EJN fens EUROPEAN JOURNAL OF NEUROSCIENCE

European Journal of Neuroscience, pp. 1–8, 2010 doi:10.1111/j.1460-9568.2010.07470.x

Neural correlates of strategy use during auditory working memory in musicians and non-musicians

K. Schulze,[1,2] K. Mueller[1] and S. Koelsch[1,3]
[1]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[2]Developmental Cognitive Neuroscience Unit, UCL Institute of Child Health, 30 Guilford Street, London, WC1N 1EH, UK
[3]Cluster of Excellence 'Languages of Emotion', Freie Universität Berlin, Berlin, Germany

Keywords: auditory working memory, inferior precentral sulcus, premotor cortex, strategy, structure in music

Abstract

Working memory (WM) performance in humans can be improved by structuring and organizing the material to be remembered. For visual and verbal information, this process of structuring has been associated with the involvement of a prefrontal–parietal network, but for non-verbal auditory material, the brain areas that facilitate WM for structured information have remained elusive. Using functional magnetic resonance imaging, this study compared neural correlates underlying encoding and rehearsal of auditory WM for structured and unstructured material. Musicians and non-musicians performed a WM task on five-tone sequences that were either tonally structured (with all tones belonging to one tonal key) or tonally unstructured (atonal) sequences. Functional differences were observed for musicians (who are experts in the music domain), but not for non-musicians – The right pars orbitalis was activated more strongly in musicians during the encoding of unstructured (atonal) vs. structured (tonal) sequences. In addition, data for musicians showed that a lateral (pre)frontal–parietal network (including the right premotor cortex, right inferior precentral sulcus and left intraparietal sulcus) was activated during WM rehearsal of structured, as compared with unstructured, sequences. Our findings indicate that this network plays a role in strategy-based WM for non-verbal auditory information, corroborating previous results showing a similar network for strategy-based WM for visual and verbal information.

Introduction

Although the amount of information that can be held in working memory (WM) is limited (Miller, 1956; Baddeley *et al.*, 1975; Baddeley, 2003), research has shown that WM performance can be improved by the use of a strategy, for example by chunking the information to be remembered (Ericsson *et al.*, 1980; Gobet *et al.*, 2001). During chunking, elements of information are organized into one unit or chunk (Miller, 1956), with strong associations between within-chunk elements and weak associations between elements belonging to different chunks (Gobet *et al.*, 2001).

Previously, such strategy-based memorization was explored using visual–spatial or verbal material (e.g. Mandler, 1967; Savage *et al.*, 2001; Bor *et al.*, 2003, 2004; Bor & Owen, 2007). For instance, recall memory for lists of spatial patterns increased for 'structured' (similar to a known shape, e.g. triangle or square) as compared with 'unstructured' sequences (Bor *et al.*, 2003). This was mainly associated with stronger activation of the lateral prefrontal cortex (LPFC) bilaterally [including the inferior frontal gyrus (IFG)] and the bilateral inferior parietal lobe (IPL) [Brodmann area (BA) 40] during encoding of 'structured' than of 'unstructured' stimuli. These results obtained with visual–spatial patterns were supported by a follow-up study (Bor *et al.*, 2004) using verbal stimuli (spoken digits), which either contained a mathematical structure or were unstructured. Participants' recall performance was significantly better and faster for structured than for unstructured sequences. This better WM performance was associated with stronger activation mainly of the LPFC bilaterally [including the IFG and the dorsolateral prefrontal cortex (DLPFC)] and the IPL bilaterally during encoding of the structured material, similar to the previous results (Bor *et al.*, 2003). Bor & Owen (2007) investigated which strategy could be used during the strategic recoding of visually presented digits – a mathematical or a mnemonic strategy. Again, the authors observed activation of the LPFC (Bor *et al.*, 2003, 2004) during both mathematical and mnemonic strategic processes, indicating that the LPFC is involved during strategic WM processes, regardless of the strategy used.

One example of strategic processing for language stimuli is semantic organization, whereby lists of randomly ordered words are better memorized by mental grouping of these words into categories (Tulving, 1962; Mandler, 1967). This semantic organization is associated with stronger activation of the left DLPFC (BA 9) and the left IFG (BA 45/46) (Savage *et al.*, 2001).

Hence, the use of a strategy for visual–spatial or verbal information during encoding seems to rely on the recruitment of a prefrontal–parietal network (Bor & Owen, 2007), including the LPFC (Savage

---

Correspondence: Katrin Schulze, [2]Developmental Cognitive Neuroscience Unit, as above.
E-mail: kschulze@ich.ucl.ac.uk

Stefan Koelsch, [3]Cluster of Excellence 'Languages of Emotion', as above
E-mail: stefan.koelsch@fu-berlin.de

Received 19 January 2010, accepted 9 September 2010

et al., 2001; Bor *et al.*, 2003, 2004; Bor & Owen, 2007) and the inferior parietal cortex (Bor *et al.*, 2003, 2004; Bor & Owen, 2007).

Surprisingly, the brain structures serving strategic WM processes for non-verbal auditory stimuli have remained elusive. The present functional magnetic resonance imaging (fMRI) study investigates which neural correlates are associated with strategy-based maintenance of non-verbal auditory information in WM. For this purpose, five-tone sequences with and without musical structure (for details see Materials and methods) were used. We made use of the scale information underlying major–minor tonal music, because previous research showed that both musicians and non-musicians remember musical stimuli consisting of pitches that belong to a tonal scale (i.e. tonal musical stimuli) better than atonal musical stimuli (Krumhansl, 1979; Dowling, 1991).

