Neuroarchitecture of Verbal and Tonal Working Memory in Nonmusicians and Musicians

Katrin Schulze,[1,2] **Stefan Zysset,**[1] **Karsten Mueller,**[1] **Angela D. Friederici,**[1] **and Stefan 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, London, United Kingdom*
[3]*Cluster of Excellence Languages of Emotion, Freie Universität Berlin, Berlin, Germany*

Abstract: Working memory (WM) for auditory information has been thought of as a unitary system, but whether WM for verbal and tonal information relies on the same or different functional neuroarchitectures has remained unknown. This fMRI study examines verbal and tonal WM in both nonmusicians (who are trained in speech, but not in music) and highly trained musicians (who are trained in both domains). The data show that core structures of WM are involved in both tonal and verbal WM (Broca's area, premotor cortex, pre-SMA/SMA, left insular cortex, inferior parietal lobe), although with significantly different structural weightings, in both nonmusicians and musicians. Additionally, musicians activated specific subcomponents only during verbal (right insular cortex) or only during tonal WM (right globus pallidus, right caudate nucleus, and left cerebellum). These results reveal the existence of two WM systems in musicians: A phonological loop supporting rehearsal of phonological information, and a tonal loop supporting rehearsal of tonal information. Differences between groups for tonal WM, and between verbal and tonal WM within musicians, were mainly related to structures involved in controlling, programming and planning of actions, thus presumably reflecting differences in action-related sensorimotor coding of verbal and tonal information. *Hum Brain Mapp 32:771–783, 2011.* © 2010 **Wiley-Liss, Inc.**

Key words: working memory; auditory processing; language; music; functional plasticity

---

INTRODUCTION

The memory system that maintains verbal information for a few moments to establish a representation of strings of verbal elements has been referred to as verbal working memory (WM). Usually, verbal WM refers to a memory system retaining phonemes, syllables, and words. However, pitch processing also plays an important role for the perception of spoken language, and recent theories of the evolution of language proposed that the origins of language are partly grounded on cognitive systems dedicated to the processing of pitch [Fitch, 2006; Wallin et al., 1999]. Although auditory WM has so far been thought of as a unitary system, these assumptions give rise to the question whether there are two separate WM systems, one

primarily serving the processing of verbal, and one primarily serving the processing of pitch information.

Surprisingly, however, the neural correlates of WM have so far mainly been investigated with regards to language [Baddeley, 2003]. Research in this area has primarily been based on the Baddeley and Hitch WM model [Baddeley, 2003; Baddeley and Hitch, 1974]. In this model, the system that processes language-related information is referred to as the phonological loop, which is further subdivided into two components: The articulatory rehearsal process, which is comparable to subvocal speech, relies on an articulatory code [Baddeley, 2003; Conrad and Hull, 1964; Jacquemot and Scott, 2006], and is suggested to be neurally implemented by Broca's area, premotor areas and the supplementary motor area (SMA) [Baddeley, 2003; Gruber and von Cramon, 2003; Paulesu et al., 1993], the cerebellum [Chen and Desmond, 2005; Ravizza et al., 2004], as well as the insular cortex [Bamiou et al., 2003; Paulesu et al., 1993]. The second component is a passive phonological store which has been suggested to comprise inferior parietal lobe (IPL) [Awh et al., 1996; Gruber and von Cramon, 2003] and perhaps superior parietal lobe (SPL) [Awh et al., 1996; Chen and Desmond, 2005], as well as the intraparietal sulcus (IPS) [Gruber and von Cramon, 2003]. The localization of the phonological store in the parietal lobe, however, is still partly controversial [Buchsbaum and D'Esposito, 2008; Hickok et al., 2003].

So far, only a few behavioral studies investigated whether a different sub-system such as a "tonal loop" [Pechmann and Mohr, 1992] exists in addition to the phonological loop. The available studies, however, do not provide a consistent picture [Chan et al., 1998; Deutsch, 1974; Salame and Baddeley, 1989; Semal et al., 1996]. Furthermore, the loci of activations described in the few functional neuroimaging studies [Gaab et al., 2003; Zatorre et al., 1994] that investigated the neural network underlying WM for tonal stimuli are remarkably similar to those observed in verbal WM experiments [Baddeley, 2003].

To our knowledge, only two neuroimaging studies have directly compared the neural networks underlying WM for tonal and verbal stimuli [Hickok et al., 2003; Koelsch et al., 2008], and both the studies reported a considerable overlap of neural resources underlying WM for verbal and tonal information: A fronto-parietal network (comprised of premotor cortex, Broca's area, and in one of the two studies also the IPL and the cerebellum [Koelsch et al., 2008]), as well as the planum temporale/area Spt (Sylvian fissure at the temporal-parietal boundary) which is assumed to play a role in the translation of auditory-sensory information into motor representations [Hickok and Poeppel, 2007; Warren et al., 2005]. These observations corroborate the idea that the formation and maintenance of articulatory sensorimotor codes does not only serve the rehearsal of verbal, but also the rehearsal of tonal WM [Hickok et al., 2003; Jacquemot and Scott, 2006; Koelsch et al., 2008]. However, both studies [Hickok et al., 2003; Koelsch et al., 2008] explored WM only in nonmusicians, who are exten-sively trained in the speech, but not in the music domain [for a study reporting an overlap between tone and language processing in musicians with Absolute Pitch see Oechslin et al., 2009]. That is, because speech is a fundamental means of human communication (and a skill humans acquire during their early childhood), both nonmusicians and musicians were considered to be trained in recognizing and producing speech. Thus, for a more balanced comparison between WM for verbal and tonal information, the present functional magnetic resonance imaging (fMRI) study examined verbal and tonal WM in both nonmusicians and highly trained musicians. Musical training includes the coupling of auditory input and motor output [Bangert et al., 2006; Drost et al., 2005], leading to long-term training effects for associations between pitch information and corresponding motor actions [Zatorre et al., 2007]. Here, we aimed at testing (a) whether instrumental musicians engage different sensorimotor-related brain structures for verbal compared to tonal WM and (b) whether musicians engage different neural networks for tonal WM compared to nonmusicians.

MATERIALS AND METHODS

Participants

Seventeen right-handed nonmusicians (9 male, age range: 21–29, $M = 25.47$ years, SEM = 0.61 years) and sixteen right-handed musicians (9 male, age range: 20–27, $M = 23.50$ years, SEM = 0.63 years) took part in this study. Musicians studied an instrument at the "University of Music and Theatre Mendelsohn Bartholdy" in Leipzig (eight participants studied the piano, four a woodwind instrument, three a string instrument, and one a brass instrument). Nonmusicians neither had formal musical training (besides regular lessons in school) nor were any of them playing an instrument. None of the musicians reported to possess absolute pitch (AP), and this was verified using an established AP test [Keenan et al., 2001]. Participants were right-handed according to the Edinburgh Handedness Inventory [Oldfield, 1971]; mean lateralization quotient was 97% for musicians, and 95% for nonmusicians, with no significant difference between both 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

Each auditory stimulus consisted of a spoken syllable and a simultaneously presented sine wave tone, both syllable and sine wave tone had the same loudness. That is, tones and syllables were presented simultaneously (instead of trials with tones only, and other trials with

syllables only), so that the auditory input was identical for the tonal and the verbal task.

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 ranged from 261 to 523 Hz (one octave), resulting in 13 different tones of 12 pitch classes. The syllables were spoken by a professional male speaker and were the German names of the scale tones (e.g., gis [g-sharp], c etc.), resulting in 12 different syllables (tones with a frequency of 261 and 523 Hz are both referred to as c). In each experimental trial, five such stimuli were presented in a sequence. Each stimulus had a duration of 400 ms, with periods of 150 ms of silence between them, resulting in a sequence length of 2,600 ms (see Fig. 1a).

