistently associated with representation and processing of face information, and more particularly with novel face information as regards to hippocampal activation (Henson et al., 2003); the present study indeed used novel face stimuli unfamiliar to the participants.

The only modality-specific activations in extra-sensory regions were confined to the left inferior frontal gyrus (Broca's area) and adjacent sensori-motor cortex for the verbal STM condition, and to a medial superior frontal area for the face STM condition. Broca's area has been shown to be involved in phonological and articulatory processes, and is critical for verbal rehearsal processes (Paulesu et al., 1993). The medial superior frontal area being specifically activated in the face item STM condition has been shown to be activated in face identification processes, when detailed face information has to be compared or recognized (Henson et al., 2003; Platek et al., 2006). Detailed face recognition processes were indeed required in the present item STM conditions since positive and negative probes differed to a very minimal extent. Hence, modality specific effects in extra-sensory regions were also related to the specific sensory requirements of the items to be processed.

Perspectives for models of STM

Our data do not support the assumption of modality-specific STM buffers, at least if we consider STM buffers as being autonomous cognitive components, as proposed by the original working memory model

(Baddeley & Hitch, 1974). The present data revealed no differences in STM networks for faces and nonwords outside those neural networks already involved in the identification and processing of faces and nonwords. However, this does not rule out the possibility that the neural substrates for face and language processing serve themselves this buffer function, as suggested by the sustained activity during early maintenance observed in superior temporal phonological processing areas during the verbal STM condition, and in hippocampal face processing areas during the visual STM condition. This alternative interpretation is however in fundamental opposition with the working memory model, which considers that STM buffers and long-term representational systems are distinct functional entities. The present study shows that if these buffers exist, their representational substrate cannot by distinguished from long-term representational systems and sensory networks. Our data therefore strongly support alternative models of STM, considering that temporary activation of long-term memory representational systems is an undissociable element of STM processing (e.g., Baddeley et al., 1998; Cowan, 1999; Fuster, 1995; Martin & Saffran, 1992; Gupta, 2003). Our data further support attentional accounts of STM processing, by showing that fronto-parietal networks centered around the left intraparietal sulcus intervene equally in verbal and visual STM tasks, and hence play a general, amodal role during STM tasks, as implied by these theoretical positions (Cowan, 1999; Fuster, 1999). Importantly, our data are the first to show that a network centered around the right anterior intraparietal sulcus supports serial order retention processes during both verbal and visual STM conditions. Current models of serial order processing have been developed preferentially for the verbal domain (e.g., Brown et al., 2000; Burgess & Hitch, 1999; Henson, 1998; Gupta, 2003). The present data suggest that the mechanisms proposed to underlie serial order processing in the verbal domain are also applicable to serial order processing in the visual domain. This calls for the development of new STM models, assuming a common serial order processing device, communicating with verbal and visual sensory processing substrates involved in item short-term storage. More generally, the present results argue for considering STM not as an autonomous and modular cognitive function, but rather as an emergent property, resulting from the interaction between modality-independent attentional-executive processes, modality-

independent serial ordering processes and the temporary activation of modality-specific long-term memory representations.

ACKNOWLEDGMENTS

REFERENCES

Baddeley, A. D. (1974). Working memory. In G.H.Bower (Ed.), *The psychology of learning and motivation* (pp. 47-90). San Diego, CA: Academic Press.

Baddeley (2000). The episodic buffer : a new component of working memory? *Trends in Cognitive Sciences, 4,* 417-423.

Baker, S.C., Frith, C.D., Frackowiak, R.S. and Dolan, R.J. (1996). Active representation of shape and spatial location in man. *Cerebral Cortex, 6*, 612–619.

Becker, J. T., MacAndrew, D. K., & Fiez, J. L. (1999). A comment on the functional localization of the phonological storage subsystem of working memory. *Brain and Cognition, 41,* 27-38.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufman, J. N. et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex, 10,* 512-528.

Brahmbhatt, S. B., McAuley, T., & Barch, D. M. (2008). Functional developmental similarities and differences in the neural correlates of verbal and nonverbal working memory tasks. *Neuropsychologia, 46,* 1020-1031.

Bricolo, E., Gianesini, T., Fanini, A., Bundesen, C., Chelazzi, L. (2002). Serial attention mechanisms in visual search: a direct behavioral demonstration. *Journal of Cognitive Neuroscience, 14*, 980–993.

Buchsbaum, B. R. & D'Esposito, M. (2008). The search for the phonological store: From loop to convolution. *Journal of Cognitive Neuroscience, 20,* 762-778.

Burgess, N. & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106,* 551-581.

Burgess, N. & Hitch, G. (2005). Computational models of working memory: putting long-term memory into

context. *Trends in Cognitive Sciences, 9,* 535-541.

Cairo, T. A., Liddle, P. F., Woodward, T. S., & Ngan, E. T. (2004). The influence of working memory load on

phase specific patterns of cortical activity. *Cognitive Brain Research, 21,* 377-387.

Corbetta, M., Miezin, F. M., Shulman, G. L., & Petersen, S. E. (1993). A PET study of visuospatial attention.

Journal of Neuroscience, 13, 1202-1226.

Cowan, N. (1999). An embedded-processes model of working memory. In A.Miyake & P. Shah (Eds.), *Models of

working memory: Mechanisms of active maintenance and executive control* (pp. 62-101). Cambridge:

Cambridge University Press.

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A. et al. (2005). On the capacity of

attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51,* 42-

100.

Donner, T., Kettermann, A., Diesch, E., Ostendorf, F., Villringer, A., Brandt, S.A., 2000. Involvement of the

human frontal eye field and multiple parietal areas in covert visual selection during conjunction search.

European Joural of Neuroscience, 12, 3407–3414.

Fiebach, C. J., Friederici, A. D., Smith, E. E., & Swinney, D. (2007). Lateral inferotemporal cortex maintains

conceptual–semantic representations in verbal working memory. *Journal of Cognitive Neuroscience, 19,* 2035-

2049.

Frackowiak, R. S. J. (1994). Functional mapping of verbal memory and language. *Brain, 17,* 109-115.

Friston, K. J., Penny, W. D., & Glaser, D. E. (2005). Conjunction revisited. *NeuroImage, 25,* 661-667.