Research on how expertise influences memory performance has mainly investigated chess experts. For example, it was observed that chess experts can remember briefly shown chess positions better than non-experts [for an overview about the theories explaining expert memory, see Gobet (1998) and Gobet *et al.* (2001)]. WM performance in musical experts has so far received little attention (e.g. Pechmann & Mohr, 1992; Williamson *et al.*, 2010; Schulze *et al.*, in press). Musical expertise is associated with both functional and structural brain plasticity [for an overview, see Munte *et al.* (2002) and Jancke (2009)]. Functional differences have mainly been investigated with the use of event-related potential components and fMRI. For example, more pronounced event-related potentials associated with the perception of irregularities in musical syntax (Koelsch *et al.*, 1999, 2002b) and with the processing of pitch and pitch patterns (e.g. Koelsch *et al.*, 1999; Pantev *et al.*, 2001; Tervaniemi *et al.*, 2001, 2005; Fujioka *et al.*, 2004, 2005) have been reported in musicians than in non-musicians. Also, fMRI studies showed the benefit of training, that is, expertise, on WM performance (Olesen *et al.*, 2004; Gaab *et al.*, 2006; Moore *et al.*, 2006). Participants showed better task-related performance after training, which was associated with an increased blood oxygen level-dependent (BOLD) signal in the LPFC and parietal regions during visual WM performance (Olesen *et al.*, 2004; Moore *et al.*, 2006) and in parietal areas during auditory WM performance (Gaab *et al.*, 2006).

In addition, there is also broad evidence that musically trained listeners have a more elaborate sense of key when listening to sequences of tones, or chords that belong to one key (e.g. Krumhansl & Shepard, 1979; Koelsch *et al.*, 2002b). By using tonal (structured) and atonal (unstructured) five-tone sequences, the present fMRI study investigated whether musical structure influences encoding and rehearsal in a non-verbal auditory WM task, and how this is reflected in the brain. To explore the potential influence of training and expertise, we compared structured and unstructured tone sequences in musicians and non-musicians. Stronger involvement of lateral prefrontal and parietal areas for structured (tonal) than for unstructured (atonal) auditory sequences (Bor *et al.*, 2003, 2004; Olesen *et al.*, 2004; Gaab *et al.*, 2006; Moore *et al.*, 2006; Bor & Owen, 2007) was expected in musicians (Koelsch *et al.*, 1999, 2002b) than in non-musicians (Koelsch *et al.*, 2000; Bigand & Poulin-Charronnat, 2006).

Materials and methods

Participants

Seventeen right-handed non-musicians (age range, 21–29 years; average age, 25.47 years; nine males) and 16 right-handed musicians (age range, 20–27 years; average age, 23.50 years; nine males) took part in this experiment. Musicians studied an instrument at the University of Music and Theatre Mendelsohn Bartholdy in Leipzig. Eight musicians studied the piano, four a woodwind instrument, three a string instrument, and one a brass instrument. Musicians started their musical training at an average age of 6.10 years [standard error of the mean (SEM), 0.19 years], and practised for several hours each day. None of the musicians possessed absolute pitch (as tested with an absolute pitch test) (Keenan *et al.*, 2001). Non-musicians had not received any formal musical training (besides regular school lessons, which did not include learning a musical instrument). Participants were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971); the mean lateralization quotient was 97% for musicians and 95% for non-musicians, with no significant difference between the two groups ($t_{31} = 0.97$, $P = 0.34$). The study was approved by the local ethics committee of the University of Leipzig, and conducted in accordance with the Declaration of Helsinki.

Stimuli

For the WM task, participants listened to sine wave tones. The frequencies of the sine wave tones corresponded to the frequencies of the tones of the Western chromatic scale (based on $A = 440$ Hz), and included tones from one octave ranging from 261 Hz (C4) to 523 Hz (C5). Each tone had a duration of 400 ms. Stimuli were presented in sequences of five tones, with 150-ms periods of silence between tones, resulting in a sequence duration of 2600 ms. There were 40 tonal and 40 atonal sequences. To investigate strategy-based WM for tones, we designed two types of sequence. In the tonal sequences, all five tones belonged to one tonal key, and of these five tones, three tones formed a triad; these three tones did not necessarily directly succeed each other. Atonal sequences consisted of neither triad nor key (Fig. 1).

Procedure

In the tonal and atonal conditions, participants listened to one of the sequences of five tones (henceforth referred to as the sample stimuli sequence), and then rehearsed the tones internally for 4200 ms up to 6200 ms. At the end of each trial, a tone (henceforth referred to as the test stimulus) was presented, and participants had to indicate by a button press whether the test stimulus had already been presented during the sample stimuli sequence.

For both tonal and atonal sequences, in 50% of the sequences the test stimulus was already presented during the sequence, and in 50% it was not, resulting in a chance level of 50%. Tonal and atonal sequences were presented pseudorandomly. Participants were repeatedly instructed not to sing or hum aloud during the scanning session. This experiment also investigated verbal WM, which is reported elsewhere (Schulze *et al.*, in press). Participants were not informed about any structure in the tonal sequences that might help them to

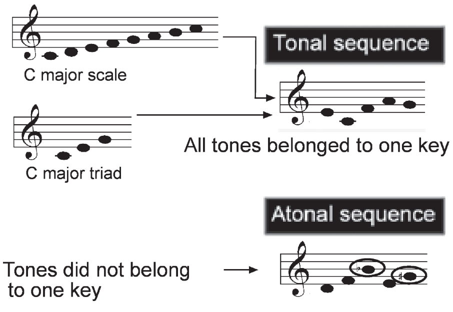

FIG. 1. Tonal and atonal sequences.

perform the task. To investigate whether participants were aware of the two types of structure (tonal and atonal), they were given a questionnaire after the scanning session. Participants indicated in this questionnaire whether some sequences were easier to remember. Performance data (as percentage of correct responses) were analysed with the SPSS statistical software package (SPSS, Chicago, IL, USA).