Procedure

There were two experimental conditions (verbal and tonal), and one control condition, which is referred to as pink noise control condition (see below for explanation). In the verbal and tonal conditions, participants subsequently listened to sequences of five auditory stimuli, the sample stimuli sequence (see Fig. 1a) and then rehearsed internally for 4,200 up to 6,200 ms either the syllables during the verbal condition or the tones during the tonal condition (see *silence* in Fig. 1a). At the end of each trial a test stimulus was presented, consisting of the simultaneous presentation of one syllable and one sine wave tone (see *test stimulus* in Fig. 1a), and participants had to indicate by a button press, whether the syllable in the verbal condition, or the sine wave tone in the tonal condition, had already been presented during the sample stimuli sequence (see *response* in Fig. 1a). It was not possible to solve the tonal task by paying attention merely to the syllables instead of to the tones, because tones and syllables (which were names of the tones) were presented together systematically in only 50% of the sequences, i.e., the frequency of tone for the tone stimulus and the tone name for the syllable stimulus were only congruent in 50% of all sequences (e.g., tone = g′, syllable = "ge"). This issue was also pointed out to the participants before the experiment.

During the pink noise control condition, in which participants did not perform a verbal or tonal WM task, pink noise was presented instead of the sample stimuli sequence, and a test stimulus of pink noise was presented to control for the auditory perceptual input. Additionally, participants pressed a predefined button after the end of the sequence to account for the motor response.

The sequences were presented pseudorandomly in a blocked design, and participants started either with a verbal or tonal block (counterbalanced across participants and groups) for a total of 10 blocks. Each block consisted of 16 experimental sequences (±1 sequence), resulting in 160 experimental sequences, 80 presented during the verbal WM task, and 80 presented during the tonal WM task. At the
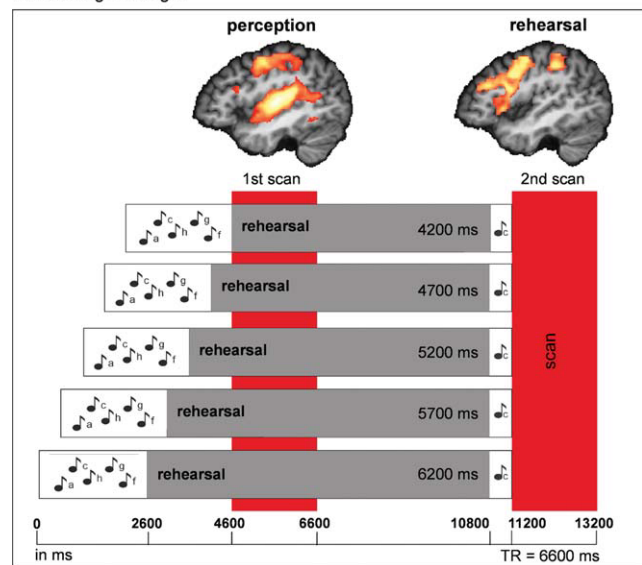

Figure 1.

Experimental task and scanning paradigm. **a**: Participants listened to sequences of five auditory stimuli (each of which consisted of a sine wave tone and a simultaneously presented spoken syllable). Subsequently, participants rehearsed either the syllables (verbal condition) or the sine wave tones (tonal condition) internally. At the end of each trial a probe stimulus was presented, and participants had to indicate by a button press whether the syllable (verbal condition) or the sine wave tone (tonal condition) had been presented in the initial sequence. **b**: The scanning paradigm was a modified version of the sparse temporal sampling technique (scans are indicated by the red bars, each scan consisted of a clustered volume acquisition covering the entire brain). Auditory stimuli were presented in the time intervals between scans (i.e., in the absence of the scanner noise), the rehearsal times differed between 4.2 s and 6.2 s (see Materials and Methods). Brain activation in the left hemisphere associated with perception (verbal and tonal perception vs. silence; $x = -43$; $z > 3.09$; corr. for mult. comp., $P < 0.05$) and rehearsal (verbal and tonal rehearsal vs. pink noise nonrehearsal; $x = -44$; $z > 3.09$; corr. for mult. comp., $P < 0.05$) for nonmusicians and musicians is visualised. During perception (first scan) significant activation was observed in Heschl's gyrus (left Heschl's gyrus $z = 8.49$ (-44, -21, 9) and right Heschl's gyrus $z = 7.944$ (43, -19, 9)). No activation was observed in Heschl's gyrus during verbal and tonal rehearsal (second scan). Note that only data of second scan (rehearsal) were investigated in this WM paper.

beginning of each block, a visual cue (see Fig. 1a) indicated whether the WM task for the next block was verbal or tonal WM. In each block, 3 pink noise control sequences (±1 sequence) were presented additionally, resulting in 30 pink noise control sequences overall. The blocks and sequences used were identical for the tonal and verbal WM task. Participants were repeatedly instructed not to sing or hum aloud during the scanning session.

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 noise. Two scans per trial were conducted, allowing scanning of the hemodynamic response correlated with the processes active during (a) the perception (first scan) and (b) the rehearsal period of the stimuli (second scan, see Fig. 1b). We only present here the results of the WM rehearsal (i.e., data obtained with the second scan), because inclusion of the perception data would go beyond the scope of this article. As depicted in Figure 1b, five different onsets of the auditory sequence relative to the first scan differing in their onsets by 500 ms were generated to allow an optimal sampling of the hemodynamic response associated with the perception, i.e. the first scan occurred at 0 ms, the second at 500 ms, the third at 1,000 ms, the fourth at 1,500 ms, and the fifth at 2,000 ms after the auditory presentation. The rehearsal time differed in length accordingly (see Fig. 1b).

The experiment was carried out on a 3T scanner (Siemens TRIO, Erlangen). Before each functional session, an anatomical reference data set was acquired for each participant, which was standardized to the Talairach stereotactic space [Talairach and Tournoux, 1988]. A bunched gradient-echo EPI sequence was used with a TE of 30 ms, a flip angle of 90° and a TR of 6,600 ms and an acquisition bandwidth of 100 kHz. Twenty-four axial slices were acquired rapidly within ∼1,600 ms, so that no scanning occurred during the rest of the TR. The matrix dimensions were 64 × 64 with a field of view (FOV) of 192 mm, resulting in a voxel size of 3 × 3 mm^2, slice thickness of 4 mm, and an interslice gap of 1 mm.

Data Analysis

The data preprocessing, the statistical analysis, and the visualization of the fMRI data were performed using the software package LIPSIA [Lohmann et al., 2001]. An off-line motion correction was performed on the functional images, using a Siemens motion correction protocol (Siemens, Erlangen, Germany) and a matching metric based on linear correlation. The cut-off frequency of the temporal high-pass filter, which was used to remove baseline drifts, was 1/300 Hz. A spatial gaussian filter with a FWHM of 5.65 mm (sigma = 0.8 voxels) was used to improve the

signal-to-noise-ratio of the data. To align the functional data with a 3D stereotactic coordinate reference system, a rigid linear registration was performed. The translational and rotational parameters were acquired using the anatomical reference brain. In the following, the calculated translational and rotational parameters were used to transform the functional data set to the stereotactic coordinate system, by using a trilinear interpolation. Then the functional data were linearly normalized. Because the brains of musicians have been shown to differ anatomically from those of nonmusicians [Munte et al., 2002], the functional images were only linearly normalized, because a nonlinear normalization would have removed the anatomical specificity between the groups. The statistical evaluation used the general linear model (GLM). The design matrix was generated using one synthetic hemodynamic response function [Friston et al., 1998]. For each participant, β-values were estimated for different contrasts for each voxel. After the individual functional datasets were all aligned to the same stereotactic reference space, the single-participant contrast-images were entered into a second-level random effects analysis for the relevant contrasts. The group analysis consisted of a one-sample t-test across the contrast images of all subjects indicating whether observed differences were significantly distinct from zero.