Fuster, J.M. (1999). *Memory in the cerebral cortex: an empirical approach to neural networks in the human and nonhuman primate* (372 pp). Cambridge: MIT Press.

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition and immediate serial recall in adults. *Quarterly Journal of Experimental Psychology, 56A,* 1213-1236.

Hautzel, H., Mottaghy, F. M., Schmidt, D., Zemb, M., Shah, N. J., Müller-Gärtner, H.-W. et al. (2002). Topographic segregation and convergence of verbal, object, shape and spatial working memory in humans. *Neuroscience Letters, 323,* 156-160.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4,* 223-233.

Henson, R. N. A. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology, 36*, 73-137.

Henson, R. N., Goshen-Gottstein, Y., Ganel, T., Otten, L. J., Quayle, A., & Rugg, M. D. (2003). Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cerebral Cortex, 13,* 793-805.

Henson, R. N. A., Burgess, N., & Frith, C. D. (2000). Recoding, storage, rehearsal, and grouping in verbal short-term memory: an fMRI study. *Neuropsychologia, 38,* 426-440.

Hulme, C., Maughan, S., & Brown, G. D. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language, 30,* 685-701.

Jonides, J., Smith, E. E., Koeppe, R. A., Awh, E., Minoshima, S., & Mintun, M. A. (1993). Spatial working memory in humans as revealed by PET. *Nature, 363,* 623-625.

Lycke, C., Specht, K., Ersland, L., & Hugdahl, K. (2008). An fMRI study of phonological and spatial working memory using identical stimuli. *Scandinavian Journal of Psychology, 49,* 393-401.

Majerus, S., Collette, F., Van der Linden, M., Peigneux, P., Laureys, S., Delfiore, G. et al. (2002). A PET investigation of lexicality and phonotactic frequency in oral language processing. *Cognitive Neuropsychology, 19,* 343-360.

Majerus, S. & Van der Linden, M. (2003). The development of long-term memory effects on verbal short-term memory : A replication study. *British Journal of Developmental Psychology, 21,* 303-310.

Majerus, S., Van der Linden, M., Mulder, L., Meulemans, T., & Peters, F. (2004). Verbal short-term memory reflects the sublexical organization of the phonological language network: Evidence from an incidental phonotactic learning paradigm. *Journal of Memory and Language, 51,* 297-306.

Majerus, S., Poncelet, M., Elsen, B., Van der Linden, M. (2006a). Exploring the relationship between new word learning and short-term memory for serial order recall, item recall and item recognition. *European Journal of Cognitive Psychology, 18*, 848–873.

Majerus, S., Poncelet, M., Van der Linden, M., Albouy, G., Salmon, E., Sterpenich, V et al. (2006b). The left intraparietal sulcus and verbal short-term memory: focus of attention or serial order? *NeuroImage, 32*, 880–891.

Majerus, S., Bastin, C., Poncelet, M., Van der Linden, M., Salmon, E., Collette, F. et al. (2007). Short-term memory and the left intraparietal sulcus: focus of attention? Further evidence from a face short-term memory paradigm. *NeuroImage, 35*, 353–367.

Majerus, S., Belayachi, S., De Smedt, B., Leclercq, A.L., Martinez, T., Schmidt, C. et al. (2008). Neural networks for short-term memory for order differentiate high and low proficiency bilinguals. *NeuroImage, 42*, 1698-1713.

Majerus, S. (2009). Verbal short-term memory and temporary activation of language representations: the importance of distinguishing item and order information. In A.S.Thorn & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 244-276). Hove, UK: Psychology Press.

Majerus, S., Heiligenstein, L., Gautherot, N., Poncelet, M., & Van der Linden, M. (2009). The impact of auditory selective attention on verbal short-term memory and vocabulary development. *Journal of Experimental Child Psychology, in press*.

Marshuetz, C., Smith, E. E., Jonides, J., DeGutis, J., & Chenevert, T. L. (2000). Order information in working memory: fMRI evidence for parietal and prefrontal mechanisms. *Journal of Cognitive Neuroscience, 12,* 130-144.

Martin, N. & Saffran, E. M. (1992). A computational account of deep dysphasia: Evidence from a single case study. *Brain and Language, 43,* 240-274.

Martin, R. C., Wu, D., Freedman, M., Jackson, E. F., & Lesch, M. (2003). An event-related fMRI investigation of phonological versus semantic short-term memory. *Journal of Neurolinguistics, 16,* 341-360.

Nairne, J. S. & Kelley, M. R. (2004). Separating item and order information through process dissociation. *Journal of Memory and Language, 50,* 113-133.

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers, 36,* 516-524.

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.

Pa, J., Wilson, S. M., Pickell, H., Bellugi, U., & Hickok, G. (2008). Neural organization of linguistic short-term memory is sensory modality-dependent: evidence from signed and spoken language. *Journal of Cognitive Neuroscience, 20,* 2198-2210.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. J. (1993). The neural correlates of the verbal component of working memory. *Nature, 362,* 342-345.

Platek, S. M., Loughead, J. W., Gur, R. C., Busch, S., Ruparel, K., Phend, N. et al. (2006). Neural substrates for functionally discriminating self-face from personally familiar faces. *Human Brain Mapping, 27,* 91-98.

Postle, B. R., Druzgal, T. J., & D'Esposito, M. (2003). Seeking the neural substrates of visual working memory storage. *Cortex, 39,* 927-946.

Postle, B. R. (2005). Delay-period activity in the prefrontal cortex: One function is sensory gating. *Journal of Cognitive Neuroscience, 17,* 1679-1690.

Prabhakaran, R., Blumstein, S. E., Myers, E. B., Hutchinson, E., & Britton, B. (2006). An event-related fMRI investigation of phonological-lexical competition. *Neuropsychologia, 44,* 2209-2221.

Price, C. & Devlin, J. T. (2003). The myth of the visual word form area. *NeuroImage, 19,* 473-481.

Rao, S. M., Mayer, A. R., & Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nature Neuroscience, 4,* 317-323.

Ravizza, S. M., Delgado, M. R., Chein, J. M., Becker, J. T., & Fiez, J. A. (2004). Functional dissociations within the inferior parietal cortex in verbal working memory. *NeuroImage, 22,* 562-573.