Data acquisition

The scanning paradigm was a modified version of the sparse temporal sampling method (Hall *et al.*, 1999), in which auditory stimuli were presented in the absence of the scanner gradient noise. Two scans per trial were acquired, allowing scanning of the haemodynamic response associated with (i) the encoding (first scan) and (ii) the rehearsal (second scan) of the sequences. Five different onsets of the auditory sequence relative to the first scan, differing in their onsets by 500 ms, were used to allow optimal sampling of the haemodynamic response associated with the encoding; that is, scans occurred 0, 500, 1000, 1500 or 2000 ms after the auditory presentation. The rehearsal time differed in length accordingly (rehearsal times were 4200, 4700, 5200, 5700, or 6200 ms). The first scan captured the BOLD response associated with encoding processes, and the second scan was associated with the WM rehearsal process (Schulze *et al.*, in press). The data of the first scan (encoding) and second scan (WM rehearsal) were analysed separately.

The experiment was carried out on a 3-T TRIO MR-scanner (Siemens, Erlangen, Germany). For each participant, a T1-weighted 3D MP-RAGE (magnetization-prepared rapid gradient echo) sequence was acquired (Mugler & Brookeman, 1990). This high-resolution anatomical dataset was standardized to the Talairach space (Talairach & Tournoux, 1988). Functional imaging was performed using bunched gradient-echo echo planar imaging with an echo time of 30 ms, a flip angle of 90°, a repetition time of 6600 ms and an acquisition bandwidth of 100 kHz. Twenty-four axial slices were acquired rapidly within approximately 1600 ms, so that no scanning occurred during the rest of the repetition time. The matrix dimensions were 64 × 64 with a field of view of 192 mm, resulting in a voxel size of 3 × 3 mm, a slice thickness of 4 mm and an interslice gap of 1 mm.

Data analysis

Pre-processing, statistical analysis and visualization of the fMRI data were performed with the software package LIPSIA (Lohmann *et al.*, 2001). An offline motion correction was performed on the functional images, using the Siemens motion correction protocol (Siemens). Thereafter, functional slices were aligned to a 3D stereotactic coordinate reference system. The registration parameters were acquired to achieve an optimal match between the functional slices and the individual 3D reference dataset, which was standardized to the Talairach stereotactic space (Talairach & Tournoux, 1988). The registration parameters were further used to transform the functional slices by using trilinear interpolation, so that the resulting functional slices were normalized with the stereotactic coordinate system. In addition to this linear transformation, a non-linear registration was performed between the anatomical 3D datasets of the group (Thirion, 1998), and the resulting deformation fields were applied to the associated functional datasets. In the last step of pre-processing, the data were smoothed with a Gaussian filter of 8-mm full-width at half-maximum.

The statistical evaluation was based on the general linear model. The design matrix was generated with the canonical haemodynamic response function (Friston *et al.*, 1998). Subsequently, contrast images were generated by computing differences between the parameter estimates (e.g. between the tonal and the atonal condition). In order to restrict the statistical analysis to relevant voxels inside the brain, a mask was applied to eliminate data from outside the brain and in the ventricular system.

Contrast images were entered into a second-level random effects analysis. One-sample *t*-tests were performed to evaluate whether observed differences were significantly different from zero (tonal vs. atonal within groups).

To protect against false-positive activation, the results were corrected for multiple comparisons by the use of cluster-size and cluster-value thresholds obtained by Monte Carlo simulations with a significance level of $P < 0.05$ (Lohmann *et al.*, 2008). Clusters were obtained with a voxel-wise threshold of $P < 0.001$ and an extent threshold of 10 voxels (Forman *et al.*, 1995). For activations that did not survive a whole-brain correction for multiple comparisons, a small volume correction (SVC) was performed, correcting the results for a restricted search volume using a sphere with a diameter of 10 voxels.

This article deals with WM for tonal and atonal auditory material. Comparisons between WM for syllables and tones (atonal and tonal sequences together) and comparisons of tonal WM between musicians and non-musicians are reported elsewhere (Schulze *et al.*, in press). Therefore, the main focus of this study was to analyse contrast images for tonal vs. atonal encoding and tonal vs. atonal rehearsal within groups: within musicians, within non-musicians, and within the pooled data from both groups.

To investigate differences between groups (musicians and non-musicians), two-sample *t*-tests for the contrast tonal vs. atonal were performed. Difference images (tonal−atonal) were compared between musicians and non-musicians to investigate whether the functional differences observed in musicians (in the contrast tonal−atonal) differed significantly between groups.

To investigate the influence of performance on the observed activation pattern, a regression analysis was conducted, partialling out the effect of the performance difference between tonal and atonal sequences. This regression analysis was performed on the tonal vs. atonal contrast images for musicians and for the pooled data of both groups during encoding and rehearsal, using a behaviourally obtained parameter of the performance difference tonal−atonal (as percentage of correct responses). For the regression analyses, no extent threshold was applied before the SVC was used.

To compare potential activation differences between the structured (tonal) and the unstructured (atonal) conditions in frontal and parietal sites of the present study with those of previous studies (Bor *et al.*, 2003, 2004; Bor & Owen, 2007), we carried out a region of interest (ROI) analysis and used ROI coordinates defined and used in other studies [DLPFC and ventrolateral prefrontal cortex (VLPFC) ROIs were used by Bor *et al.* (2003, 2004); all ROIs were used by Bor & Owen (2007)]. Contrast values between the tonal and atonal conditions within musicians and non-musicians were compared for the following ROIs: DLPFC (left, −40, 28, 19; right, 35, 31, 22); VLPFC (left, −41, 20, 0; right, 37, 20, 3); anterior cingulate gyrus (0, 26, 31); and intraparietal sulcus (IPS) (left, −37, −53, 40; right, 37, −53, 40). All ROIs were 10-mm-radius spheres centred on these coordinates.

