

The song system of the human brain

By: Steven Brown, Michael J. Martinez, Donald A. Hodges, Peter T. Fox, and Lawrence M. Parsons

Brown, S., Martinez, M., [Hodges, D.](#), & Fox, P., & Parsons, L. (2004) The song system of the human brain. *Cognitive Brain Research*, 20: 363-375. doi:10.1016/j.cogbrainres.2004.03.016

Made available courtesy of ELSEVIER: <http://www.elsevier.com/locate/brainres>

*****Note: Figures may be missing from this format of the document**

Abstract:

Although sophisticated insights have been gained into the neurobiology of singing in songbirds, little comparable knowledge exists for humans, the most complex singers in nature. Human song complexity is evidenced by the capacity to generate both richly structured melodies and coordinated multi-part harmonizations. The present study aimed to elucidate this multi-faceted vocal system by using 15O-water positron emission tomography to scan “listen and respond” performances of amateur musicians either singing repetitions of novel melodies, singing harmonizations with novel melodies, or vocalizing monotonically. Overall, major blood flow increases were seen in the primary and secondary auditory cortices, primary motor cortex, frontal operculum, supplementary motor area, insula, posterior cerebellum, and basal ganglia. Melody repetition and harmonization produced highly similar patterns of activation. However, whereas all three tasks activated secondary auditory cortex (posterior Brodmann Area 22), only melody repetition and harmonization activated the planum polare (BA 38). This result implies that BA 38 is responsible for an even higher level of musical processing than BA 22. Finally, all three of these “listen and respond” tasks activated the frontal operculum (Broca's area), a region involved in cognitive/motor sequence production and imitation, thereby implicating it in musical imitation and vocal learning.

Author Keywords: Singing; Song system; Brain; Music; Melody; Harmony, Motor Systems and Sensorimotor Integration, Cortex

Article:

Singing is a specialized class of vocal behavior found in a limited number of animal taxa, including humans, gibbons, humpback whales, and about half of the nine thousand species of bird. Various functions have been attributed to singing, including territorial defense, mate attraction, pair bonding, coalition signaling, and group cohesion [5, 25, 46 and 76]. Song production is mediated by a specialized system of brain areas and neural pathways known as *the song system*. This system is also responsible for song learning, as most singing species acquire their songs via social learning during development [30 and 31]. In some species, known as “age-limited learners”, song learning occurs once during a critical period; in “open-ended learners”, song learning occurs throughout much of the life span (e.g., [70]). In many species of bird, singing is a sexually dimorphic behavior, one that is performed mainly by males [12]. In these species, the vocal centers of males tend to be three to five times larger than those of females [41]. However, in species where both sexes sing, the vocal centers of the two sexes tend to be of comparable size [14]. Importantly, the components of the forebrain song system are absent in even taxonomically close bird species that either do not sing or that acquire their songs in the absence of vocal learning [37]. This highlights the notion that song learning through vocal imitation is an evolutionary novelty, one that depends on the emergence of new neural control centers.

Although humans are by far the most complex singers in nature, the neurobiology of human song is much less well understood. A deeper understanding of singing may benefit from a comparative approach, as human singers show features that are both shared with, and distinct from, birds and other singers in nature [22]. Common features include the following: (1) both absolute and relative pitch processing are used for song [42]; (2) combinatorial pitch codes are used for melody generation [44]; (3) there is a capacity for phonatory improvisation and invention [38]; (4) the song is treated as the fundamental unit of communication [68]; (5)

songs are organized into repertoires [69]; (6) imitative vocal learning is important for song acquisition [1]; (7) there is year-round rather than seasonal singing [7]; and (8) there is a capacity for acquisition of songs throughout the life span [16]. Along these lines, although there is no systematic evidence for a critical period in human song learning, it is conceivable that the common incidence of “poor pitch singing” (often mislabeled as “tone deafness”) reflects the possibility that vocal behavior (or its absence) during childhood has a strong effect on adult singing abilities.

At the same time, human music has several features distinct from singing in other animals, most notably choral singing and harmony. The temporal synchronization of voices that underlies human choral singing bears little relation to the “dawn chorus” of birds, in which vocal blending is little more than random simultaneity. While there is clear evidence for synchronization of parts in the songs of duetting species, such as gibbons and many tropical birds, none shows the kind of vertical alignment of parts that is the defining feature of harmonic singing in humans. Vertically integrated, multi-part singing is absent in non-human species, thereby suggesting that the human song system is different from that of other species, one specialized for coordinated multi-person blending. Harmonic singing is a characteristic musical style of several distinct regions of the world. Such singing is generally a cooperative behavior, often serving to reinforce collective norms and group actions. Our closest genetic relatives, chimpanzees and bonobos, do not engage in any kind of vocalizations reminiscent of song. Singing, therefore, cannot be seen as an ancestral trait of hominoid species but instead a derived feature of humans.

Such considerations are consistent with the hypothesis that the human song system is an evolutionary novelty and neural specialization, analogous to the song system of birds. However, this hypothesis is difficult to evaluate at the present time as human singing has been little researched. While music and song were the subjects of intense speculation by Enlightenment thinkers (e.g., [10 and 63]), modern neurobiology provides limited pertinent information. There are few studies of vocal amusias but instead various reports of Broca's aphasics whose singing ability, even for lyrics, is spared (e.g., [28, 77 and 79]). Such findings are probably more common than reports of spared musical function in the face of language deficits, as a knowledge of baseline musical-production skills is absent in most “non-musicians” and because neurologists do not generally examine musical capacities in patients who are not musicians. Most noninvasive functional brain imaging studies of music have focused on perceptual rather than productive aspects.