Subsequently, t-values were transformed into z scores and statistical parametric maps (SPMs) were generated. The results were corrected for multiple comparisons using cluster-size (minimum cluster size of 10 voxels) and cluster-value thresholds obtained by Monte-Carlo simulations using a significance level of $P < 0.05$ (clusters in the resulting maps were obtained using a z-value threshold of 2.57) [Forman et al., 1995].

Data from tonal and from verbal rehearsal were compared separately to the pink noise nonrehearsal condition (tonal rehearsal vs. nonrehearsal, verbal rehearsal vs. nonrehearsal). In addition, a conjunction analysis was computed for these two contrasts (tonal rehearsal vs. nonrehearsal, verbal rehearsal vs. nonrehearsal). The conjunction analysis indicates which areas are activated during both verbal and tonal WM, and which are significantly activated during verbal (but not tonal) or tonal (but not verbal) WM.

These SPMs were also used for the analysis of regions of interest (ROI). Areas that were found to be significantly activated were subjected to further post hoc analyses. Specifically, we tested whether contrast values in ROIs (spheres with a radius of 6 mm) were significantly different from zero in these different conditions. We also performed ROI analyses between groups in order to investigate the difference between musicians and nonmusicians. The ROI coordinates are provided in Tables I and II; all ROIs were spheres with a radius of 6 mm centered around these coordinates.

A regression analysis, to investigate correlations between performance and activations in nonmusicians and musicians, is described in the Supporting Information.

TABLE I. Activations for verbal (vs. nonrehearsal) and tonal (vs. nonrehearsal) in nonmusicians (and comparison nonmusicians > musicians)

Verbal WM in Nonmusicians										
	Left hemisphere					Right hemisphere				
	Verbal > nonrehearsal			Verbal > tonal	Nonmusicians > musicians	Verbal > nonrehearsal			Verbal > tonal	Nonmusicians > musicians
Area	BA	Coordinate (x, y, z)	z-value (SPM)	P-value (ROI)	P-value (ROI)	BA	Coordinate (x, y, z)	z-value (SPM)	P-value (ROI)	P-value (ROI)
Pre-SMA[a]	6	−2, 19, 45	5.242	.003[v]	n.s.					
SMA	6	−2, 4, 57	4.845	.004[v]						
Mid-DLPFC	9/46	−44, 25, 33	4.455	.02[v]						
Pars opercularis[b]	44	−50, 16, −3	3.785	.004[v]						
IFS	46	−44, 31, 21	4.531	.003[v]						
Ant. insula		−32, 25, 3	4.351	.0001[v]						
PMd	6	−32, −5, 54	3.923	.03[v]		6	34, −14, 54	4.272	n.s.	.005[nm]
PMv	6	−50, 7, 27	5.641	.0001[v]		6	34, −2, 36	3.459	.003[v]	
						6	46, −5, 42	3.480	n.s.	
M1	4	−50, −11, 45	3.856	.03[v]						
IPL	40	−41, −41, 54	4.371	.003[v]		40	40, −29, 39	3.658	.02[v]	n.s.
	40	−35, −38, 39	5.906	.0001[v]						
IPS	7/40	−26, −53, 39	4.272	.0002[v]						
Cerebellum		−8, −47, −9	3.745	n.s.	n.s.		22, −65, −18	4.752	.006[v]	
		−17, −47, −18	4.009	n.s.	n.s.		16, −71, −33	3.727	.05[v]	
		−35, −65, −18	3.465	n.s.	.007[nm]					
Brainstem							10, −35, −24	3.530	n.s.	

Tonal WM in Nonmusicians									
	Left hemisphere					Right hemisphere			
	Tonal > nonrehearsal			Tonal > verbal	Nonmusicians > musicians	Tonal > nonrehearsal		Tonal > verbal	Nonmusicians > musicians
Area	BA	Coordinate (x, y, z)	z-value (SPM)	P-value (ROI)	P-value (ROI)	BA	Coordinate (x, y, z)	z-value (SPM)	
Cingulate gyrus	32	−8, 19, 33	3.420	n.s.	n.s.				
Pre-SMA/SMA	6	−5, 4, 57	4.591	n.s.	n.s.				
Pars orbitalis	47	−41, 28, −3	4.926	n.s.	.02[nm]				
Pars triangularis	46	−50, 37, 6	2.883	n.s.	n.s.				
PMv[c]	6	−47, −11, 48	3.586	n.s.	n.s.				
PMv	6	−38, −5, 36	3.189	n.s.	n.s.				
	6	−56, 1, 30	3.877	n.s.	n.s.				
PMv[d]	6	−59, 4, 18	4.173	n.s.	n.s.				
IPL	40	−47, −32, 48	3.732	n.s.	n.s.				
	40	−35, −35, 39	3.495	n.s.	n.s.				

Coordinates refer to standard stereotactic space [Talairach and Tournoux, 1988].

BA: Brodmann Area; z-values (SPM) indicate activations for verbal > nonrehearsal or tonal > nonrehearsal ($z > 2.57$, corrected for multiple comparisons, minimum cluster size of 10 voxels); P-value verbal > tonal (ROI): P-value obtained in the ROI analysis comparing verbal > tonal rehearsal for verbal WM in nonmusicians; P-value tonal > verbal (ROI): P-value obtained in the ROI analysis comparing tonal > verbal rehearsal for tonal WM in nonmusicians; P-value nonmusicians > musicians (ROI): P-value obtained in the ROI analysis comparing nonmusicians > musicians; superscript indices: v: verbal > tonal; t: tonal > verbal; nm: nonmus > mus; n.s.: not significant (blank fields indicate that no comparison was conducted, see Materials and Methods); underlined P-values were significant after Bonferroni correction; abbreviations: ant. = anterior, c = cortex.

[a]This activation also included activation in the cingulate gyrus.
[b]This activation also included activation of the anterior STG.
[c]This activation also included activation in PMd and M1.
[d]This activation also included activation in the ant. insula and pars opercularis.

TABLE II. Activations for verbal (vs. nonrehearsal) and tonal (vs. nonrehearsal) in musicians (and comparison musicians > nonmusicians)

Verbal WM in Musicians

| | | Left hemisphere | | | | | Right hemisphere | | | |
| | | Verbal > nonrehearsal | | | Verbal > Tonal | Musicians > nonmusicians | Verbal > Nonrehearsal | | | Verbal > tonal | Musicians > nonmusicians |
Area	BA	Coordinate (x, y, z)	z-value (SPM)	P-value (ROI)	P-value (ROI)	BA	Coordinate (x, y, z)	z-value (SPM)	P-value (ROI)	P- value (ROI)
Cingulate gyrus						24	13, 13, 30	4.345	n.s.	
Pre-SMA/SMA						6	4, 4, 54	5.483	n.s.	
Pars opercularis[a]	44	−50, 7, 12	4.845	.0001[v]						
Ant. insula		−29, 19, 6	5.258	.002[v]			31, 19, 3	4.617	.02[v]	.002[m]
PMd	6	−26, −17, 57	3.734	n.s.						
	6	−26, −2, 54	3.608	n.s.						
PMv[b]	6	−47, 4, 36	4.715	.006[v]		6	28, −5, 48	4.448	n.s.	
	6	−32, 1, 33	4.010	n.s.						
Paracentral lobule						6	7, −29, 63	4.183	n.s.	n.s.
Somatosensory cortex						3	31, −29, 66	3.794	n.s.	n.s.
						3	31, −29, 48	3.966	n.s.	n.s.
IPL	40	−53, −38, 48	3.394	n.s.						
IPS	7/40	−47, −44, 57	3.895	n.s.						
	7/40	−35, −50, 48	4.105	n.s.						
	7/40	−29, −62, 45	3.989	n.s.						
SPL	7	−20, −65, 57	3.987	n.s.	n.s.					
Putamen		−20, −8, 12	3.787	n.s.	n.s.					
Cerebellum							25, −65, −15	4.401	n.s.	