Results

Behavioural data

Musicians had, on average, 72.50% (SEM, 2.28%) correct responses for the tonal (structured) sequences, and 66.56% (SEM, 3.32%) for the atonal (unstructured) sequences (Table 1). Non-musicians had, on

TABLE 1. Mean performance (percentage correct responses) for non-musicians and musicians for tonal and atonal sequences (numbers in parenthese are SEMs)

	Tonal	Atonal	Tonal vs. atonal
Non-musicians	57.87 (1.65)	54.70 (2.60)	$t_{16} = 1.24$, $P = 0.23$
Musicians	72.5 (2.28)	66.56 (3.32)	$t_{15} = 2.78$, $P = 0.014$*

Paired-samples *t*-tests were used to compare performance for tonal and atonal sequences. *Significant difference after a Bonferroni correction.

average, 57.87% (SEM, 1.65%) correct responses for the tonal sequences, and 54.70% (SEM, 2.60%) for the atonal sequences (Table 1). Performance data (percentage of correct responses for tonal and atonal sequences, for non-musicians and musicians) were normally distributed, as shown by a Shapiro–Wilk test ($P > 0.05$; non-significant *P*-values indicate a normal distribution in this test). Musicians' performance was significantly better than chance (50%) during both the tonal and atonal conditions, as indicated by one-sample *t*-tests ($P < 0.001$ in each test). Non-musicians' performance was significantly better than chance for the tonal sequences ($P < 0.001$), whereas for the atonal condition, non-musicians showed only a slight tendency towards a better-than-chance performance ($P = 0.09$).

An ANOVA with factors tonality (tonal and atonal) as within-subject factor and group (non-musicians and musicians) as between-subjects factor showed a main effect of tonality ($F_{1,31} = 7.38$, $P = 0.011$), and a main effect of group ($F_{1,31} = 17.90$, $P < 0.01$), but no two-way interaction ($F_{1,31} = 0.68$, $P = 0.42$).

To investigate task difficulty in more detail, we analysed participants' answers given in the questionnaire (see Materials and methods): 15 of 16 (94%) musicians but only eight of 17 (47%) non-musicians indicated that some sequences were less difficult to remember. This difference between groups proved to be significant ($P < 0.005$), as indicated by a Pearson chi-square test.

fMRI data

fMRI results will be reported in their sequential order: first for encoding processes (scan 1), and then for WM rehearsal (scan 2).

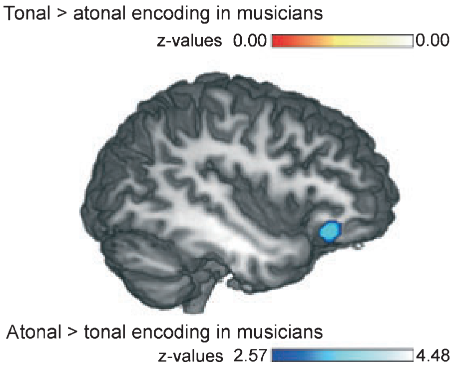

FIG. 2. Stronger activation of the right pars orbitalis during atonal than during tonal encoding in musicians ($z > 2.57$, corrected for multiple comparisons).

Furthermore, because activation differences between the tonal and atonal conditions were mainly expected for the musicians, results will be presented first for musicians, then for non-musicians, and then for the pooled data from both groups.

Encoding

In musicians, the pars orbitalis of the right IFG (BA 47) was activated more strongly during encoding of atonal than of tonal sequences (Fig. 2; Table 2). No differences were observed for the non-musicians. In the pooled data from both groups, stronger activation was observed in the contrast tonal > atonal in the right parahippocampal gyrus during tonal rehearsal (Table 2).

WM rehearsal

As depicted in Fig. 3 (see also Table 2), musicians showed an increased BOLD response during tonal as compared with atonal rehearsal in the right superior frontal gyrus (SFG), right inferior precentral sulcus (IPCS), right premotor cortex (PMC) and left IPS. An increased BOLD response was observed in the left middle occipital gyrus (MOG) during atonal as compared with tonal rehearsal. No functional difference for the contrast tonal−atonal rehearsal was

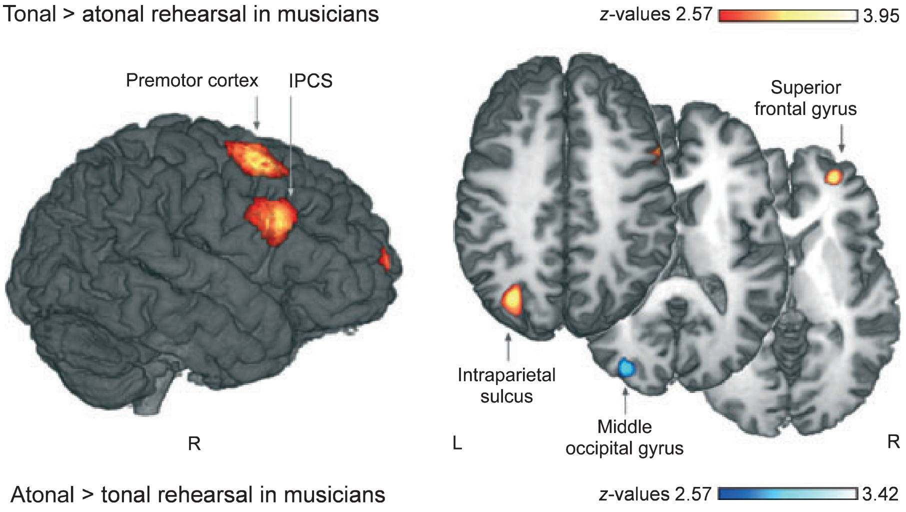

FIG. 3. Stronger activation ($z > 2.57$, uncorrected) of the right superior frontal gyrus (SFG), right inferior precentral sulcus (IPCS), right premotor cortex (PMC) and left intraparietal sulcus (IPS) during tonal than during atonal rehearsal and stronger activation of the left middle occipital gyrus (MOG) during atonal than during tonal rehearsal in musicians (R: right hemisphere, L: left hemisphere).

observed in non-musicians, and no difference was observed for the pooled data from both groups.

Comparison between groups

Difference images (tonal−atonal) were compared between musicians and non-musicians to investigate whether the functional differences observed in musicians (in the contrast tonal−atonal; Table 2) differed significantly between groups. The only structure in which such a group difference was observed was the right PMC (Table 2) during rehearsal.