Building on the foregoing achievements and considerations, we designed the current PET study to elucidate the audiovocal system underlying basic musical production processes and to compare the functional neuroanatomy of melody production with that for harmony production. The study was designed to examine these issues more comprehensively than did the two previous studies of song production. Perry et al. [53] looked only at monotone singing, and Riecker et al. [60] looked only at the singing of a single highly familiar melody. In the present investigation, we were interested in examining the vocal processing of *novel* melodies, as they would serve as more richly engaging stimuli with which to probe the audiovocal system. Amateur musicians performed four tasks while being scanned in this study: (1) *Melody Repetition*: subjects sang repetitions of novel, one-line, rhythmically varied melodies; (2) *Harmonization*: subjects sang harmonizations in coordination with novel, chordal, rhythmically varied melodies; (3) *Monotonic Vocalization*: the two preceding conditions were contrasted with a lower-level task in which subjects sang isochronous monotone sequences in alternation with isochronous sequences of the same piano pitch; and (4) *Rest*: eyes-closed rest was used as a silent, non-motor baseline condition. A distinct feature of this design compared to the previous studies was an element of imitative vocalizing. The Melody Repetition condition involved tandem repetition of heard melodies, the Monotonic Vocalization condition involved a matching of the pitch and rhythm of a monotone sequence, and the Harmonization condition—while not requiring direct imitation of the presented melodic sequence—required a shadowing of that sequence at a displaced location in tonal pitch space (e.g., a major third above the original melodic line). For terminological purposes, we are using the words “repetition” and “imitation” more or less interchangeably, with “repetition” being used more in the context of our tasks and “imitation” more in the context of general cognitive processing.

We hypothesized that secondary and tertiary auditory areas would be increasingly recruited as the complexity of the pitch, rhythmic and musical aspects of the production task increased from basic monotonic vocalizing to melodic and harmonic singing. We also hypothesized that the Repetition and Harmonization tasks would engage brain areas involved in working memory, compared to the Monotone task. Finally, we hypothesized that regions thought to underlie higher-level motor planning for vocalization—such as the supplementary motor area, Broca's area, and the anterior insula [15, 19, 33 and 86]—would be involved not only in the motor control of song production but in musical imitation as well.

1. Materials and methods

1.1. Subjects

Five male and five female neurologically healthy amateur musicians, with a mean age of 25 years (range 19–46 years), participated in the study after giving their informed consent (Institutional Review Board of the University of Texas Health Science Center). Each individual was right-handed, as confirmed by the Edinburgh Handedness Inventory [49]. All subjects were university students, many in their first or second years as music education majors, with a mean of 5.3 years of formal music instruction in voice or instrument. Subjects began music instruction at a mean age of 12.4 years old, having had an involvement in musical production (e.g., school bands, church choirs) for an average of 12.6 years prior to the study. None of them had absolute pitch, as based on self-report. Their musical specializations included voice, flute, trumpet, trombone, piano, drums, bass, guitar, percussion, and clarinet. Subjects underwent a detailed behavioral screening procedure in order to determine their suitability for the study. Each potential subject was presented with 35 melody repetition samples and 26 harmonization samples. Criteria for inclusion in the study included the following: (1) a proficiency at singing in key, (2) an ability to sing at least 50% of the repetition samples with perfect accuracy, and (3) an ability to sing at least 50% of the harmonization samples in such a manner that the melodic contour of the original melody was shadowed perfectly, in accordance with the rules of tonal harmony (see Tasks below). The 10 subjects who were used in this study were taken from a pool of 36 amateur musicians who underwent the screening procedure.

1.2. Tasks

Stimuli for the vocal tasks were sequences of digitized piano tones generated using Finale 2001 (Coda Music Technology). Subjects performed three vocal tasks and eyes-closed rest (see Fig. 1). The carrier syllable /da/ was used for all the singing tasks; this was done to avoid humming, to control head and mouth movement, and to permit adequate respiration during performance of the tasks. (1) *Monotonic Vocalization*. Subjects heard a piano tone (147 Hz; D below middle C), played 4 to 11 times isometrically (in an equal-interval, regular rhythm). The notes were played at a rate of 100 beats per minute, or 1.67 Hz, with a note duration of 600 ms. Subjects had to sing back the same pitch at the same tempo and rate (i.e., isochronously) whenever the piano stopped playing the note, doing so in continuous alternation with the piano. As with each sequence of piano tones, the response period allowed time for the singing of 4–11 tones. Each successive sequence was different in the number of tones from the prior one. The goal of this arrangement was to ensure that subjects, in attempting to match pitch and rhythm, were not cognitively engaged in counting piano tones; subjects did not need to count piano tones because their singing was interrupted when the piano tones of the succeeding trial began. Hence their goal was simply to match the pitch and rhythm of these tones. (2) *Melody Repetition*. Subjects listened to a series of tonal melodies, and had to sing back each one after it was played. Each melody was 6 s in duration, followed by a 6-s period for response generation. The inter-trial interval was 1 s. Consecutive samples were never in the same key. (3) *Harmonization*. Subjects listened to a series of melodies accompanied by chords and had to spontaneously sing a harmonization with each melody as it was being replayed. Each melody was 6 s in duration. A “prompt tone” was provided after the first presentation of each melody, which subjects were instructed to use as the first note of their harmonization. This tone was typically a major third above the first note of the melody, which itself was frequently the tonic pitch of the scale. When

melodies started on the third degree of the scale, the prompt tone was a perfect fifth above the tonic. The loudness of the stimulus heard during harmonization was reduced by 67% so that subjects could hear their singing. The inter-trial interval was 1 s. Consecutive samples were never in the same key. Subjects were instructed to create harmonizations that conformed to the rules of tonal harmony. While they generally sang the harmonizations in thirds, there were points in the melody where the rules of harmony dictated the use of other intervals, such as fourths, as a function of the implicit harmonic structure of the melody at that point.



Fig. 1. Representative stimuli for the three singing tasks performed in this study: Monotonic Vocalization, Melody Repetition, and Harmonization. The note with the asterisk over it in Harmonization is the “prompt tone” that was provided to subjects as the first note of their harmonization (see Materials and methods).

1.3. Stimuli

All stimuli for the vocal tasks were presented to both ears as piano tones, and were generated using Finale 2001. The source material consisted of folk-music samples from around the world, modified to fit the time and musical constraints of the stimulus set. Pilot testing ($n=7$) confirmed that all stimulus material was novel for our subject population. A hypothetical standard for the stimulus set consisted of a sample with 10 quarter-notes at a tempo of 100 beats per minute in 4/4 time. The stimuli for the Melody Repetition and Harmonization conditions were varied with regard to tempo (slower and faster than the standard), number of notes (fewer or more notes than the standard), tonality (major and minor), rhythm (duple [2/4, 6/8], triple and quadruple time), motivic pattern (e.g., dotted vs. non-dotted rhythms), and melodic contour (ascending and descending patterns). The samples covered a wide range of keys. Volume was approximately constant among the stimuli. The Monotonic Vocalization task consisted of a single tone (147 Hz) in a comfortable vocal range of males and females, although subjects were given the option of singing the tone one octave higher. This task was designed to control for the average number of notes that a subject would both hear and produce in the other two singing conditions.