Tonal WM in Musicians

| | | Left hemisphere | | | | | Right hemisphere | | | |
| | | Tonal > nonrehearsal | | | Tonal > verbal | Musicians > nonmusicians | Tonal > nonrehearsal | | | Tonal > verbal | Musicians > nonmusicians |
Area	BA	Coordinate (x, y, z)	z-value (SPM)	P-value (ROI)	P-value (ROI)	BA	Coordinate (x, y, z)	z-value (SPM)		
Cingulate gyrus	32	−17, 19, 33	3.822	.03[t]	.006[m]					
Pre-SMA/SMA						6	1, 1, 54	4.960	n.s.	.04[m]
Lateral orbitofrontal c.	10	−26, 40, 12	3.391	n.s.	n.s.					
Pars opercularis	44	−56, 7, 9	3.999	n.s.	.03[m]					
Ant. insula		−29, 22, 3	4.660	n.s.	.003[m]					
		−35, 13, 9	4.514	n.s.	.005[m]					
IFS	46	−44, 25, 27	3.944	n.s.	n.s.					
	46	−44, 46, 9	3.449	.04[t]	n.s.					
IFS	8	−50, 13, 33	3.741	n.s.	.03[m]					
PMd	6	−26, −2, 51	3.897	n.s.	.002[m]	6	25, −11, 60	4.181	.05[t]	n.s.
						6	34, −17, 54	4.799	.0006[t]	n.s.
PMv[c]	6	−47, −11, 42	4.535	n.s.	n.s.	6	55, −2, 39	3.933	.02[t]	n.s.
	6	−47, 1, 36	4.366	n.s.	.03[m]					
	6	−32, 1, 30	3.738	n.s.	.006[m]					
Paracentral lobule						7	1, −32, 57	4.923	.005[t]	n.s.
IPS/IPL	40	−35, −50, 54	4.734	n.s.	.0004[m]					
SPL	7	−20, −65, 54	4.272	n.s.	.009[m]					
Cuneus	31	−26, −68, 6	3.070	.02[t]	.02[m]					
Putamen		−26, −8, 12	4.325	.03[t]	n.s.					
Globus pallidus							19, −8, 0	4.153	.01[t]	n.s.

TABLE II. (Continued)

Tonal WM in Musicians

		Left hemisphere					Right hemisphere			
		Tonal > nonrehearsal		Tonal > verbal	Musicians > nonmusicians		Tonal > nonrehearsal		Tonal > verbal	Musicians > nonmusicians
Area	BA	Coordinate (x, y, z)	z-value (SPM)	P-value (ROI)	P-value (ROI)	BA	Coordinate (x, y, z)	z-value (SPM)		
Caudate nucleus							16, −2, 21	4.479	.02^t	.02^m
Cerebellum		−20, −68, −9	3.793	.0001^t	n.s.		4, −80, −9	3.105	.02^t	n.s.
		−17, −38, −12	3.933	n.s.	n.s.		28, −71, −12	5.133	.04^t	n.s.
		−17, −53, −12	3.795	n.s.	n.s.		43, −65, −18	4.257	.02^t	n.s.
		−23, −44, −21	4.121	n.s.	n.s.					

Coordinates refer to standard stereotactic space [Talairach and Tournoux, 1988].
BA: Brodmann Area; z-values (SPM) indicate activations for verbal vs. nonrehearsal or tonal vs. nonrehearsal ($z > 2.57$, corrected for multiple comparisons, minimum cluster size of 10 voxels); P-value verbal > tonal (ROI): P-value obtained in the ROI analysis comparing verbal > tonal rehearsal for verbal WM in musicians; P-value tonal > verbal (ROI): P-value obtained in the ROI analysis comparing tonal > verbal rehearsal for tonal WM in musicians; P-value musicians > nonmusicians (ROI): P-value obtained in the ROI analysis comparing musicians > nonmusicians; superscript indices: v: verbal > tonal; t: tonal > verbal; m: mus > nonmus; n.s.: not significant (blank fields indicate that no comparison was conducted, see Materials and Methods); underlined P-values were significant after Bonferroni correction; abbreviations: ant. = anterior, c. = cortex.
[a]This activation also included activation in the IFS.
[b]The left activation also included activation in M1, the right activation also included activation in PMd.
[c]The left activation also included activation in M1.

RESULTS

Nonmusicians

Behaviorally, nonmusicians had on average 84.49% (SEM = 2.09%) correct responses in the verbal, and 56.28% (SEM = 1.76%) in the tonal condition, with the difference between the conditions being significant ($t(16) = 17.322$, $P < 0.001$, two-tailed paired t-test). Although the performance during the tonal condition was relatively low, it was significantly above chance level ($t(16) = 3.563$, $P < 0.005$, two-tailed one-sample t-test).

The conjunction analysis of the fMRI data for tonal and verbal WM in nonmusicians is shown in the left columns of Figures 2a and 3. Both conditions activated the left pars opercularis (posterior part of Broca's area), the left insular cortex, the left ventro- and dorsolateral premotor cortex (PMv and PMd), pre-SMA, SMA proper, the cingulate gyrus, and the left IPL (see also Fig. 2b and Table I for a complete list of activations; during tonal WM, the activation in the left PMv encroached on the pars opercularis). Although significantly activated during both verbal and tonal WM, all of these structures were activated more strongly during verbal compared with tonal WM (see also Table I), as shown by the ROI analysis.

The conjunction analysis also showed activations for verbal, but not tonal, WM in the left mid-dorsolateral prefrontal cortex (mid-DLPFC), the left inferior frontal sulcus (IFS), the right PMv, the left IPS, the right IPL, as well as the right cerebellum. The ROI analysis confirmed that the differences in activation between conditions (verbal and tonal WM) in these regions were significant (see also Table I).

Finally, the conjunction analysis suggests activity in a few voxels within the left pars orbitalis of the inferior frontal gyrus (IFG) and the left IPL only during tonal WM. However, the ROI analyses did not indicate any significant difference in activation between tonal and verbal rehearsal in these structures. That is, there was no brain area activated exclusively, or more strongly, during tonal compared to verbal WM in nonmusicians.

Musicians

Musicians had on average 88.13% (SEM = 1.98%) correct responses in the verbal, and 69.53% (SEM = 2.64%) in the tonal condition, with the difference between conditions being significant ($t(15) = 6.003$, $P < 0.001$, two-tailed paired t-test). Whereas musicians did not differ from nonmusicians in the verbal condition ($t(31) = 1.258$, $P > 0.25$), a two-tailed independent-samples t-test showed that they scored significantly higher than nonmusicians in the tonal condition ($t(31) = 4.218$, $P < 0.001$).