Regression analysis

To investigate the influence of performance on the observed activation pattern, a regression analysis was conducted, partialling out the effect of the performance difference between tonal and atonal sequences. For musicians, activation differences were observed during encoding (right pars orbitalis for atonal > tonal) and rehearsal (right SFG, right IPCS, right PMC and left IPS for tonal > atonal; and left MOG for atonal > tonal; Table 2). For the pooled data of both groups, the right parahippocampal gyrus was more strongly activated during the encoding of tonal than of atonal sequences (Table 2). With the regression analysis, it was investigated whether these activations could still be observed after partialling out the performance difference between both conditions (Table 2). All structures that showed an activation difference in the tonal vs. atonal comparison (Table 2 and structures listed above) were still observed in the regression analysis. This indicates that the reported differences of the activation patterns between tonal and atonal sequences were indeed driven by the use of a strategy, and not simply by performance differences.

ROI analysis

In addition to the group-wise analysis of the statistical parametric maps (SPMs), ROI analyses were conducted to compare contrast values between the tonal and the atonal conditions between musicians and non-musicians (see Materials and Methods for coordinates). For musicians, a tendency towards a significantly stronger activation was observed in the right DLPFC ROI ($t = 1.79$; $P = 0.037$), the right IPS ROI ($t = 1.78$; $P = 0.038$) and the left IPS ROI ($t = 1.40$; $P = 0.081$) during the rehearsal of tonal, as compared with atonal, sequences (note that the Bonferroni-corrected threshold for statistical significance corresponds to a P-value of 0.007). No differences were found for the VLPFC ROI and the anterior cingulate gyrus ROI in musicians. Differences between the tonal and atonal conditions for the same ROIs were observed neither for non-musicians nor for the pooled data from both musicians and non-musicians.

Discussion

Musicians showed better performance for tonal than for atonal sequences. Presumably, musicians' knowledge about musical regularities (Krumhansl, 1979; Krumhansl & Shepard, 1979; Koelsch *et al.*, 1999, 2002c) contributed to keeping the structured (tonal) sequences in WM.

During WM rehearsal of structured (tonal) as compared with unstructured (atonal) sequences, the right IPCS and PMC, as well as the left IPS, were activated more strongly in musicians. In the

TABLE 2. Coordinates and significant values of the contrasts tonal vs. atonal (tonal > atonal and atonal > tonal) for the pooled data from both groups (musicians and non-musicians) and for musicians only. Furthermore, difference images (tonal−atonal) were compared between musicians and non-musicians to investigate whether the functional differences observed in musicians (in the contrast tonal−atonal) differed significantly between groups (musicians > non-musicians). To investigate the influence of performance on the observed activation pattern, a regression analysis was conducted, partialling out the effect of the performance difference between tonal and atonal sequences. Coordinates refer to standard stereotactic space (Talairach & Tournoux, 1988). To protect against false-positive activation, the results were corrected for multiple comparisons by using cluster-size and cluster-value thresholds obtained by Monte Carlo simulations with a significance level of $P < 0.05$. Activations that survived correction for multiple comparisons with a significance level of $P < 0.05$ are indicated in bold and marked with an asterisk. Clusters were obtained using a voxel-wise threshold of $P < 0.001$ and an extent threshold of 10 voxels. For activations that did not survive a whole-brain correction for multiple comparisons, a small volume correction (SVC) was performed, correcting the results for a restricted search volume using a sphere with a diameter of 10 voxels (these activations are indicated in italic). Blank fields indicate non-significant results.

Structure		Tonal > atonal		Atonal > tonal		Musicians > non-musicians		Regression	
ENCODING	BA	Coordinate (*x*, *y*, *z*)	*z*-value (SPM)	Coordinate (*x*, *y*, *z*)	*z*-value (SPM)	Coordinate (*x*, *y*, *z*)	*z*-value (SPM)	Coordinate (*x*, *y*, *z*)	*z*-value (SPM)
Musicians and non-musicians									
Right anterior parahippocampal gyrus	36	*34, −30, −15*	*3.67*					*34, −33, −15*	*3.35*
Musicians									
Right pars orbitalis	47			**43, 27, −9***	**4.68**			**40, 27, −9***	**4.36**
REHEARSAL									
Musicians									
Left IPS	7/40	*−29, −69, 39*	*3.92*					*−29, −69, 39*	*3.68*
Right IPCS/PMC	6/9	*52, 12, 27*	*3.71*					*52, 12, 24*	*3.21*
Right PMC	6	**43, 3, 48***	**4.18**			*40, 9, 48*	*3.87*	*40, 9, 48*	*3.75*
Right SFG	10	*25, 51, 6*	*4.01*					*27, 51, 6*	*3.28*
Left MOG	19			*−29, −87, 18*	*3.53*			*−29, −87, 18*	*3.50*

following, we will discuss the functions related to the structures constituting this network.

The PMC is important for verbal WM [for an overview, see Baddeley (2003)] and rehearsal of pitch information (Koelsch *et al.*, 2009; Schulze *et al.*, in press). Furthermore, this structure seems to be involved in strategy-based WM of auditory–verbal sequences (Bor *et al.*, 2004), but not in strategy-based WM of visual–spatial material (Bor *et al.*, 2003). In addition, the part of the PMC (Talairach coordinate: 43, 3, 48) that was activated more strongly during the tonal > atonal contrast in the present study has been suggested to facilitate serial prediction for auditory sequences (Schubotz *et al.*, 2003; Schubotz, 2007).

Previous studies showed the involvement of the IPCS in strategy-based WM in different modalities. Stronger activation of the right IPCS was observed during the encoding of structured than of unstructured spatial patterns (Bor *et al.*, 2003), and during the encoding of structured than of unstructured auditory–verbal sequences (Bor *et al.*, 2004). Furthermore, activation of the IPCS increased bilaterally after visual WM training (Moore *et al.*, 2006). Interestingly, studies investigating neural correlates of the processing of musical structure, using syntactically regular vs. irregular music chord functions, have also repeatedly reported IPCS activation (Koelsch *et al.*, 2002a; Koelsch, 2005; Koelsch & Siebel, 2005).