1.4. Procedure

During the PET session, subjects lay supine in the scanning instrument, with the head immobilized by a closely fitted thermal-plastic facial mask with openings for the eyes, ears, nose, and mouth. Auditory stimuli were presented through the earpieces of headphones taped over the subjects' ears. During scanning, subjects were told to close their eyes, lie motionless, and to clench their teeth lightly so as to make the syllable /da/ when singing. Pre-scan training enabled the subjects to perform the vocalization tasks with minimal head movement. Each subject had two PET scans for each of the vocal tasks and one of rest. Task order was counterbalanced pseudo-randomly across subjects. The subjects began each task 30 s prior to injection of the bolus. Bolus uptake

required approximately 20 s to reach the brain, at which time a 40-s scan was triggered by a sufficient rate of coincidence-counts, as measured by the PET camera. At the end of the 40-s scan, the auditory stimulus was terminated and the subject was asked to lie quietly without moving during a second scan (50 s). From the initiation of the task until the start of the second scan, each subject had responded to six to seven stimuli.

1.5. Imaging

PET scans were performed on a GE 4096 camera, with a pixel spacing of 2.0 mm, and inter-plane, center-to-center distance of 6.5 mm, 15 scan planes, and a z -axis field of view of 10 cm. Images were reconstructed using a Hann filter, resulting in images with a spatial resolution of approximately 7 mm (full-width at half-maximum). The data were smoothed with an isotropic 10-mm Gaussian kernel to yield a final image resolution of approximately 12 mm. Anatomical MRI scans were acquired on an Elscint 1.9 T Prestige system with an in-plane resolution of 1 mm^2 and 1.5-mm slice thickness.

Imaging procedures and data analysis were performed exactly as described by Parsons and Osherson [52], according to the methods of Raichle et al. [57], Fox et al. [18] and Mintun et al. [47]. Briefly, local extrema were identified within each image with a 3-D search algorithm [47] using a 125 voxel search cube (2 mm^3 voxel). A beta-2 statistic measuring kurtosis and a beta-1 statistic measuring skewness of the extrema histogram [17] were used as omnibus tests to assess overall significance [11]. Critical values for beta statistics were chosen at $p < 0.01$. If the null hypothesis of omnibus significance was rejected, then a post hoc (regional) test was done [17 and 18]. In this algorithm, the pooled variance of all brain voxels is used as the reference for computing significance. This method is distinct from methods that compute the variance at each voxel but is more sensitive [71], particularly for small samples, than the voxel-wise variance methods of Friston et al. [20] and others. The critical-value threshold for regional effects ($z > 2.58$, $p < 0.005$, one-tailed) is not raised to correct for multiple comparisons since omnibus statistics is established before post hoc analysis.

1.6. Task performance

As noted above, we selected subjects who were able to perform the tasks with competence. Analysis of recorded task performance confirmed that subjects performed in the scanner in a manner qualitatively identical to their performance during the screening session. Our use of a stringent screening procedure for subject inclusion meant that our subject sample was rather homogeneous, producing minimally variable task performance across individuals. Therefore, by design, we were not in a position to employ covariance analysis to look at the relationship between brain activation and task performance.

2. Results

The mean cerebral blood flow increases for the Monotonic Vocalization task, as contrasted with Rest (Fig. 2, Table 1), showed bilateral activations in the primary auditory cortex (Brodmann Area [BA] 41) and the mouth region of the primary motor cortex (BA 4). Bilateral activations were observed in the auditory association cortex (BA 42 and posterior BA 22), frontal operculum (inferior parts of BA 44, 45 and 6), and supplementary motor area (SMA; medial BA 6), with trends towards greater right hemisphere activations; it is important to note that for the frontal operculum, the left hemisphere activation was reproducibly more posterior than that in the right hemisphere, extending into BA 6. The anterior cingulate cortex (BA 24) was also seen to be activated in this task. Other notable activations occurred in left anterior putamen, right globus pallidus, and posterior cerebellar hemispheres. The activations in basal ganglia—putamen on the left and globus pallidus on the right—most likely supported processes in the ipsilateral cerebral hemispheres [13]. Broadly speaking, then, this task produced bilateral activations in primary auditory and vocal areas and more right-lateralized activations in higher-level cortical areas.

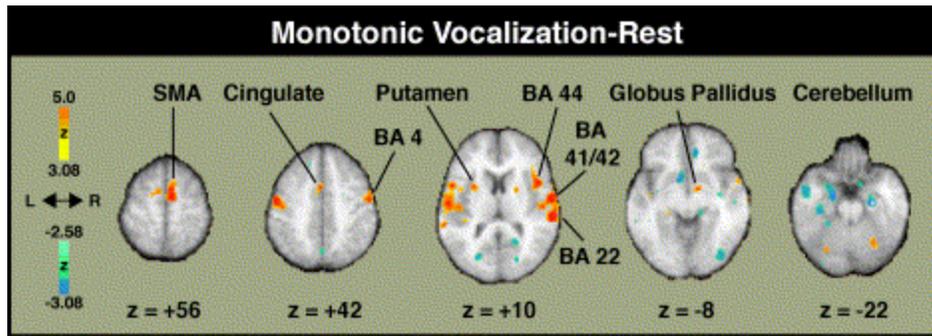


Fig. 2. Axial views of cerebral blood flow changes during Monotonic Vocalization contrasted to Rest. The Talairach coordinates of the major activations (contrasted to Rest) are presented in Table 1. The averaged activations for 10 subjects are shown registered onto an averaged brain in all the figures. The right side of the figure is the right side of the brain in all the figures. At the left end of the figure are two color codes. The upper one (yellow to red) is a scale for the intensity of the *activations* (i.e., blood flow increases), whereas the lower one (green to blue) is a scale for the intensity of the *deactivations* (i.e., blood flow decreases). The group mean blood-flow decreases showed no obvious pattern related to the tasks or to the blood-flow increases and are thus not reported in the text. Note that the same set of five slice-levels is shown in Fig. 2, Fig. 3 and Fig. 4. Note also that bilateral activations are labeled on only one side of the brain. The label SMA stands for supplementary motor area. The intensity threshold in Fig. 2, Fig. 3 and Fig. 4 for all tasks is $z > 2.58$, $p < 0.005$ (one-tailed).