The conjunction analysis of the fMRI data for tonal and verbal WM in musicians is shown in the right columns of Figures 2a and 3, and in Figure 4. Similar to the findings in nonmusicians, both conditions activated Broca's area, the left insular cortex, the left PMv and PMd, pre-SMA, SMA proper, the cingulate gyrus, and the left IPL (see also
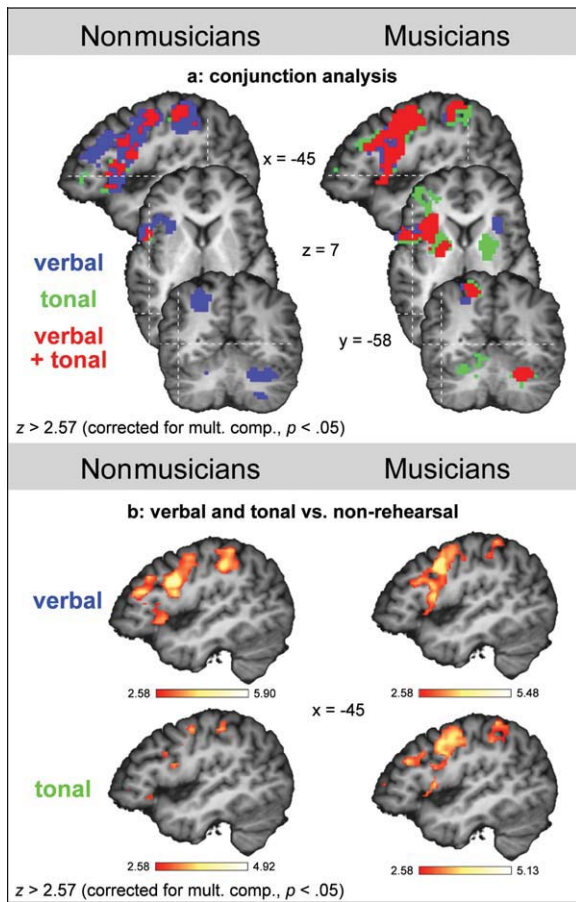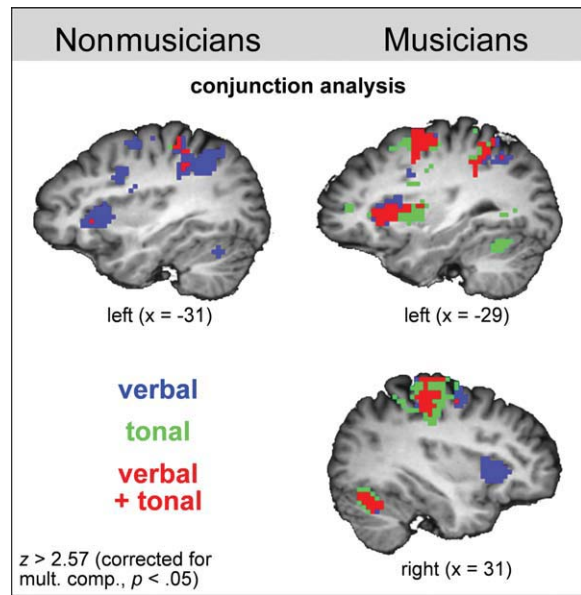
Figure 2.

fMRI results. **a:** Conjunction analysis for verbal and tonal rehearsal, separately for nonmusicians (left column) and musicians (right column). The images show areas that were significantly activated during the verbal (blue), the tonal (green), or during both verbal and tonal rehearsal (red). All activations were corrected for multiple comparisons ($P < 0.05$). **b:** Statistical parametric maps of activations during verbal (upper panel), and during tonal rehearsal (lower panel), both contrasted against the control condition (nonrehearsal), separately for nonmusicians (left column) and musicians (right column). All activations were corrected for multiple comparisons ($P < 0.05$).

Fig. 2b and Table II). In contrast to the results in nonmusicians, both verbal and tonal WM activated in musicians the left IFS, right PMd and PMv, right paracentral lobule, left IPS and left SPL, left putamen, and the right cerebellum (see also Table II and statistical evaluation of group differences in the next section).

Although significantly activated during both verbal and tonal WM, the ROI analysis showed that several structures were activated more strongly either during verbal or during tonal WM (see also Table II). During verbal compared with tonal WM, the left pars opercularis (encroaching on the IFS), the left insular cortex, and the left PMv were acti-
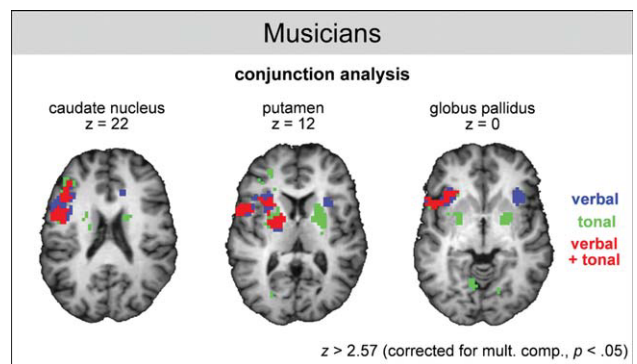

Figure 3.

fMRI results. Conjunction analysis for verbal and tonal rehearsal. The images illustrate activations in the insula in nonmusicians (left panel) and musicians (right panel) during the verbal (blue), the tonal (green), or during both verbal and tonal rehearsal (red). All activations were corrected for multiple comparisons ($P < 0.05$).

vated more strongly. During tonal compared with verbal WM, stronger activations were observed in the left IFS, right PMd and right PMv, the left cingulate gyrus, right paracentral lobule, left putamen, and right cerebellum.

Moreover, the conjunction analysis showed activations for tonal, but not verbal, WM in the left cuneus, the right

Figure 4.

fMRI results. Conjunction analysis for verbal and tonal rehearsal in musicians. The images illustrate activations in the basal ganglia during the verbal (blue), the tonal (green), or during both verbal and tonal rehearsal (red). All activations were corrected for multiple comparisons ($P < 0.05$).

globus pallidus and the right caudate nucleus, as well as the left cerebellum. The ROI analysis confirmed that the difference in activation between conditions was significant in these regions (see also Fig. 2a and Table II). By contrast, verbal, but not tonal, WM activated the right insular cortex, an observation that was also confirmed by the ROI analysis (see Fig. 2a and Table II).

Comparison Between Nonmusicians and Musicians

For all structures that were activated during tonal WM in either of the groups (as indicated by the conjunction analysis), ROI analyses were performed to test for differences in the activations of these structures between groups. In nonmusicians, only the left pars orbitalis of the IFG was activated more strongly than in musicians, whereas in musicians the left pars opercularis, left IFS, left PMv and PMd, left insular cortex, right caudate nucleus, pre-SMA and SMA, the cingulate gyrus, left IPL, left IPS, left SPL, and left cuneus were activated more strongly than in nonmusicians (see column musicians vs. nonmusicians in Tables I and II).

Finally, to confirm that structures showing a difference in activation in the ROI analysis also show a difference in the direct comparisons of the SPMs a direct comparison between musicians and nonmusicians on a voxel-by-voxel level for the data of the tonal condition was conducted. The differences found to be significant in the ROI analysis were also significant with the same significance level of $P < 0.05$ in the direct comparison (voxel-by-voxel analysis) of the SPMs.

Interestingly, many of the areas observed to be activated more strongly in nonmusicians during verbal compared to tonal WM, were activated more strongly in musicians compared to nonmusicians during tonal WM, namely Broca's area, left PMv and left PMd, left insular cortex, pre-SMA and SMA, the cingulate gyrus, and left IPL. That is, the cerebral network used by nonmusicians for verbal WM was used by musicians also for tonal WM.

An analogous ROI analysis for between-group differences was also performed for the verbal condition (for all structures that were activated either only in nonmusicians or only in musicians). Although we did not have hypotheses about group differences during verbal WM, we report these results here to generate hypotheses for future studies: Results showed no differences between groups, except that the right PMd and the left cerebellum were activated more strongly in nonmusicians than in musicians, and the right anterior insula was activated more strongly in musicians compared to nonmusicians.