Neuroimaging results typically suggest that during WM performance, frontal and parietal structures are co-activated (Baddeley, 2003; Bor *et al.*, 2003, 2004; Bor & Owen, 2007]. That is, these structures are suggested to form a functional network facilitating WM. In the present study, the left IPS was activated more strongly in musicians in the tonal condition (in which maintenance was facilitated because of the musical structure) than in the atonal condition. The IPS has been reported to be involved in tonal WM (Koelsch *et al.*, 2009; Schulze *et al.*, in press) and encoding of structured material (Bor *et al.*, 2004), and it has been shown that IPS activity increases after participants trained on a visual WM task (Olesen *et al.*, 2004; Moore *et al.*, 2006). This indicates that the IPS, together with the PMC and the IPCS, forms a (pre)frontal–parietal network subserving WM for structured material.

Bor *et al.* (2003, 2004) observed stronger involvement of the DLPFC ROIs (significant for the left hemisphere; tendency for the right hemisphere) during the encoding of structured material. We observed a tendency for there to be stronger activation of the right DLPFC ROI during the WM rehearsal of tonal (structured) material. This corroborates the view that the DLPFC is involved in online performance monitoring and the manipulation of information during WM tasks (Petrides *et al.*, 1993a,b; D'Esposito *et al.*, 1999; Owen *et al.*, 1999; Petrides, 2000; Fletcher & Henson, 2001; Curtis & D'Esposito, 2003). No activation difference between the structured (tonal) and unstructured (atonal) condition was observed for the VLPFC, in accordance with previous findings (Bor *et al.*, 2003).

But what cognitive mechanism or strategy might underlie musicians' superior performance and different activation pattern during WM processing of tonal as compared with atonal sequences? Bor & Owen (2007) investigated which strategy could be used during the strategic recoding of visually presented digits – a mathematical or a mnemonic strategy. In the mathematical condition, participants could additionally use mathematical relations between digits. Thus, mathematical procedures such as addition and subtraction allowed participants to deduce some digits on the basis of a stored subset of digits. Hence, fewer digits had to be remembered. For the mnemonic condition, participants had been given the opportunity to learn the presented sequences before. The activation pattern associated with the mathematical, but not the mnemonic, strategy included activation of the PMC and activation of the IPCS (Bor & Owen, 2007). In the present study, these structures were also involved in the rehearsal of tonal as compared with atonal sequences, suggesting similarities with the previously investigated mathematical strategy (Bor & Owen, 2007). Musicians might have been able to extract the interval information, that is, the relations between the tones, for the structured tonal, but not the unstructured, sequences. Thus, the interval information could have enabled musicians to cluster the tonal sequences, leading to better WM performance for tonal sequences.

Another important aspect that deserves discussion is the potential influence of task difficulty. Whereas musicians showed better behavioural performance for the tonal condition than for the atonal condition, as well as functional differences between the tonal and the atonal conditions, neither behavioural nor neuroimaging differences were observed for non-musicians. Although numerous studies have shown that non-musicians, using their implicit knowledge of musical regularities, are also capable of determining the key information of pitches (e.g. Koelsch *et al.*, 2000; Bigand & Poulin-Charronnat, 2006), there is also broad evidence that musically trained listeners have a more elaborate sense of key when listening to sequences of tones, or chords, that belong to one key (e.g. Krumhansl & Shepard, 1979; Koelsch *et al.*, 2002b). We propose that the lack of behavioural and neuroimaging differences for non-musicians during tonal and atonal rehearsal suggests that the structural differences between the tonal (all tones belong to one tonality, and three of them form a triad) and atonal (consisting of neither triad nor key) sequences were too subtle to facilitate auditory WM in non-musicians.

Even for musicians, the overall performance was not very high, and the performance difference between the structured and the unstructured conditions was quite small as compared with other studies (Bor *et al.*, 2003, 2004; Moore *et al.*, 2006; Bor & Owen, 2007). This indicates that the WM task presented in this study was relatively difficult, and might have required more active rehearsal processes than other WM tasks described in the literature (Bor *et al.*, 2003, 2004; Moore *et al.*, 2006; Bor & Owen, 2007). This could explain why activation of the structures (IPCS, PMC and IPS) that we observed during the WM rehearsal delay had been reported previously already during WM encoding (Bor *et al.*, 2003, 2004; Bor & Owen, 2007).

Importantly, it has been shown that increasing task difficulty is reflected by an increase in the BOLD response in prefrontal (Duncan & Owen, 2000; Wager & Smith, 2003) and parietal (Wager & Smith, 2003) areas. Because musicians showed more activations during the tonal than during the atonal WM task (in the presence of better task performance), it is unlikely that these differences are simply attributable to differences in task performance. This has also been observed in other studies, where an increase in activation was associated with a decrease in task difficulty (Bor *et al.*, 2003, 2004; Bor & Owen, 2007).

During the encoding of atonal as compared with tonal sequences, an increased BOLD response was observed in the pars orbitalis of the right IFG (BA 47) in musicians, but not in non-musicians. This activation was unexpected, and not included in our hypotheses; therefore, it is necessary for future studies to replicate this finding, and specify the role that BA 47 might play in the encoding of tonal and atonal pitch sequences. Perhaps the activation of BA 47 observed in the present study is attributable to the greater unexpectedness of tones in the atonal condition, where tones could be less well predicted than in the tonal condition. This might correspond to a previous observation of BA 47 activation during the detection of structurally unexpected as compared with expected music chords (Koelsch, 2005; Koelsch *et al.*, 2005).

In summary, our data in musicians showed that a lateral (pre)frontal–parietal network (including the right IPCS and PMC, as well as the left IPS) is more strongly involved during strategy-based WM processing for non-verbal auditory stimuli. A similar network has previously been reported to be involved during strategy-based WM processing for visual and auditory–verbal stimuli (Bor *et al.*, 2003, 2004; Bor & Owen, 2007), pointing towards a modality-independent (pre)frontal–parietal network subserving strategy-based WM.