Raz, A., & Buhle, J. (2006). Typologies of attentional networks. *Nature Reviews Neuroscience, 7,* 367-379.

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. *Cognitive, Affective, & Behavioral Neuroscience, 1,* 161-171.

Rowe, J. B., Toni, I., Josephs, O., Frackowiak, R. S. J., & Passingham, R. E. (2000). The prefrontal cortex: Response selection or maintenance within working memory? *Science, 288,* 1656-1660.

Ruchkin, D. S., Berndt, R. S., Johnson, R. Jr., Grafman, J., Ritter, W., & Canoune, H. L. (1999). Lexical contributions to retention of verbal information in working memory: Event-related brain potential evidence. *Journal of Memory and Language, 41,* 345-364.

Salmon, E., Van der Linden, M., Collette, F., Delfiore, G., Maquet, P., Degueldre, C. et al. (1996). Regional brain activity during working memory tasks. *Brain, 119,* 1617-1625.

Saint-Aubin, J., & Poirier, M. (2000). Immediate serial recall of words and nonwords: Tests of a retrieval-based hypothesis. *Psychonomic Bulletin and Review, 7*, 332 -340.

Scott, S. K., Catrin Blank, C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain, 123,* 2400-2406.

Smith, E. E. & Jonides, J. (1995). Working memory in humans: neuropsychological evidence. In M.S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1009-1020). Cambridge: MIT Press.

Smyth, M. M., Dennis, C. H., & Hitch, G. (2005). Serial position memory in the visual-spatial domain: Reconstructing sequences of unfamiliar faces. *Quarterly Journal of Experimental Psychology, 58A,* 909-930.

Todd, J. J. & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428,* 751-754.

Todd, J. J., Fougnie, D., & Marois, R. (2005). Visual short-term memory load suppresses temporo-parietal junction activity and induces inattentional blindness. *Psychological Science, 16,* 965-972.

Tubach, J. L. & Boë, L. J. (1990). *Un corpus de transcription phonétique*. France: Telecom.

Turconi, E., Jemel, B., Rossion, B., & Seron, X. (2004). Electrophysiological evidence for differential processing of numerical quantity and order in humans. *Cognitive Brain Research, 21,* 22-38.

Ungerleider, L.G. and Haxby, J.V. (1994). 'What' and 'where' in the human brain. *Current Opinion of Neurobiology, 4*, 157–165.

Vitevitch, M. S. & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40,* 374-408.

Worsley, K. J., Marrett, S., Neelin, P., & Evans, A. C. (1996). Searching scale space for activation in PET images. *Human Brain Mapping, 4,* 74-90.

Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science, 256,* 846-849.

Table 1. <mark>Mean and standard deviations for reponse accuracy and latencies in the</mark> four STM conditions.

	Accuracy	Response latencies (ms)
Nonwords$_{Order}$	.79 (.12)	1976 (296)
Nonwords$_{Item}$	.74 (.12)	1557 (296)
Faces$_{Order}$	.78 (.09)	2096 (325)
Faces$_{Item}$	.75 (.08)	1542 (345)

Table 2. Common activation for the four experimental conditions (Nonword$_{Order}$ ∩ Nonword$_{Item}$ ∩ Faces$_{Order}$ ∩ Faces$_{Item}$; null conjunction), as a function of STM phase. If not otherwise stated, all regions are significant at p < .05, corrected for whole brain volume (at the voxel level).

Anatomical region	No. voxels	Left/ right	x	y	z	BA area	SPM {Z}- value
Encoding							
SMA / anterior cingulate	219	B	-6	20	48	6/32	5.28
Middle frontal gyrus	758	L	-42	30	30	9	5.17
Middle frontal gyrus	309	R	44	40	28	9	4.33*
Inferior frontal gyrus / Insula	456	L	-30	22	0	47	4.69
Intraparietal sulcus (ant)	14	L	-42	-48	46		3.47*
Caudate (head)	706	L	-6	6	24		4.60
		R	10	2	24		5.19
Caudate (tail)	425	L	-24	-36	12		5.43
Caudate (tail)	422	R	24	-32	14		5.87
Cerebellum	453	L	-38	-64	-32	CrI	4.63
Cerebellum	976	R	40	-68	-30	CrI	7.80
Cerebellum	1706	R	6	-80	-36	VIIB	6.54
Maintenance							
Precentral gyrus	10	L	-36	-18	58	6	3.23**
Retrieval							
SMA / anterior cingulate	1092	B	-12	32	32	6/32	5.56
Inferior frontal gyrus	4881	R	34	26	-10	47	>7.80
		R	38	26	14	45	6.13
		R	34	10	30	44	5.03
Middle frontal gyrus		R	52	30	26	9/46	5.78
Inferior frontal gyrus / Insula	3827	L	-30	22	-6	47	7.69
		L	-42	18	28	44	6.27
Middle frontal gyrus		L	-46	48	-2	10	4.64
Intraparietal sulcus	908	L	-32	-52	42	7	5.36
Intraparietal sulcus	913	R	42	-50	44	7/40	5.27
Cerebellum	1764	R	32	-48	-34	VI	6.27
		L	42	-66	-30	Cr1	4.22
Cerebellum	2482	L	-8	-80	-32	VIIB	6.13
		R	12	-78	-38	VIIB	5.66
		L	-38	-68	-34	CrI	4.91

* p< .05, small volume corrections
** p< .001, uncorrected

Table 3. Differential main effects of STM conditions and respective null conjunctions. If not otherwise stated, all regions are significant at p < .05, corrected for whole brain volume (at the voxel level).