DISCUSSION

In nonmusicians, both WM tasks activated areas typically reported in previous studies on either verbal [Awh et al., 1996; Baddeley, 2003; Gruber and von Cramon, 2003; Paulesu et al., 1993] or tonal WM [Gaab et al., 2003; Hickok et al., 2003; Koelsch et al. 2008; Zatorre et al., 1994] (see *phonological and tonal loop nonmusicians* in Fig. 5). The fact that these core structures of auditory WM were activated in both tasks corroborates previous results showing considerable overlap of the WM networks underlying the rehearsal of verbal and tonal information [Hickok et al., 2003; Koelsch et al., 2008]. However, it is important to note that, only in nonmusicians, all structures involved in tonal WM were also involved in verbal WM, whereas verbal, but not tonal, WM activated additional structures which have previously been implicated in verbal WM [Baddeley, 2003; Cabeza and Nyberg, 2000; Petrides et al., 1993; see also *phonological loop nonmusicians* in Fig. 5]. Thus, in nonmusicians the structures involved in tonal WM showed considerable overlap with the structures involved in verbal WM (but not vice versa).

This difference in activation of WM resources in nonmusicians corresponds to the behavioral data, which showed superior performance during verbal compared with tonal WM. The behavioral and neurophysiological differences between WM for verbal and tonal information in nonmusicians are presumably due to more extensive production and rehearsal of verbal information in everyday life, an issue that will be discussed further below.

Musicians showed an even larger overlap of WM structures for verbal and tonal WM (see *phonological and tonal loop musicians* in Fig. 5). Importantly, in contrast to nonmusicians, musicians recruited a number of structures exclusively for the rehearsal in either of the two domains: for tonal information these were the left cuneus, the right globus pallidus, the right caudate nucleus, as well as the left cerebellum, and for verbal information this was the right insular cortex [for previous studies showing that basal ganglia play a role for WM see Brown et al., 1997; Gruber et al., 2006; Koelsch et al., 2008].

Moreover, in musicians, the amount of activation differed in a number of structures between verbal and tonal WM. These results indicate the existence of two WM systems in musicians: A phonological loop dedicated to the maintenance of phonological information, and a tonal loop dedicated to the maintenance of pitch information [for a recent behavioural study reporting that verbal and tonal WM performance is more similar in nonmusicians compared with musicians see Williamson et al., 2010]. Both systems show considerable overlap in that they activate the same core structures of WM (although partly with different structural weightings), and both systems also differ in that they include different neural subcomponents. However, our data indicate that nonmusicians rely during the tonal WM on a network subserving verbal WM, and that thus both systems strongly overlap. We assume that musical training enables individuals to engage additional structures for tonal WM tasks (these structures appear to be related to sensorimotor coding, see below for further details).
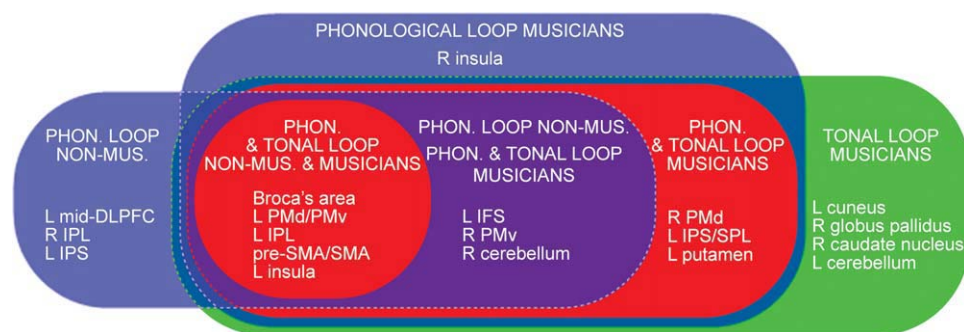
Figure 5.

Summary of main structures involved in verbal and tonal WM in nonmusicians and musicians; red areas indicate structures that are recruited for both verbal and tonal WM (referred to in the main text as core structures); blue areas indicate structures that are recruited for verbal WM in addition to the core structures; green areas indicate structures that are recruited for tonal WM in addition to the core structures; purple area indicates structures that are recruited for verbal WM in nonmusicians, as well as for both verbal and tonal WM in musicians. Phonological and tonal loop in nonmusicians: Broca's area, left (L) dorsolateral and ventrolateral premotor cortex (PMd and PMv), L inferior parietal lobe (IPL), presupplementary motor area (pre-SMA) and SMA, L insula. Phonological and tonal loop in musicians: Broca's area, L PMd/PMv, L IPL, pre-SMA/SMA, L insula, L inferior fron- tal sulcus (IFS), right (R) PMv/PMd, R cerebellum, L intraparietal sulcus (IPS), and (superior parietal lobe) SPL, L putamen. Phono- logical loop in nonmusicians: Broca's area, L PMd/PMv, L IPL, pre-SMA/SMA, L insula, L mid-dorsolateral prefrontal cortex (mid-DLPFC), R IPL, L IPS, L IFS, R PMv, R cerebellum. Phono- logical loop in musicians: Broca's area, L PMd/PMv, L IPL, pre- SMA/SMA, L insula, L IFS, R PMv, R cerebellum, R PMd, L IPS/ SPL, L putamen, R insula (the R insula was selectively recruited for verbal WM). Tonal loop in musicians: Broca's area, L PMd/ PMv, L IPL, pre-SMA/SMA, L insula, L IFS, R PMv, R cerebellum, R PMd, L IPS/SPL, L putamen, as well as four additional struc- tures that were selectively recruited for tonal WM: L cuneus, R globus pallidus, R caudate nucleus, L cerebellum.

Interestingly, the neural network used by nonmusicians for verbal WM was used by musicians for both verbal and tonal WM: First, the left IFS, the right PMv, left IPS, and the right cerebellum were involved in nonmusicians in verbal, but not tonal WM. In musicians, by contrasts, these structures were involved in both verbal and tonal WM (in musicians, the right PMv and the right cerebellum were activated even more strongly during tonal than during verbal rehearsal). Future studies could determine in how far this phenomenon might be responsible for positive transfer effects of musical training on language skills in children [e.g., Magne et al., 2006; Jentschke and Koelsch, 2009]. Second, a fronto-parietal network was activated more strongly in nonmusicians during verbal compared to tonal WM (see *phonological loop nonmusicians* in Fig. 5), but was activated more strongly in musicians compared to nonmusicians during tonal WM (see column musicians vs. nonmusicians in Tables I and II). Thus, it is likely that the better behavioral performance of musicians for tonal WM was related in part to a more extensive recruitment of resources that are used by nonmusicians for verbal WM. Note that musicians' performance during the tonal WM task was not influenced by the fact that simultaneously presented verbal stimuli were the names of tones, for two reasons: (i) none of the musicians possessed AP and (ii) musicians did not show a performance difference ($P = 0.26$) between congruent and incongruent trials (see Meth- ods for explanation of congruent and incongruent trials).

Therefore, for musicians the rehearsal of the tones was not affected by hearing sine wave tones and note names together.

It is remarkable that the better behavioral performance in nonmusicians during verbal compared to tonal WM, and the superior behavioral performance in musicians compared to nonmusicians during tonal WM, was mainly related to differences in activation of structures that are known to be involved in the control, programming and planning, as well as execution of actions (such as Broca's area, premotor cortex (PMC), pre-SMA/SMA, left insular cortex, IPS and IPL, and the cerebellum). This was corro- borated by positive correlations between (a) nonmusicians' performance in the verbal WM task and activation in the right premotor cortex, the right IPL, right SMA and the cerebellum bilaterally and (b) musicians' performance in the tonal WM task and activation in the right premotor cortex and the cerebellum bilaterally.