Acknowledgements

This study was supported by a grant from the German Research Foundation (Deutsche Forschungsgemeinschaft) awarded to S. Koelsch (KO 2266/4-1), and by the Max Planck Society.

Abbreviations

BA, Brodmann area; BOLD, blood oxygen level-dependent; DLPFC, dorsolateral prefrontal cortex; fMRI, functional magnetic resonance imaging; IFG, inferior frontal gyrus; IPCS, inferior precentral sulcus; IPL, inferior parietal lobe; IPS, intraparietal sulcus; LPFC, lateral prefrontal cortex; MOG, middle occipital gyrus; PMC, premotor cortex; ROI, region of interest; SEM, standard error of the mean; SFG, superior frontal gyrus; SPM, statistical parametric map; SVC, small volume correction; VLPFC, ventrolateral prefrontal cortex; WM, working memory.

References

Baddeley, A.D. (2003) Working memory: looking back and looking forward. *Nat. Rev. Neurosci.*, **4**, 829–839.

Baddeley, A.D., Thomson, N. & Buchanan, L. (1975) Word length and the structure of short-term memory. *J. Verb. Learn. Verb. Behav.*, **14**, 575–589.

Bigand, E. & Poulin-Charronnat, B. (2006) Are we ‘experienced listeners’? A review of the musical capacities that do not depend on formal musical training. *Cognition*, **100**, 100–130.

Bor, D. & Owen, A.M. (2007) A common prefrontal–parietal network for mnemonic and mathematical recoding strategies within working memory. *Cereb. Cortex*, **17**, 778–786.

Bor, D., Duncan, J., Wiseman, R.J. & Owen, A.M. (2003) Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron*, **37**, 361–367.

Bor, D., Cumming, N., Scott, C.E. & Owen, A.M. (2004) Prefrontal cortical involvement in verbal encoding strategies. *Eur. J. Neurosci.*, **19**, 3365–3370.

Curtis, C.E. & D’Esposito, M. (2003) Persistent activity in the prefrontal cortex during working memory. *Trends Cogn. Sci.*, **7**, 415–423.

D’Esposito, M., Postle, B.R., Ballard, D. & Lease, J. (1999) Maintenance versus manipulation of information held in working memory: an event-related fMRI study. *Brain Cogn.*, **41**, 66–86.

Dowling, W.J. (1991) Tonal strength and melody recognition after long and short delays. *Percept. Psychophys.*, **50**, 305–313.

Duncan, J. & Owen, A.M. (2000) Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.*, **23**, 475–483.

Ericsson, K.A., Chase, W.G. & Faloon, S. (1980) Acquisition of a memory skill. *Science*, **208**, 1181–1182.

Fletcher, P.C. & Henson, R.N. (2001) Frontal lobes and human memory: insights from functional neuroimaging. *Brain*, **124**, 849–881.

Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A. & Noll, D.C. (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn. Reson. Med.*, **33**, 636–647.

Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M.D. & Turner, R. (1998) Event-related fMRI: characterizing differential responses. *Neuroimage*, **7**, 30–40.

Fujioka, T., Trainor, L.J., Ross, B., Kakigi, R. & Pantev, C. (2004) Musical training enhances automatic encoding of melodic contour and interval structure. *J. Cogn. Neurosci.*, **16**, 1010–1021.

Fujioka, T., Trainor, L.J., Ross, B., Kakigi, R. & Pantev, C. (2005) Automatic encoding of polyphonic melodies in musicians and nonmusicians. *J. Cogn. Neurosci.*, **17**, 1578–1592.

Gaab, N., Gaser, C. & Schlaug, G. (2006) Improvement-related functional plasticity following pitch memory training. *Neuroimage*, **31**, 255–263.

Gobet, F. (1998) Expert memory: a comparison of four theories. *Cognition*, **66**, 115–152.

Gobet, F., Lane, P.C., Croker, S., Cheng, P.C., Jones, G., Oliver, I. & Pine, J.M. (2001) Chunking mechanisms in human learning. *Trends Cogn. Sci.*, **5**, 236–243.

Hall, D.A., Haggard, M.P., Akeroyd, M.A., Palmer, A.R., Summerfield, A.Q., Elliott, M.R., Gurney, E.M. & Bowtell, R.W. (1999) ‘Sparse’ temporal sampling in auditory fMRI. *Hum. Brain Mapp.*, **7**, 213–223.

Jancke, L. (2009) The plastic human brain. *Restor. Neurol. Neurosci.*, **27**, 521–538.

Keenan, J.P., Thangaraj, V., Halpern, A.R. & Schlaug, G. (2001) Absolute pitch and planum temporale. *Neuroimage*, **14**, 1402–1408.

Koelsch, S. (2005) Neural substrates of processing syntax and semantics in music. *Curr. Opin. Neurobiol.*, **15**, 207–212.

Koelsch, S. & Siebel, W.A. (2005) Towards a neural basis of music perception. *Trends Cogn. Sci.*, **9**, 578–584.

Koelsch, S., Schroger, E. & Tervaniemi, M. (1999) Superior pre-attentive auditory processing in musicians. *Neuroreport*, **10**, 1309–1313.

Koelsch, S., Gunter, T., Friederici, A.D. & Schroger, E. (2000) Brain indices of music processing: ‘nonmusicians’ are musical. *J. Cogn. Neurosci.*, **12**, 520–541.

Koelsch, S., Gunter, T.C., v Cramon, D.Y., Zysset, S., Lohmann, G. & Friederici, A.D. (2002a) Bach speaks: a cortical ‘language-network’ serves the processing of music. *Neuroimage*, **17**, 956–966.

Koelsch, S., Schmidt, B.H. & Kansok, J. (2002b) Effects of musical expertise on the early right anterior negativity: an event-related brain potential study. *Psychophysiology*, **39**, 657–663.