Table 1. Stereotaxic coordinates and z -score values for activations in the Monotonic Vocalization task contrasted with Rest

Region		x	y	z	z-score
<i>Frontal</i>					
Right	Supplementary Motor Area (6)	6	0	60	5.58
	Supplementary Motor Area (6)	8	-8	56	5.54
	Motor Cortex (4)	46	-10	36	4.88
	Frontal Operculum (44)	44	8	10	4.71
	Motor Cortex (4)	58	-8	24	4.17
Left	<i>Motor Cortex</i> (4)	-46	-16	36	5.92
	<i>Motor Cortex</i> (4)	-48	-10	44	5.25
	<i>Premotor Cortex</i> (6)	-48	6	6	5.00
	<i>Supplementary Motor Area</i> (6)	-6	-4	62	4.88
	<i>Motor Cortex</i> (4)	-54	-10	24	4.50
	<i>Prefrontal Cortex</i> (46/9)	-32	36	28	3.88
	<i>Frontal Operculum</i> (45)	-42	24	4	3.80
<i>Temporal</i>					
Right	Superior Temporal Gyrus (22)	58	-28	6	6.04
	Secondary Auditory Cortex (42)	58	-8	8	5.75
	Primary Auditory Cortex (41)	48	-18	10	4.25
	Planum Polare (22/38)	58	4	-4	4.17
	Middle Temporal Gyrus (21)	66	-16	-4	4.01
	Fusiform Gyrus (37)	46	-46	-18	3.84
Left	Primary Auditory Cortex (41)	-34	-34	14	4.42
	Secondary Auditory Cortex (42)	-54	-14	10	4.42
	Superior Temporal Gyrus (22)	-58	-38	10	4.17
	Primary Auditory Cortex (41)	-42	-18	10	4.13
	Superior Temporal Sulcus (21)	-42	-12	-6	3.96
<i>Other</i>					
Right	Globus Pallidus	12	-4	-6	4.38
	Putamen	20	0	8	4.01
	Putamen	27	-20	-4	3.84
Left	Insula	-38	-2	4	4.46
	Cingulate Gyrus (24)	-2	2	44	4.42
	Putamen	-22	6	4	4.38
<i>Cerebellum (Posterior Lobe)</i>					
Right	Quadrangular Lobule (VI)	26	-58	-24	4.05
Left	Quadrangular Lobule (VI)	-26	-68	-22	3.88
	Uvula (Vermis IX)	-4	-56	-30	3.84
	Inf. Semilunar Lobule (Crus II)	-12	-72	-34	3.71
	Quadrangular Lobule (VI)	-36	-52	-28	3.71

Brain atlas coordinates are in millimeters along the left-right (x), anterior–posterior (y), and superior–inferior (z) axes. In parentheses after each brain region is the Brodmann area, except in the case of the cerebellum, in which the anatomical labels of Schmahmann et al. [67] are used. The intensity threshold is $z > 3.72$, $p < 0.0001$ (one-tailed).

Melody Repetition minus Rest (Fig. 3a, Table 2), compared to the results with Monotonic Vocalization, showed no cingulate activation, much less activation in the primary auditory cortex, and activation in the superior part of the temporal pole (planum polare, BA 38). In general, the pattern of activation for Melody Repetition closely overlapped that for Monotonic Vocalization. Thus, when Monotonic Vocalization was subtracted from Melody Repetition (Fig. 3b), there was little signal above threshold in most auditory and motor areas. Only the activation in the planum polare (BA 38) remained after this subtraction, implicating this area in higher-level musical processing.

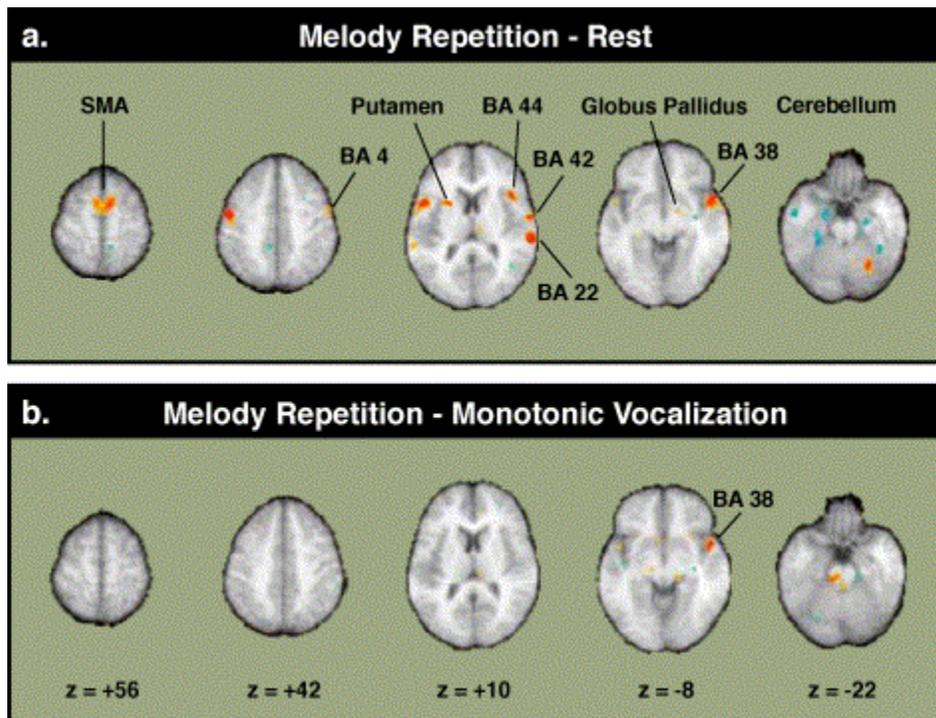


Fig. 3. Axial views of cerebral blood flow changes during Melody Repetition contrasted with (a) Rest and (b) Monotonic Vocalization. The Talairach coordinates of the major activations (contrasted to Rest) are presented in Table 2. Subtraction of Monotonic Vocalization from Melody Repetition eliminates many of the significant activations but leaves the signal in the planum polare (BA 38) at $z = -8$. The peak voxel for BA 38 in the Melody Repetition minus Monotonic Vocalization subtraction (panel b) was located at (48, 6, -6) in the right hemisphere and (-42, 4, -7) in the left.