Action-related sensorimotor codes have previously been proposed to assist WM processes by providing resources for the representation and maintenance of information in WM [Hickok et al., 2003; Koelsch et al., 2008; Wilson, 2001]: The phonological loop is conceived of as a memory system involving internal articulatory speech actions implemented by motor-related areas such as Broca's area, premotor and insular cortex [Baddeley, 2003; Bamiou et al., 2003], the SMA [Baddeley, 2003; Gruber and von Cramon, 2003; Paulesu et al., 1993; Ravizza et al., 2004],

and the cerebellum [Chen and Desmond, 2005; Ravizza et al., 2004]. A detailed review of the functional significance of these areas is beyond the scope of this article, but it has been established that: (a) Broca's area and the premotor cortex are involved in the planning, sequencing, and control of vocal and hand actions [Petrides et al., 2005; Watkins and Paus, 2004], as well as in auditory-motor mapping [Lahav et al., 2007]; (b) the anterior insula hosts movement representations for both speech and music [Hickok and Poeppel, 2007; Indefrey and Levelt, 2004; Mutschler et al., 2007]; and (c) numerous neurons in these structures are involved in cortico-basal ganglia-thalamo-cortical and cerebellar loops that serve voluntary motor control, and contribute to the programming, initiation, and execution of movements [Hoover and Strick, 1999; Leblois et al., 2006; Middleton and Strick, 2000; Parent and Hazrati, 1995]. Notably, action representations are also coded by parietal areas which have been shown to provide representations of actions with specific sensory information, and to integrate sensory and motor information for both the planning and the sensory guidance of movements [Fogassi and Luppino, 2005].

Differences between sensorimotor codes involved in the rehearsal of verbal and tonal material, as well as differences between skills in establishing and maintaining sensorimotor codes can explain a number of the present findings. First, the finding that Broca's area, SMA, premotor and insular cortex, IPL and the IPS, as well as the cerebellum were activated more strongly in nonmusicians during verbal compared with tonal WM is likely to be due to the fact that nonmusicians have more elaborate sensorimotor representations for the rehearsal of verbal than for the rehearsal of tonal information, and it is likely that these more elaborate representations underlie the superior behavioral performance of nonmusicians during the verbal WM task.

Second, in musicians, the structures that were specifically involved in the tonal loop were also comprised of motor-related structures, namely, the right globus pallidus, the right caudate nucleus, and the left cerebellum, which were activated only during tonal WM. Moreover, the right PMC, left putamen, as well as the right cerebellum were activated more strongly during tonal than during verbal WM (and we also observed a correlation between musicians' performance and activation in these latter structures, see the regression analysis in the Supporting Information). These structures presumably provided sensorimotor resources that were specifically used for the rehearsal of tonal information, and it is likely that the differences between tonal and phonological loops in musicians are at least partly due to different sensorimotor codes used for the rehearsal of verbal information involving articulatory codes on the one hand, and tonal information involving in addition hand- and finger codes on the other.

Third, with regard to differences between musicians and nonmusicians, the right premotor cortex was activated more strongly during tonal compared to verbal WM in musicians only, and when comparing musicians and nonmusicians directly during the tonal WM task, both left PMv and PMd were activated more strongly in musicians. This suggests that musicians have more elaborate sensorimotor codes for the rehearsal of tonal information compared to nonmusicians, presumably due to long-term learning of associations between pitch information and motor actions. This assumption is supported by behavioral results [Drost et al., 2005] indicating that action representations are triggered in pianists when they perceive sounds of piano chords (so-called learned action-effect coupling). Moreover, functional neuroimaging results showed coactivation of both motor related areas and auditory areas when musicians listen to musical stimuli played on their instrument [Bangert et al., 2006; D'Ausilio et al., 2006], when musicians perform on their instrument [Bangert et al., 2006], and when musicians observe their instrument being played [Haslinger et al., 2005].

In contrast to the tonal loop, the phonological loop of musicians included the right anterior insula, and verbal compared to tonal WM activated more strongly the left insular cortex. This indicates that the left insula plays a more important role for verbal than for tonal WM, consistent with findings showing that the left insula is involved in auditory and motor functions related to speech production [Indefrey and Levelt, 2004] and speech processing [Hickok and Poeppel, 2007], and that the right insula may be part of the verbal, but not of the tonal loop in musicians. However, note that the left insula was activated more strongly during tonal WM in musicians than in nonmusicians. This indicates that, besides its role in verbal production and processing, the left insula is involved in tonal rehearsal in musicians, and that this involvement can be shaped by musical experience. This finding is consistent with a previous study showing that the left anterior insula is involved in re-activating learned movement representations during the processing of pitch information [Mutschler et al., 2007].

An inherent problem when comparing musicians and nonmusicians with musical tasks is the performance difference between groups, which potentially influences, and interacts with, the fMRI activation pattern. Wager and Smith [2003] reported that the activation pattern associated with WM performance is particularly influenced by continuous updating, order memory, manipulation of information, and selective attention (referred to as "executive demands" by the authors). An increase of such executive demand increases (bilateral) activation in frontal (e.g., including IFG and premotor cortex), and parietal cortices during verbal WM [Wager and Smith, 2003]. Because musicians showed more activations and a more bilateral activation pattern than nonmusicians during the tonal WM task (in the presence of better task performance), it is unlikely that these group differences are simply due to differences in the mentioned executive demands. As discussed above, we presume that these differences in task difficulty are at least partly related to different

performance skills, such as the ability of musicians to produce tones, intervals, and tone sequences (partly using sensorimotor coding mechanisms) that enable them to perform better in the tonal WM task compared to nonmusicians. The neural differences underlying such performance differences are reflected in the activation patterns associated with the tonal WM task e.g., stronger activation of the left premotor cortex in musicians compared to nonmusicians. With regards to differences between verbal and tonal tasks, it is important to note that verbal as well as tonal WM processes recruited several brain structures selectively in musicians (see Figs. 2, 3, and 4), and not the same network to a different degree [which would have to be expected if different degrees of executive demand were responsible for the activation differences, Wager and Smith, 2003]. Thus, it appears unlikely that simply differences in executive demand are responsible for the structural differences between the phonological and the tonal loop in musicians.

SUMMARY AND CONCLUSION

Rehearsal of both verbal and tonal information recruited core areas of WM (Broca's area, pre-SMA/SMA, left PMC, left insular cortex, and left parietal lobe), although with significantly different structural weightings, in both nonmusicians and musicians. Nonmusicians recruited a more extensive network for verbal than for tonal WM, and the networks underlying verbal and tonal WM overlapped more strongly in musicians than in nonmusicians. In addition, the data from musicians indicate that these networks also include different neural subcomponents: Verbal (but not tonal) WM involved the right insular cortex (and Broca's area, left insular cortex, and left PMv were involved more strongly during verbal compared to tonal rehearsal), whereas tonal (but not verbal) WM involved the right globus pallidus and the right caudate nucleus, as well as the left cerebellum (and the right premotor cortex, the left putamen, and the right cerebellum were involved more strongly in tonal compared to verbal rehearsal). Therefore, the data from musicians (who are trained in both the verbal and the tonal domain) indicate that auditory WM is not a unitary system, but that maintenance of verbal information relies on a phonological loop, whereas maintenance of pitch information relies on a tonal loop. Both WM systems involve overlapping core structures, but also engage different neural subcomponents. Superior behavioral performance in nonmusicians during verbal (vs. tonal) WM, and in musicians (vs. nonmusicians) for tonal WM, was mainly related to differences in activation of structures that are also involved in the control, programming and planning of actions (such as PMC, pre-SMA/SMA, Broca's area, and cerebellum). This suggests that action-related sensorimotor codes are involved in the representation and maintenance of information in WM, and that different sensorimotor codes account for differences

between verbal and tonal WM in musicians, as well as for differences of tonal WM between nonmusicians and musicians.