Anatomical region	No. voxels	Left/ right	x	y	z	BA area	SPM {Z}-value	No. voxels	x	y	z	BA area	SPM {Z}-value
Order > Item								**Conjunction** (Nonwords$_{Order}$ > Nonwords$_{Item}$) ∩ (Faces$_{Order}$ > Faces$_{Item}$)					
Superior frontal gyrus	2064	L	-22	5	55	6	6.72	280	-20	-2	50	6	3.80*
Superior frontal gyrus	3155	R	28	14	52	6	7.29	484	26	16	50	6	3.92
Middle frontal gyrus		R	38	38	30	9	4.45*						
Middle frontal gyrus	816	L	-34	32	38	9	4.92						
Intraparietal sulcus	429	L	-40	-44	48		4.87						
Intraparietal sulcus	515	R	48	-40	46		6.13	64	50	-44	56		3.56*
Precuneus	513	R	12	-68	58	7	7.79	1661	14	-70	60	7	5.18
	449	L	-8	-56	56	7	6.59						
Superior occipital gyrus	442	L	-42	-76	28	19	4.70						
Superior occipital gyrus	303	R	44	-78	38	19	7.51	321	46	-76	34	19	4.69
Cerebellum	702	L	-38	-42	-48	VI	5.07	52	-38	-42	48	VI	3.47*
	668	R	34	-62	-34	CrI	4.41						
Item > Order								**Conjunction** (Nonwords$_{Item}$ > Nonwords$_{Order}$) ∩ (Faces$_{Item}$ > Faces$_{Order}$)					
Superior frontal gyrus	175	R	14	52	50	8	3.86**	No suprathreshold activation					
Verbal > Visual								**Conjunction** (Nonwords$_{Order}$ > Faces$_{Order}$) ∩ (Nonwords$_{Item}$ > Faces$_{Item}$)					
SMA/anterior cingulate	492	L	-4	4	64	6	5.42	77	-4	4	64	6	3.67*
Middle frontal gyrus	384	L	-56	2	42	6	6.59	248	-52	-6	48	6	4.12
Inferior frontal gyrus	515	L	-52	6	16	44	> 7.80	479	-54	6	12	44	6.08
Postcentral gyrus	434	L	-64	-16	20	43	> 7.80	388	-64	-16	20	43	5.93
Superior temporal gyrus	421	L	-58	-36	10	22	5.59						
Superior temporal sulcus	368	L	-66	-28	2	22	5.34						
Inferior temporal gyrus	837	L	-48	-58	-6	19	5.67	82	-44	-58	-4	19	3.72*
Supramarginal gyrus	485	L	-54	-34	32	40	5.82						

Region	Voxels	Hem	x	y	z	BA	Z	Voxels	x	y	z	BA	Z
Cerebellum	1177	R	26	-66	-26	IV	>7.80	351	26	-64	-26	VI	5.73
	152	R	18	-72	-50	VIIIA	4.90						

Visual > Verbal

Conjunction (Faces$_{Order}$ > Nonwords$_{Order}$) $\cap$ (Faces$_{Item}$ > Nonwords$_{Item}$)

Region	Voxels	Hem	x	y	z	BA	Z	Voxels	x	y	z	BA	Z
Superior frontal gyrus	3986	B	-4	60	38	9	5.07						
Medial frontal gyrus	588	B	0	54	-20	11	5.18						
Middle temporal gyrus	2231	L	-52	-68	24	39	7.12	957	-54	-72	26	39	4.39
Middle temporal gyrus	2254	R	46	-66	30	39	5.37	309	46	-66	28	39	3.71*
Inferior temporal gyrys	1214	L	-60	-8	-30	20	4.88						
Inferior temporal gyrus	987	R	28	-6	-46	20	5.67	86	28	-6	-46	20	3.69*
Fusiform gyrus	1624	R	42	-44	-28	37	7.49	436	44	-46	-28	37	5.26
Hippocampus	293	R	22	2	-26		4.42*						
Precuneus	2529	L	-10	-50	38	31	5.41						

* $p < .05$, small volume corrections

Table 4. Regions showing a significant interaction between experimental condition and STM phase. If not otherwise stated, all regions are significant at p < .05, corrected for whole brain volume (at the voxel level).

Anatomical region	No. voxels	Left/ right	x	y	z	BA area	SPM {Z}- value
Stronger response during nonword item and order STM conditions (encoding + retrieval)							
SMA / anterior cingulate	431	B	-6	2	62	6	6.01
Inferior frontal gyrus	1458	L	-56	4	24	44	>7.80
Postcentral gyrus		L	-62	-14	20	40	>7.80
Praecentral gyrus	1144	L	-50	-8	44	6/4	>7.80
Superior temporal gyrus	346	L	-64	-32	4	22	4.18*
Middle temporal gyrus	1414	L	-46	-66	20	39	7.07
Middle temporal gyrus	1180	R	44	-66	20	39	5.91
Inferior temporal gyrus	510	L	-50	-58	-8	37	5.15
Inferior occipital gyrus	292	L	-28	-102	-10	18	6.95
Cerebellum	502	R	22	-64	-26	VI	7.48
Stronger response during face item and order STM conditions (encoding + retrieval)							
Superior frontal gyrus	497	L	-14	56	38	8	4.79
Superior frontal gyrus	344	B	0	50	-20	11	4.92
Fusiform gyrus	1607	R	40	-70	-14	37	>7.80
Fusiform gyrus	1355	R	28	-84	-10	18	>7.80
Fusiform gyrus	1138	L	-28	-78	-10	37	6.69
Hippocampal gyrus	738	R	20	-36	-12	36	4.82
Hippocampal gyrus	358	R	20	-4	-24	28	5.35
Hippocampus	412	R	22	-30	-2		4.70
Precuneus	758	B	2	-60	32	7	4.71
Middle occipital gyrus	844	R	30	-96	20	18/19	7.67
Inferior occipital gyrus	513	R	8	-92	-10	18	>7.80
Cuneus	1141	L	16	-102	12	18	>7.80
Cuneus	819	R	-12	-108	6	18	>7.80
Cerebellum	1315	R	42	-48	-26	VI	>7.80
Cerebellum	558	L	-36	-48	-24	VI	6.28
Stronger response during nonword and face order STM conditions (encoding + retrieval)							
Inferior parietal lobule	215	R	42	-38	48	40	4.61*
Superior parietal lobule	1210	L	-12	-66	60	7	5.23
Superior parietal lobule	1040	R	14	-68	58	7	5.61
Superior occipital gyrus	962	R	42	-78	36	19	6.40
Stronger response during all STM conditions except the face item STM condition (encoding + retrieval)							
Inferior parietal lobule	972	L	-40	-36	40	40	6.61

  * p< .05, small volume corrections

Table 5. Main effects for sustained activity during the four STM conditions. If not otherwise stated, all regions are significant at p < .05, corrected for whole brain volume (at the voxel level).