Table 2. Stereotaxic coordinates and z -score values for activations in the Melody Repetition task contrasted with Rest

	Region	x	y	z	z-score
<i>Frontal</i>					
Right	Supplementary Motor Area (6)	6	2	62	5.10
	Frontal Operculum (44)	40	14	10	4.47
	Premotor Cortex (6)	54	-6	40	4.17
Left	Motor Cortex (4)	-48	-8	40	5.80
	Frontal Operculum (44/6)	-48	6	10	4.54
	Supplementary Motor Area (6)	-6	0	52	4.17
<i>Temporal</i>					
Right	Planum Polare (22/38)	50	8	-4	5.87
	Superior Temporal Gyrus (22)	60	-26	4	5.73
	Superior Temporal Gyrus (22)	56	-6	4	5.32
	Middle Temporal Gyrus (21)	62	0	-11	3.80
Left	Planum Polare (22/38)	-52	0	-2	5.43
	Superior Temporal Gyrus (22)	-60	-24	6	4.10
	Primary Auditory Cortex (41)	-34	-34	16	4.06
	Superior Temporal Gyrus (22)	-58	-34	6	3.91
<i>Other</i>					
Right	Globus Pallidus	16	-4	-6	3.80
Left	Putamen	-22	6	6	4.99
	Insula	-26	6	14	4.39
	Insula	-42	12	2	3.99
<i>Cerebellum (Posterior Lobe)</i>					
Right	Quadrangular Lobule (VI)	26	-58	-24	4.47

Legend as in Table 1.

Harmonization minus Rest, as compared to the results for Melody Repetition, showed more intense activations in the same song-related areas (Fig. 4a, Table 3). In addition, there appeared to be a nonsignificant trend toward greater bilaterality of the temporal lobe activations (including BA 38) for the Harmonization task compared to the Melody Repetition task. However, when the Melody Repetition task was subtracted from the Harmonization task, no activations remained above threshold (data not shown). This can be explained in part by the results of the contrast with Monotonic Vocalization (Fig. 4b). Interestingly, even the activation in the planum polare (BA 38) was eliminated in this subtraction (not shown). In sum, harmony generation and melody generation produced closely overlapping patterns of activation. Interestingly, we predicted that dorsolateral prefrontal cortex (BA 46 and 9) would be activated in the Repetition and Harmonization tasks due to the need for subjects to keep the melodic template of the stimulus in working memory. However, such activations, while present, were below the z threshold used in our tables.

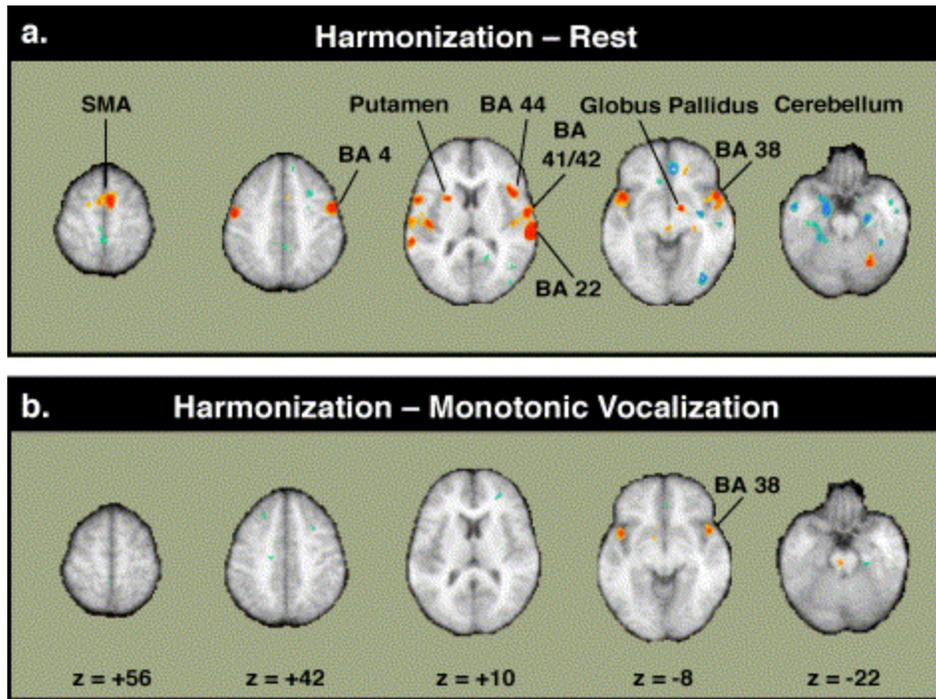


Fig. 4. Axial views of cerebral blood flow changes during Harmonization contrasted with (a) Rest, and (b) Monotonic Vocalization. The Talairach coordinates of the major activations (contrasted to Rest) are presented in Table 3. The peak voxel for BA 38 in the Harmonization minus Monotonic Vocalization subtraction (panel b) was located at (46, 8, -6) in the right hemisphere and (-42, 6, -10) in the left.

Table 3. Stereotaxic coordinates and z-score values for activations in the Harmonization task contrasted with Rest

	Region	x	y	z	z-score
<i>Frontal</i>					
Right	Motor Cortex (4/6)	52	-2	40	5.83
	Supplementary Motor Area (6)	8	-2	58	5.49
	Frontal Operculum (44/45)	40	16	8	5.10
	Frontal Operculum (44)	44	10	10	5.01
Left	Motor Cortex (4)	-50	-8	40	5.83
	Frontal Operculum (45)	-42	24	6	3.76
<i>Temporal</i>					
Right	Superior Temporal Gyrus (22)	60	-28	6	7.69
	Planum Polare (22/38)	48	6	-2	6.39
	Planum Polare (22/38)	54	0	-2	5.96
	Superior Temporal Sulcus (21)	50	-8	-2	5.87
	Secondary Auditory Cortex (42)	58	-8	8	5.57
Left	Planum Polare (22/38)	-50	4	-2	5.62
	Superior Temporal Gyrus (22)	-60	-20	4	5.49
	Superior Temporal Gyrus (21/22)	-54	-12	2	5.23
	Superior Temporal Gyrus (22)	-60	-38	8	4.84
	Planum Polare (38)	-41	4	-14	4.84
	Primary Auditory Cortex (41)	-40	-20	8	4.71
<i>Other</i>					
Right	Globus Pallidus	14	-4	-6	5.05
	Calcarine Cortex (17)	16	-68	18	4.11
	Midbrain	2	-24	-10	3.89
	Putamen	26	8	0	3.85
	Parahippocampal Gyrus (27)	28	-26	-6	3.85
	Cingulate Gyrus (32)	18	32	-8	3.76
Left	Insula	-42	10	0	5.27
	Putamen	-30	22	2	4.19
<i>Cerebellum (Posterior Lobe)</i>					
Right	Quadrangular Lobule (VI)	24	-64	-20	4.80
Left	Uvula (Vermis IX)	-4	-56	-28	4.23
	Quadrangular Lobule (VI)	-36	-50	-32	3.89

Legend as in Table 1.