REFERENCES

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

Baddeley AD (2003): Working memory: Looking back and looking forward. Nat Rev Neurosci 4:829–839.

Baddeley AD, Hitch GJ (1974): Working memory. In: Bower GA, editor. Recent Advances in Learning and Motivation. New York: Academic Press. pp 47–89.

Bamiou DE, Musiek FE, Luxon LM (2003): The insula (Island of Reil) and its role in auditory processing: Literature review. Brain Res Brain Res Rev 42:143–154.

Bangert M, Peschel T, Schlaug G, Rotte M, Drescher D, Hinrichs H, Heinze HJ, Altenmuller E (2006): Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. Neuroimage 30:917–926.

Brown LL, Schneider JS, Lidsky TI (1997): Sensory and cognitive functions of the basal ganglia. Curr Opin Neurobiol 7:157–163.

Buchsbaum BR, D'Esposito M (2008): The search for the phonological store: From loop to convolution. J Cogn Neurosci 20:762–778.

Cabeza R, Nyberg L (2000): Imaging cognition. II. An empirical review of 275 PET and fMRI studies. J Cogn Neurosci 12:1–47.

Chan AS, Ho YC, Cheung MC (1998): Music training improves verbal memory. Nature 396:128.

Chen SH, Desmond JE (2005): Cerebrocerebellar networks during articulatory rehearsal and verbal working memory tasks. Neuroimage 24:332–338.

Conrad R, Hull AJ (1964): Information, acoustic confusion and memory span. Br J Psychol 55:429–432.

D'Ausilio A, Altenmuller E, Olivetti Belardinelli M, Lotze M (2006): Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. Eur J Neurosci 24:955–958.

Deutsch D (1974): Generality of interference by tonal stimuli in recognition memory for pitch. Q J Exp Psychol 26:229–234.

Drost UC, Rieger M, Brass M, Gunter TC, Prinz W (2005): Action-effect coupling in pianists. Psychol Res 69:233–241.

Fitch WT (2006): The biology and evolution of music: a comparative perspective. Cognition 100:173–215.

Fogassi L, Luppino G (2005): Motor functions of the parietal lobe. Curr Opin Neurobiol 15:626–631.

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (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 KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R (1998): Event-related fMRI: Characterizing differential responses. Neuroimage 7:30–40.

Gaab N, Gaser C, Zaehle T, Jancke L, Schlaug G (2003): Functional anatomy of pitch memory—An fMRI study with sparse temporal sampling. Neuroimage 19:1417–1426.

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

Gruber AJ, Dayan P, Gutkin BS, Solla SA (2006): Dopamine modulation in the basal ganglia locks the gate to working memory. J Comput Neurosci 20:153–166.

Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW (1999): "Sparse" temporal sampling in auditory fMRI. Hum Brain Mapp 7:213–223.

Haslinger B, Erhard P, Altenmuller E, Schroeder U, Boecker H, Ceballos-Baumann AO (2005): Transmodal sensorimotor networks during action observation in professional pianists. J Cogn Neurosci 17:282–293.

Hickok G, Poeppel D (2007): The cortical organization of speech processing. Nat Rev Neurosci 8:393–402.

Hickok G, Buchsbaum B, Humphries C, Muftuler T (2003): Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. J Cogn Neurosci 15:673–682.

Hoover JE, Strick PL (1999): The organization of cerebellar and basal ganglia outputs to primary motor cortex as revealed by retrograde transneuronal transport of herpes simplex virus type 1. J Neurosci 19:1446–1463.

Indefrey P, Levelt WJ (2004): The spatial and temporal signatures of word production components. Cognition 92:101–144.

Jacquemot C, Scott SK (2006): What is the relationship between phonological short-term memory and speech processing? Trends Cogn Sci 10:480–486.

Jentschke S, Koelsch S (2009): Musical training modulates the development of syntax processing in children. Neuroimage 47:735–744.

Keenan JP, Thangaraj V, Halpern AR, Schlaug G (2001): Absolute pitch and planum temporale. Neuroimage 14:1402–1408.

Koelsch S, Schulze K, Sammler D, Fritz T, Muller K, Gruber O (2008): Functional architecture of verbal and tonal working memory: An FMRI study. Hum Brain Mapp 30:859–873.

Lahav A, Saltzman E, Schlaug G (2007): Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. J Neurosci 27:308–314.

Leblois A, Boraud T, Meissner W, Bergman H, Hansel D (2006): Competition between feedback loops underlies normal and pathological dynamics in the basal ganglia. J Neurosci 26:3567–3583.

Lohmann G, Muller K, Bosch V, Mentzel H, Hessler S, Chen L, Zysset S, von Cramon DY (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.

Magne C, Schon D, Besson M (2006): Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. J Cogn Neurosci 18:199–211.

Middleton FA, Strick PL (2000): Basal ganglia output and cognition: Evidence from anatomical, behavioral, and clinical studies. Brain Cogn 42:183–200.

Munte TF, Altenmuller E, Jancke L (2002): The musician's brain as a model of neuroplasticity. Nat Rev Neurosci 3:473–478.

Mutschler I, Schulze-Bonhage A, Glauche V, Demandt E, Speck O, Ball T (2007): A rapid sound-action association effect in human insular cortex. PLoS ONE 2:e259.

Oechslin MS, Meyer M, Jancke L (2009): Absolute Pitch—Functional evidence of speech-relevant auditory acuity. Cereb Cortex 20:447–455.

Oldfield RC (1971): The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 9:97–113.

Parent A, Hazrati LN (1995): Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. Brain Res Brain Res Rev 20:91–127.

Paulesu E, Frith CD, Frackowiak RS (1993): The neural correlates of the verbal component of working memory. Nature 362:342–345.

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

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

Petrides M, Cadoret G, Mackey S (2005): Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. Nature 435:1235–1238.

Ravizza SM, Delgado MR, Chein JM, Becker JT, Fiez JA (2004): Functional dissociations within the inferior parietal cortex in verbal working memory. Neuroimage 22:562–573.

Salame P, Baddeley AD (1989): Effects of background music on phonological short-term memory. Q J Exp Psychol 41:107–122.

Semal C, Demany L, Ueda K, Halle PA (1996): Speech versus nonspeech in pitch memory. J Acoust Soc Am 100:1132–1140.

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

Wager TD, Smith EE (2003): Neuroimaging studies of working memory: a meta-analysis. Cogn Affect Behav Neurosci 3:255–274.

Wallin NL, Merker B, Brown S (1999): The Origins of Music. Cambridge: MIT Press.

Warren JE, Wise RJ, Warren JD (2005): Sounds do-able: Auditory-motor transformations and the posterior temporal plane. Trends Neurosci 28:636–643.

Watkins K, Paus T (2004): Modulation of motor excitability during speech perception: The role of Broca's area. J Cogn Neurosci 16:978–987.

Williamson VJ, Baddeley AD, Hitch GJ (2010): Musicians' and nonmusicians' short-term memory for verbal and musical sequences: Comparing phonological similarity and pitch proximity. Mem Cognit 38:163–175.

Wilson M (2001): The case for sensorimotor coding in working memory. Psychon Bull Rev 8:44–57.

Zatorre RJ, Evans AC, Meyer E (1994): Neural mechanisms underlying melodic perception and memory for pitch. J Neurosci 14:1908–1919.

Zatorre RJ, Chen JL, Penhune VB (2007): When the brain plays music: Auditory-motor interactions in music perception and production. Nat Rev Neurosci 8:547–558.