Anatomical region	No. voxels	Left/ right	x	y	z	BA area	SPM {Z}- value
Nonwords$_{Order}$							
Superior frontal gyrus (1)	1212	L	-22	12	58	6	4.52
Superior frontal gyrus (2)	155	R	30	6	60	6	4.23*
Middle frontal gyrus (3)	2776	L	-44	32	32	9	6.25
Inferior frontal gyrus (4)		L	-52	8	14	44	5.74
Middle frontal gyrus (5)	590	R	42	40	30	9	5.44
Intraparietal sulcus (6)	1172	L	-38	-44	46		5.36
Intraparietal sulcus (7)	573	R	58	-42	56		4.84
Caudate (Tail) (8)	743	L	-22	-44	10		5.77
Caudate (Tail) (9)	1326	R	36	-44	0		5.50
Caudate Head (10)		L	-2	4	20		4.98
Cerebellum (11)	1841	R	32	-64	-32	CrI	5.93
Faces$_{Order}$							
Superior frontal gyrus (1)	1240	R	32	8	60	6	6.20
Middle frontal gyrus (2)	4208	L	-46	30	34	9	7.08
Superior frontal gyrus (3)		L	-28	6	64	6	6.45
Middle frontal gyrus (4)	787	R	44	32	32	9	5.82
Middle frontal gyrus (5)	1377	L	-30	52	8	10	4.45
Intraparietal sulcus (6)	1711	L	-40	-48	44		6.28
Intraparietal sulcus (7)	858	R	46	-42	48		5.20
Precuneus (8)	775	R	10	-66	58	7	6.19
Precuneus (9)		L	-6	-64	52	7	4.83
Caudate (Head) (10)	484	R	6	2	22		4.87
Caudate (Tail) (11)	401	L	-22	-36	10		5.17
Caudate (Tail) (12)	401	R	22	-32	12		5.48
Cerebellum (13)	645	R	40	-66	-34	CrI	5.41
Nonwords$_{Item}$							
SMA / anterior cingulate (1)	64	B	-4	10	58	6	3.92*
Middle frontal gyrus (2)	126	L	-44	32	30	9	4.34*
Inferior frontal gyrus (3)	2376	L	-52	6	18	44	6.53
Superior temporal sulcus (4)	54	L	-68	-26	-2	22	4.06*
Caudate (Head) (5)	312	L	-2	2	20		5.22
Caudate (Tail) (6)	338	L	-22	-44	10		5.29
Caudate (Tail) (7)	270	R	36	-44	0		4.01
Cerebellum (8)	394	R	26	-64	-28	VI	5.37
Faces$_{Item}$							
Middle frontal gyrus (1)	109	L	-46	28	34	9	3.77*
Hippocampus (2)	13	R	18	-10	-24		3.54*
Caudate (Head) (3)	202	L	-2	4	20		5.06

Caudate (Tail) (4)	110	L	-22	-44	10		3.98
Caudate (Tail) (5)	75	R	20	-38	12		4.01

* p< .05, small volume corrections

Table 6. Conjunction analyses for sustained effects during the four STM conditions. If not otherwise stated, all regions are significant at p < .05, corrected for whole brain volume (at the voxel level).

Anatomical region	No. voxels	Left/ right	x	y	z	BA area	SPM {Z}-value
Nonwords$_{Order}$ ∩ Nonwords$_{Item}$ ∩ Faces$_{Order}$ ∩ Faces$_{Item}$							
Middle frontal gyrus	37	L	-46	30	34	9	3.72*
Caudate (Head)	47	L	-8	4	20		4.80*
Caudate (Head)	98	R	8	4	20		4.80*
Caudate (Tail)	90	L	-22	-44	10		3.98*
	47	R	20	-38	12		4.01*
Nonwords$_{Order}$ ∩ Faces$_{Order}$							
Superior frontal gyrus	1004	L	-22	12	58	6	4.52
Superior frontal gyrus	356	R	30	6	60	6	4.23
Middle frontal gyrus	719	L	-44	32	32	9	6.25
Middle frontal gyrus	455	R	44	40	30	9	5.41
Intraparietal sulcus	782	L	-38	-44	46		5.36
Intraparietal sulcus	420	R	50	-42	50		4.47
Caudate (Head)	377	R	2	4	20		4.83
Caudate (Tail)	401	L	-22	-38	10		5.15
Caudate (Tail)	354	R	24	-34	12		4.99
Cerebellum	422	R	40	-66	-34	CrI	5.41
Nonwords$_{Item}$ ∩ Faces$_{Item}$							
Middle frontal gyrus	37	L	-46	30	34	9	3.72*
Caudate (Head)	180	L	-2	4	20		5.01
Caudate (Head)		R	8	4	20		4.80*
Caudate (Tail)	110	L	-22	-44	10		3.98*
	54	R	20	-38	12		4.01*

* p< .05, small volume corrections

Table 7. Differential main effects for sustained main effects between the visual and verbal STM conditions, as a function of item and order conditions. If not otherwise stated, all regions are significant at p < .05, corrected for whole brain volume (at the voxel level).

Anatomical region	No. voxels	Left/ right	x	y	z	BA area	SPM {Z}- value
Nonwords > Faces							
Order							
Inferior frontal gyrus	2956	L	-54	6	18	44	5.51
Postcentral gyrus		L	-64	-16	20	40	4.69
Superior temporal gyrus		L	-62	-30	6	22	4.19*
Mid-fusiform gyrus	118	L	-44	-58	-2	37	4.04*
Cerebellum	396	R	26	-66	-24	VI	5.01
Item							
SMA/anterior cingulate	427	B	-6	4	62	6	5.39
Inferior frontal gyrus	6551	L	-52	6	20	44	>7.60
Postcentral gyrus		L	-64	-16	22	40	6.02
Superior temporal gyrus		L	-58	-36	10	22	4.77
Mid-fusiform gyrus	309	L	-48	-62	-6	37	4.56*
Cerebellum	809	R	24	-64	-26	VI	6.36
Faces > Nonwords							
Order							
Fusiform gyrus	863	R	42	-48	-24	37	5.42
		R	42	-62	-18	37	5.14
Item							
Superior frontal gyrus	1557	B	0	64	30	9	4.69
Middle temporal gyrus	2213	L	-44	-64	22	39	5.84
Inferior temporal gyrus	1509	R	32	-6	-44	20	5.05
Fusiform gyrus	5091	R	40	-46	-24	37	7.03
		R	42	-62	-18	37	6.62
		R	50	-80	-8	18	5.48
Fusiform gyrus	840	L	-60	-12	-26	20	4.69
Hippocampus	176	R	22	-4	-26		4.78
Hippocampus	59	L	-22	-6	-24		3.76*
Precuneus	1928	B	-4	-58	34	7	4.93
Cuneus	310	L	-10	-110	6	18	6.25

* p< .05, small volume corrections

Table 8. Hemodynamic response function differences for the regions of interest, as a function of STM condition. The table shows the regions significantly differing in activation level relative to other regions, for a given time point of the hemodynamic response function (all p<.05, corrected for multiple comparisons – Bonferroni method). For ease of reading, the regions of interested are designated by their number labels as listed in Table 5.