3. Discussion

3.1. The human song system

These data provide a picture of the auditory and vocal components of the human song system as well as those neural areas involved in imitation, repetition, and the pitch-tracking processes underlying harmonization. The cortical activations observed here can be grouped hierarchically in terms of primary auditory and vocal areas, secondary auditory and vocal areas, and higher-level cognitive areas. All three vocal tasks showed strong activations in the primary auditory cortex (BA 41) and in the mouth region of the primary motor cortex (BA 4) [19]. Furthermore, all three vocal tasks showed activations in the auditory association cortex (BA 42 and BA 22), supplementary motor area (BA 6), frontal operculum (BA 44/6), and left insula. An activation in the anterior cingulate cortex (BA 24) was seen exclusively in the monotonic vocalization task. Finally, the two high-level music tasks, but not monotonic vocalization, showed activations in the planum polare (BA 38), implicating this area in higher-level musical processing. Interestingly, although the stimuli for the melody repetition and harmonization tasks changed key from sample to sample, we did not observe activations in the ventromedial prefrontal region identified as being important for tracking key changes [29].

Although we observed only a single occipital activation in this study—in calcarine cortex for the harmonization task—several studies of music perception and musical imagery have shown cortical activations in parietal and occipital areas [e.g., [26, 29, 40, 66 and 84]. In addition to cortical activations, we observed several activations in non-cortical areas. The left-lateralized putamen activations in all three of our vocalization tasks are consistent with findings on vocalization processes in animals and humans [33, 34 and 78]. The right globus pallidus was likewise activated in all three tasks. Further research is required to determine the exact function of this area for these tasks. Activation was detected in midbrain but only for harmonization (minus rest). At the resolution of PET used here, this activity may originate in substantia nigra or nucleus ambiguus, structures involved in the motor control of vocalization. Finally, the posterior cerebellum, especially the quadrangular lobule (VI), was active in all three tasks, as discussed below.

Overall, our results are in broad agreement with the two other studies of song production. In the PET study of Perry et al. [53], non-musicians sang simple monotone sequences using the vowel /ä/ at a target rate of 0.8 Hz, based on a presented target pitch. The activation profile seen by Perry et al. was quite similar to that observed here, with major activations occurring in the primary and secondary auditory cortices, primary motor cortex, supplementary motor area, anterior cingulate cortex, insula, and frontal operculum. In an fMRI study by Riecker et al. [60], non-musicians either overtly or covertly sang a familiar melody without words. As in both the present study and that of Perry et al., major activations occurred in the primary motor cortex, supplementary motor area, anterior insula, and posterior cerebellum. Each of the latter areas has been implicated in vocalization. The primary motor cortex is, of course, a critical mediator of voluntary vocalization. Our major focus of activation for the primary motor cortex was in the mouth area [19]. While it is possible that there were activations as well in the larynx area, we were not able to distinguish them from the activations in the frontal operculum at the spatial resolution of this study. Interestingly, nonhuman primates lack a direct connection between the larynx representation of the primary motor cortex and the nucleus ambiguus, the major peripheral neural center for vocalization [32], and as a result, no primate except the human is capable of phonatory vocal learning, such as that underlying the acquisition of song. Moreover, there is firm evidence that the supplementary motor area (SMA) plays a key role in higher-level motor control, and it is often activated during overt speech tasks in imaging experiments [75]. Direct electrical stimulation of SMA produces vocalization in humans but not other mammals [33], and damage to SMA (as with many other structures) is associated with mutism [86]. The anterior insula has long been associated with vocalization processes, and damage to this structure has been linked to disorders of articulation [15]. Its role in vocalization has been confirmed by imaging studies of counting, nursery-rhyme recitation, and propositional speech [9]. Finally, the posterior cerebellum has been implicated in vocalization processes, particularly the quadrangular lobule (VI) [75], observed both in the present study and that of Perry et al. to be activated during singing. The exact contribution of the cerebellum to song is unclear because activations in this structure could be involved in motor, auditory or somatosensory processing [51].

Activations in the primary and secondary auditory regions were seen for all three singing tasks in this study, as with Perry et al.'s monotone task. Activation in the primary auditory cortex was less in the repetition task than either the monotonic vocalization or harmonization task, for reasons that are not currently clear to us. The activations in the superior temporal gyrus (BA 22) were strongly right-lateralized for all three tasks. Activations in this region could have been due to at least two major sources: the presented stimuli and the subject's own voice. The superior temporal gyrus has been implicated in melody processing, most especially the right hemisphere [26, 81, 84 and 85] (see also Zatorre et al. [87] for a discussion of right-hemisphere dominance of the primary auditory cortex for spectral processing). Indeed, the peak activations observed here at (60, -28, 6) for harmonization and at (60, -26, 4) for melody repetition correspond to that at (64, -26, 5) when musically experienced listeners tracked a melody as it changed keys [29]. However, the elimination of BA 22 in the subtractions of monotonic vocalization from both melody repetition and harmonization suggests that BA 22 sits at a lower position in the auditory-processing hierarchy than BA 38, which was not eliminated in the same contrasts. This suggests that BA 38 might, in fact, be a form of tertiary auditory cortex. Additional studies will be needed to determine the relative contributions of the posterior and anterior regions of the superior temporal gyrus to musical processing.