Koelsch, S., Schroger, E. & Gunter, T.C. (2002c) Music matters: preattentive musicality of the human brain. *Psychophysiology*, **39**, 38–48.

Koelsch, S., Fritz, T., Schulze, K., Alsop, D. & Schlaug, G. (2005) Adults and children processing music: an fMRI study. *Neuroimage*, **25**, 1068–1076.

Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Muller, K. & Gruber, O. (2009) Functional architecture of verbal and tonal working memory: an FMRI study. *Hum. Brain Mapp.*, **30**, 859–873.

Krumhansl, C.L. (1979) The psychological representation of musical pitch in a tonal context. *Cogn. Psychol.*, **11**, 346–374.

Krumhansl, C.L. & Shepard, R.N. (1979) Quantification of the hierarchy of tonal functions within a diatonic context. *J. Exp. Psychol. Hum. Percept. Perform.*, **5**, 579–594.

Lohmann, G., Muller, K., Bosch, V., Mentzel, H., Hessler, S., Chen, L., Zysset, S. & von Cramon, D.Y. (2001) LIPSIA – a new software system for the evaluation of functional magnetic resonance images of the human brain. *Comput. Med. Imaging Graph.*, **25**, 449–457.

Lohmann, G., Neumann, J., Mueller, K., Lepsien, J. & Turner, R. (2008) The multiple comparison problem in fMRI – a new method based on anatomical priors. *Proceedings of the MICCAI Workshop on Analysis of Functional Images, September 10*, 2008, New York.

Mandler, G. (1967). The development of free and constrained conceptualization and subsequent verbal memory. In Spence, K.W. & Spence, J.T. (Ed.), *The Psychology of Learning and Motivation*. Academic Press, New York, pp. 327–372.

Miller, G.A. (1956) The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.*, **63**, 81–97.

Moore, C.D., Cohen, M.X. & Ranganath, C. (2006) Neural mechanisms of expert skills in visual working memory. *J. Neurosci.*, **26**, 11187–11196.

Mugler, J.P. III & Brookeman, J.R. (1990) Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP RAGE). *Magn. Reson. Med.*, **15**, 152–157.

Munte, T.F., Altenmuller, E. & Jancke, L. (2002) The musician’s brain as a model of neuroplasticity. *Nat. Rev. Neurosci.*, **3**, 473–478.

Oldfield, R.C. (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, **9**, 97–113.

Olesen, P.J., Westerberg, H. & Klingberg, T. (2004) Increased prefrontal and parietal activity after training of working memory. *Nat. Neurosci.*, **7**, 75–79.

Owen, A.M., Herrod, N.J., Menon, D.K., Clark, J.C., Downey, S.P., Carpenter, T.A., Minhas, P.S., Turkheimer, F.E., Williams, E.J., Robbins, T.W., Sahakian, B.J., Petrides, M. & Pickard, J.D. (1999) Redefining the functional organization of working memory processes within human lateral prefrontal cortex. *Eur. J. Neurosci.*, **11**, 567–574.

Pantev, C., Roberts, L.E., Schulz, M., Engelien, A. & Ross, B. (2001) Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport*, **12**, 169–174.

Pechmann, T. & Mohr, G. (1992) Interference in memory for tonal pitch: implications for a working-memory model. *Mem. Cognit.*, **20**, 314–320.

Petrides, M. (2000) The role of the mid-dorsolateral prefrontal cortex in working memory. *Exp. Brain Res.*, **133**, 44–54.

Petrides, M., Alivisatos, B., Evans, A.C. & Meyer, E. (1993a) Dissociation of human mid-dorsolateral from posterior dorsolateral frontal cortex in memory processing. *Proc. Natl. Acad. Sci. USA*, **90**, 873–877.

Petrides, M., Alivisatos, B., Meyer, E. & Evans, A.C. (1993b) Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proc. Natl. Acad. Sci. USA*, **90**, 878–882.

Savage, C.R., Deckersbach, T., Heckers, S., Wagner, A.D., Schacter, D.L., Alpert, N.M., Fischman, A.J. & Rauch, S.L. (2001) Prefrontal regions supporting spontaneous and directed application of verbal learning strategies: evidence from PET. *Brain*, **124**, 219–231.

Schubotz, R.I. (2007) Prediction of external events with our motor system: towards a new framework. *Trends Cogn. Sci.*, **11**, 211–218.

Schubotz, R.I., von Cramon, D.Y. & Lohmann, G. (2003) Auditory what, where, and when: a sensory somatotopy in lateral premotor cortex. *Neuroimage*, **20**, 173–185.

Schulze, K., Zysset, S., Mueller, K., Friederici, A.D. & Koelsch, S. (in press) Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. *Hum. Brain Mapp.*, doi: 10.1002/hbm.21060.

Talairach, P. & Tournoux, J. (1988) *A Stereotactic Coplanar Atlas of the Human Brain*. Thieme, Stuttgart.

Tervaniemi, M., Rytkonen, M., Schroger, E., Ilmoniemi, R.J. & Naatanen, R. (2001) Superior formation of cortical memory traces for melodic patterns in musicians. *Learn. Mem.*, **8**, 295–300.

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A. & Schroger, E. (2005) Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. *Exp. Brain Res.*, **161**, 1–10.

Thirion, J.P. (1998) Image matching as a diffusion process: an analogy with Maxwell's demons. *Med. Image Anal.*, **2**, 243–260.

Tulving, E. (1962) Subjective organization in free recall of 'unrelated' words. *Psychol. Rev.*, **69**, 344–354.

Wager, T.D. & Smith, E.E. (2003) Neuroimaging studies of working memory: a meta-analysis. *Cogn. Affect. Behav. Neurosci.*, **3**, 255–274.

Williamson, V.J., Baddeley, A.D. & Hitch, G.J. (2010) Musicians' and nonmusicians' short-term memory for verbal and musical sequences: comparing phonological similarity and pitch proximity. *Mem. Cognit.*, **38**, 163–175.