Time	Region labels
Nonwords$_{Order}$	
Encoding	
T1	Cerebellum (11) > 1, 2, 3, 8, 9, 10
T2	/
Maintenance	
T3	Caudate (8, 9, 10) < 5
T4	Caudate (10) < 6, 11
T5	Caudate (8, 9, 10) > 1, 2, 3, 4, 5, 6, 7, 11
T6	/
Retrieval	
T7	/
T8	/
T9	Caudate (Head)10 > 3, 4, 5, 6, 8, 9
T10	/
Faces$_{Order}$	
Encoding	
T1	Precuneus (8) > 10, 11
T2	/
Maintenance	
T3	Caudate (10) < 1, 2, 4, 7, 8, 9
T4	Caudate (11) < 4, 5, 7
T5	Caudate (10, 11), Precuneus (9) > 1, 2, 4
T6	Caudate (11) > 8
Retrieval	
T7	Superior frontal (1) > 11, 13
T8	/
T9	/
T10	/
Nonwords$_{Item}$	
Encoding	
T1	Caudate (Tail) (6,7) < 1, 4, 5, 8
T2	Caudate (Head) (5) < 1, 3
Maintenance	
T3	Caudate Tail (7) < 1
T4	/
T5	Caudate (5, 6, 7) > 1, 2, 3, 4
T6	Caudate (5) > 8
Retrieval	
T7	Caudate (5, 6) < 1, 2, 3

T8	/
T9	/
T10	/
Faces$_{Item}$	
Encoding	
T1	/
T2	/
Maintenance	
T3	/
T4	Caudate (4,5), Hippocampus (2) < 1
T5	Caudate (4) > 1
T6	/
Retrieval	
T7	/
T8	/
T9	Middle frontal gyrus > 2, 3, 4, 5
T10	/

FIGURE LEGENDS

Figure 1. Experimental design and timing of the probe recognition tasks. For each condition, a negative probe trial is illustrated. Translation of task instructions: "*retenir ordre*"- remember the order, "*retenir identité*" – remember the identity, "*regarder*" – look (at the picture / nonwords).

Figure 2. Main effects for transient effects in the four STM conditions, as a function of STM phase, rendered on 3D MRI template (results are shown at p<.001, uncorrected). The upper part displays main effects in the left hemisphere and the lower part displays main effects in the right hemisphere.

Figure 3. Average beta coefficient estimates for regions presenting a phase-by-STM condition interaction in the transient effect analyses (beta coefficients have been averaged over regions showing the same pattern of activation as a function of STM condition and phase; see text and Table 4 for more details). Results are shown at p<.001, uncorrected, rendered on a 3D MRI template. E = encoding, M = maintenance, R = retrieval.

Figure 4. Main effects for sustained activity in the four STM conditions, rendered on a 3D MRI template (results are shown at p<.001, uncorrected). The upper part of the figure shows sustained activation for the verbal (red) and visual (green) order STM conditions; the lower part of the figure shows sustained activation for the verbal (red) and visual (green) item STM conditions.