In general, activations in primary and secondary auditory areas (BAs 41, 42, 22, 21) in the left hemisphere were more posterior than those in the right hemisphere. Over the group of three tasks, the mean *y* location of activations in these areas was -25 on the left and -15 on the right. A similar effect was observed in an fMRI study of non-musicians passively listening to melodies presented in different timbres [45]. In that study, the mean *y* location of activations was -24 on the left and -8 on the right. The 10-mm difference observed across our tasks is in accord with the morphological difference of 8–11 mm in the location of left and right auditory cortex [56]. However, this difference was much more pronounced for melodic repetition than for the other two tasks. Specifically, the average *y* value for these areas in the melodic repetition task was -31 on the left and -11 on the right; however, in the monotonic vocalization task, it was -23 on the left and -17 on the right, and in the harmonization task, it was -22 on the left and -15 on the right. Further research is necessary to clarify whether there is in fact such a functional asymmetry in auditory areas for music-related tasks.

Another source of auditory stimulation in this study was the subject's own vocalization. Voice-selective cortical areas have been demonstrated along the extent of the superior temporal sulcus (between BA 21 and 22), with a dominance for the right hemisphere [3 and 4]. Such work represents an important perceptual counterpart to our work on song production, particularly since any evolutionary account of the song system must take into account parallel communicative adaptations for perception and production. Although a distinction has been observed between speech and non-speech vocal sounds in these voice-selective areas [3], it will be important to examine whether there is specificity for singing versus other non-speech phonatory sounds in these regions. Such an investigation would be a fruitful counterpart to similar work with songbirds.

All three singing tasks also showed strong activations in the frontal operculum. This region, along with the more dorsal part of Broca's area proper, has been observed to be active in several neuroimaging studies of music, typically in discrimination tasks (discussed below). In addition, strong activations in right frontal operculum are observed when subjects are asked to imagine continuations of the opening fragments of familiar songs without words [26]. Previous work has established that mental imagery for motor behavior, vision, or audition can activate similar brain areas as actual action or perception. Therefore, mental imagery for melodic continuations can be viewed as a form of covert music production, in other words covert singing. Such results, in combination with the present findings and those of Riecker et al. [60] and Langheim et al. [40], suggest that musical imagery tasks can produce activations similar to those for music perception and production tasks. Activations of the frontal operculum during covert singing tasks may provide further support for a key role of this area in the human song-control system (and conceivably instrumental performance as well). This may be especially true for tasks that require active musical processing (e.g., imitation, discrimination, improvisation) rather than automatic processing based on long-term storage [40]. The frontal operculum has been shown to be activated during tasks that involve the processing of rhythm and time-intervals in addition to the processing of pitch (see below). So it is conceivable that rhythm processing contributed to the activations seen in the frontal

operculum in this study. Further studies will be needed to distinguish pitch and rhythm effects in this region. At the same time, prefrontal cortex is thought to be involved generally in temporal sequencing of actions as well as in planning and expectancy [21]. Thus, the effects observed here in the frontal operculum may be due to basic aspects of temporal and sequence expectancies [73] (see later section on antiphonal imitation).

A hierarchical feature of the song system revealed in this study was the activation of the planum polare (BA 38) during complex musical tasks but not monotonic vocalization. This accords well with the results of Griffiths et al. [23], who performed a parametric analysis of brain regions whose activity correlated with increasing musical complexity using iterated rippled noise, which produces a sense of pitch by means of temporal structure. The planum polare was one of only two regions whose activity correlated with the degree of musical complexity, especially vis-à-vis monotonic sequences. Moreover, in another parametric analysis, Zatorre and Belin [82] demonstrated that activity in this region co-varied with the degree of spectral variation in a set of pure-tone patterns. The anterior temporal region has been implicated in a host of findings related to musical processing. For example, it has been shown that surgical resection of the anterior temporal lobe of the right hemisphere that includes the planum polare often results in losses in melodic processing [65 and 80]. Koelsch et al. [35] demonstrated strong bilateral activations in planum polare during discrimination tasks involving complex chord sequences, in which discriminations were based on oddball chords or timbres. Likewise, bilateral activations were observed in this region when expert pianists performed Bach's Italian Concerto from memory [51]. Finally, in a related study, this area was observed to be active in discrimination tasks for both melody and harmony [8]. The coordinates of our BA 38 activation—at (50, 8, -4) in melody repetition—is nearly identical to the right hemisphere activation reported by Zatorre and Belin [82] in their subtraction analysis of spectral vs. temporal processing for pure tones (with coordinates at [50, 10, -6]), and were just inferior to those reported by Koelsch et al. [35] during discrimination processing for chord clusters, deviant instruments and modulations (with coordinates at [49, 2, 2]).

In sum, a convergence of results suggests that the superior part of the temporal pole bilaterally may be a type of tertiary auditory cortex specialized for higher-level pitch processing such as that related to complex melodies and harmonies, including the affective responses that accompany such processing. However, the exact nature of the processing in the planum polare during these tasks is unclear. The responses may reflect increases in memory load per se for musical information across the contrasted tasks, or reflect processing for musical grammar used to organize the musical production for the melody repetition and harmonization tasks. This area was not active when musically experienced listeners tracked a melody as its tonality changed [29], suggesting that the area is not involved in this aspect of tonality. Further investigation is necessary to refine functional accounts of this area in humans. There may not be a strict homologue of this region in non-human primates. The anterior superior temporal gyrus in monkey seems to be involved in auditory processing, possibly converging with limbic inputs [48 and 54], and so the activations seen here in the current study might even reflect a role in emotional processing. In addition, because cells in monkey anterior superior temporal gyrus are selectively responsive to monkey calls, it has been proposed that this region may be part of the “what” stream of auditory processing [72].

None of the activations for the melodic repetition and harmonization tasks (minus rest) coincided with the areas reported in prior studies of musical rhythm, such as anterior cerebellum, left parietal cortex, or left frontal cortex (lateral BA 6) [51 and 64]. This suggests that the rhythmic variation present in the melody repetition and harmonization tasks, but absent in monotonic vocalization, was not affecting the pattern of activations attributed to singing. This validates our intuition in designing the study to illuminate brain representations of melodic and harmonic, rather than rhythmic, information. We selected an isometric monotonic control task expecting that differences in brain activity for isometric versus variable-rhythm stimuli would be so small as to not obscure the differences in brain activity between monotonic sequences and musical sequences.