Figure 5. Time course analyses for modeled hemodynamic response, as a function of STM condition. The leftward figures (labels with suffix A) show mean time curves for sustained response in all regions except for caudate nuclei, as a function of STM condition (see text for further details). The rightward figures (labels with suffix B) show mean time curves for sustained response in caudate nuclei, as a function of STM condition. * indicates responses significantly different from baseline (t-tests, all p<.05).
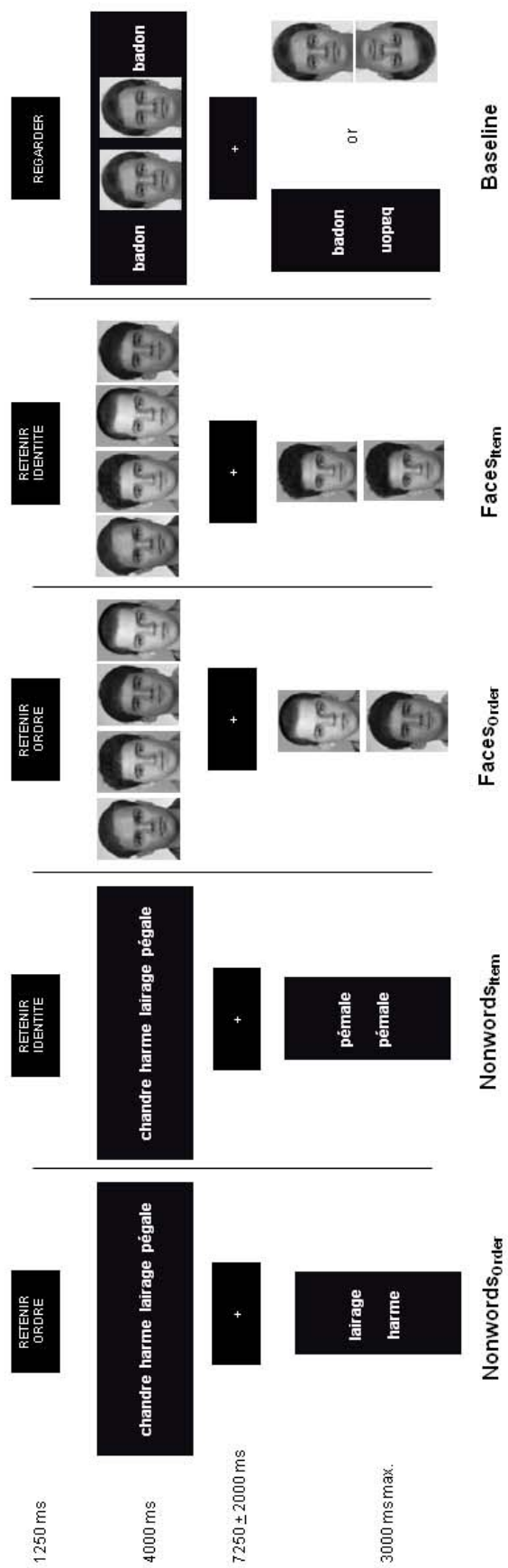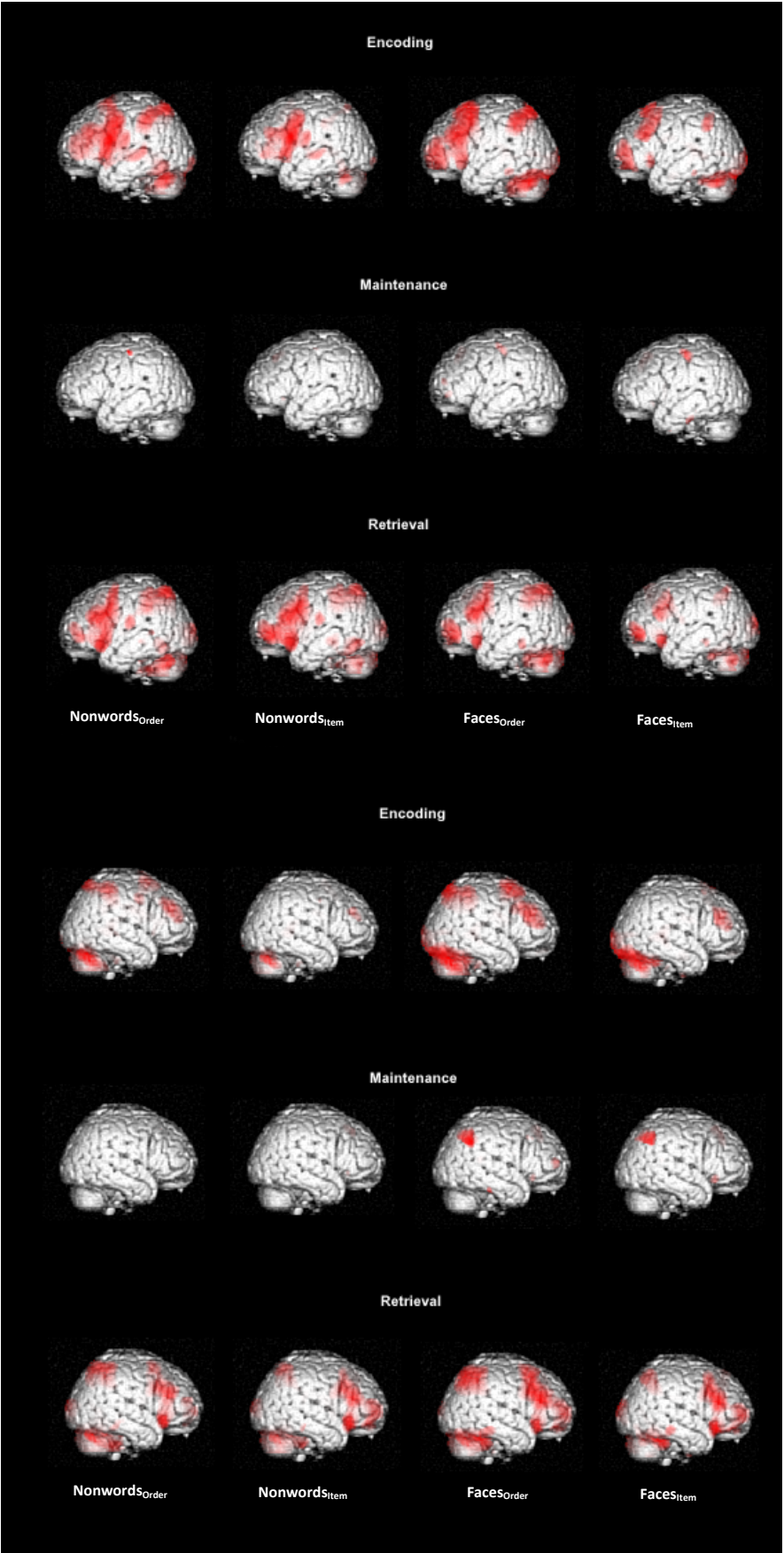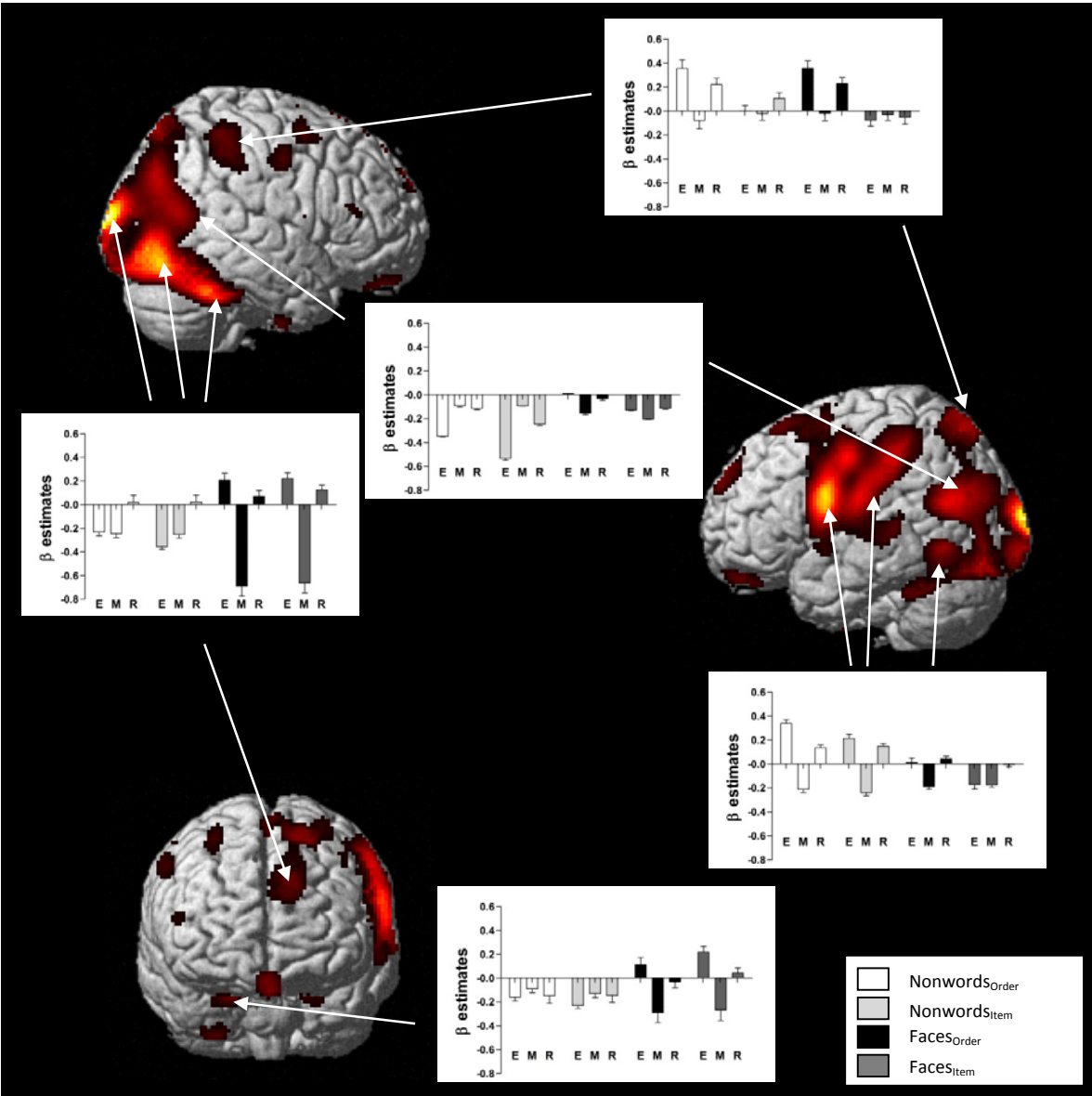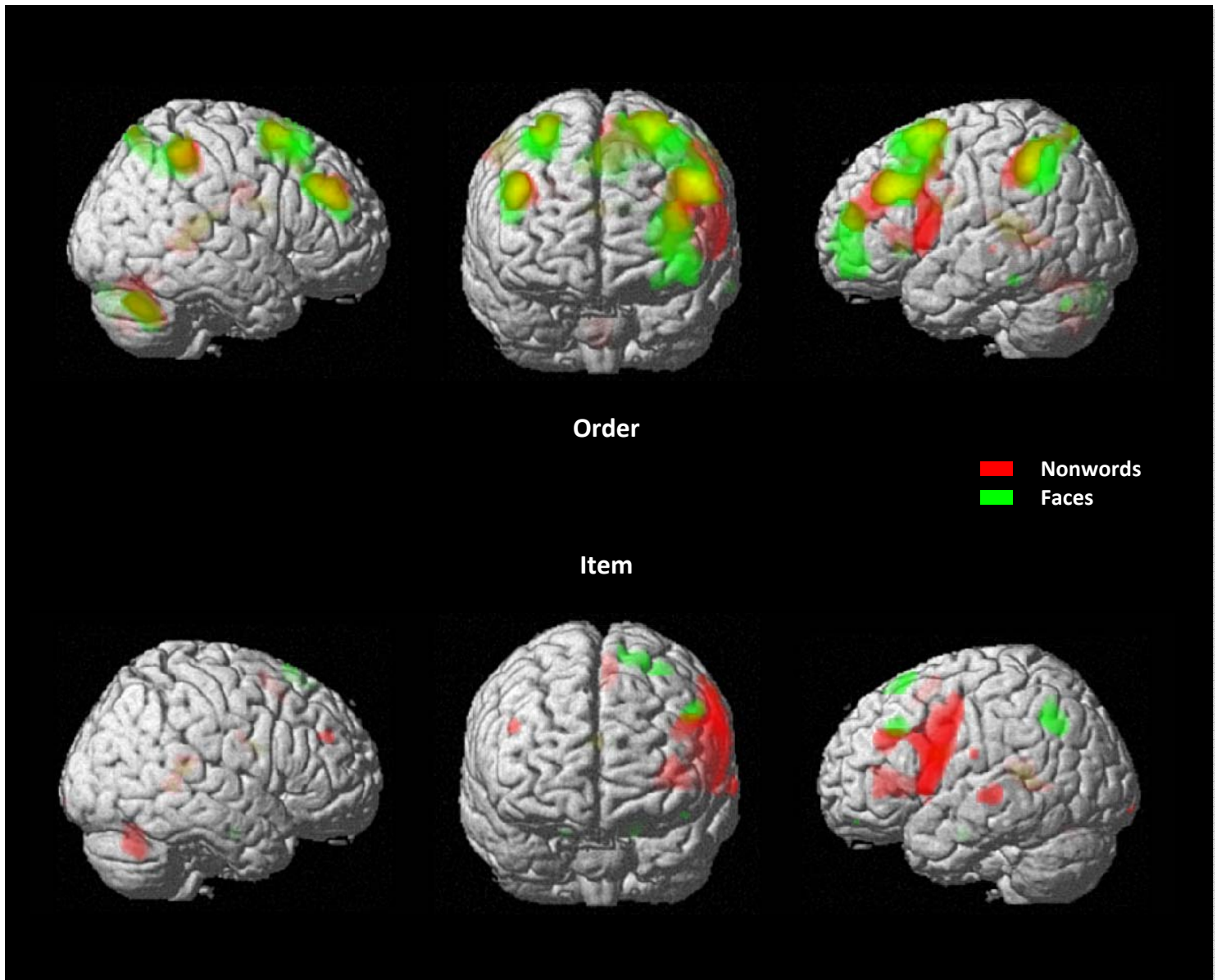
Figure 1.

Figure 2.

Figure 3.

Figure 4.

Order

Nonwords
Faces

Item

Figure 5.