An unexpected finding of this study was the robust overlap in activity amongst the monotonic vocalization, melody repetition, and harmonization tasks. This overlap suggests that, despite the use of the carrier syllable /da/ in our monotone task (Materials and methods), monotonic vocalization is more “musical” than “syllabic” in

nature. Indeed, this monotonic vocalization task may embody most of the cardinal features of human music. Seven aspects of this task connect it more with simple music than simple speech: (1) a fixed pitch (i.e., spectral specificity) was employed; (2) the vocalization tempo was relatively slow (with an overall syllable rate of 1.67 Hz compared to a rate of around 8–10 Hz for connected speech) and the vowel was extended in duration; (3) the vocalizing was repetitive; (4) the vocalization rhythm was isometric; (5) the subject was required to match pitch; (6) the subject was required to match rhythm; and (7) the subject was required to sing in alternation with another “musician” (i.e., a digital piano). So, this antiphonal monotonic vocalization task should not be seen as a non-musical control but instead as a model of some of the most important features of music. Monotones, in fact, are an integral component of the world's music, as seen in many chants and drones.

Another unexpected result was the absence of strong activations in the dorsolateral prefrontal cortex or associated areas during the melody repetition and harmonization tasks, tasks that clearly required the storage of pitch information in working memory. We observed an activation in the dorsolateral prefrontal cortex (BA 46/9) for the monotonic vocalization task. Weaker activations were found in the identical location in the melody repetition and harmonization tasks, but these were below the threshold of significance for our tables (z values of 3.13 and 3.63, respectively). Despite the presence of activations in these regions, we are still surprised by their weakness, especially given the requirements of the tasks. For the moment, we do not have a good explanation for these results. Another manner in which such prefrontal activations might have shown up was in relation to the abrupt transitions occurring the monotone task, but again these were not seen when monotonic vocalization was compared to rest.

In sum, our results differ in the following respects from those of the previous studies of singing. First, Perry et al.'s [53] study of monotone singing did not show activations in BA 38. This was the case for our monotone task as well. The BA 38 activations were observed only when complex musical stimuli involving full melodies were used, as in our melody repetition and harmonization tasks. Second, Riecker et al.'s [60] study of the singing of familiar songs did not produce activations in the frontal operculum. As we argue throughout, the frontal operculum activations appear to be related to specific features of our tasks, namely a requirement for matching musical templates. Recalling familiar melodies from long-term memory does not seem to activate this process, whereas all three of our imitative tasks require subjects to match the pitch and rhythm of novel sequences. Overall, then, the use of complex and novel melodies enabled identification of the roles of two regions of the musical brain, namely the superior part of the temporal pole (BA 38) and the opercular part of the inferior frontal gyrus (BA 44/6).

3.2. The neural basis of polyphony

Harmonization resembles monophonic singing in that both involve the creation of a single melodic line. Harmonization differs from simple melody formation, however, in that it is done in coordination with a simultaneous musical template. One complication in interpreting activation differences between the harmony and melody tasks in this study is that both tasks activated functional brain regions closely overlapping with those elicited by monotonic vocalization. Bearing this in mind, there appeared to be a trend toward greater bilaterality in higher-level auditory areas (both BA 22 and BA 38) for harmony than for melody (when contrasted with rest). It is not possible to determine yet whether the bilaterality seen in our harmony tasks is due to a true specialization of left-hemisphere auditory areas for harmony processing or merely a quantitative acoustic effect due to the presence of a greater number of notes and a thicker musical texture in the harmony condition [35]. A study in which note number is directly controlled for will be needed to resolve this issue. Although there are computational differences between the processing of melody and harmony in particular tasks, neuroimaging studies at the current limits of resolution provide limited support for the view that harmony is mediated by brain areas distinct from those underlying melody. If this null hypothesis were to be confirmed by studies at higher resolution and with a variety of other paradigms of comparison, it would imply that the capacity to perceive and produce harmony is essentially contained within a basic melodic system, perhaps suggesting that the human harmony system emerged from a basic melodic system in which individual parts were temporally blended with one another following developments in temporal processing. This line of

investigation may provide insight into a classic debate regarding whether the origins of music are to be found in melody [63] or in harmony [58].

3.3. Neural systems for antiphonal imitation

It has been proposed that there is a system of “mirror neurons” specialized for the kinds of imitative behaviors that underlie such things as antiphonal imitation, or what have been referred to as resonance behaviors [62]. During resonance behaviors, organisms act by mirroring the activities of others, either behaviorally or cognitively. The focus of such a mirror system has generally been on visual/manual matching (but see [36]); however, such a system would be an equally plausible foundation for *audiovocal* matching functions such as song and speech. Both music and speech, like many forms of bird song, develop ontogenetically through a process of imitation of adult role models during critical periods in brain development [39, 50, 55 and 74]. These are additional instances of vocal learning, wherein developing organisms acquire their species-specific communication repertoires through imitative processes [30 and 31].

One region of the monkey brain that has been shown to possess mirror neurons is a premotor area thought to be the human homologue of Broca's area bilaterally. This region overlaps the opercular area identified bilaterally in the current study as being important for the template-matching processes underlying the antiphonal production of song. From this point of view, then, the frontal operculum may be part of a mirror system involved in audiovocal template matching for both pitch and rhythm. Template matching is also essential to discrimination processes, and tasks in which music-related stimuli are discriminated often show activations in the frontal operculum. This has been shown to be the case for the discrimination of pitch [24, 81, 83 and 84], chords [43], durations [24], rhythms [51], time intervals [59], sound intensities [2], chords, keys and timbres [35], melodies and harmonies [8], and melody and harmony performance during score reading [51]. Hence, it appears that the frontal operculum is equally important for pitch and rhythm processing in music, and that its functional role transcends motor aspects of vocalization. In sum, a mirror function for Broca's area may have as much explanatory power for imitative audiovocal processes underlying music and speech as it does for visuomanual matching processes underlying a proposed gestural origin of language [61] (see also [27]). If so, this might suggest that the song system of the human brain evolved from a vocalization system based on antiphonal imitation [6] in which the frontal operculum developed a specialized role to mediate this function.

Acknowledgements

We are grateful to Tim Griffiths, Carol Krumhansl, Aniruddh Patel, Frederic Theunissen, Barbara Tillmann, and Patrick Wong for their insightful comments on the manuscript. This work was supported by a grant from the ChevronTexaco Foundation.

References

1. L.F. Baptista, Nature and its nurturing in avian vocal development. In: D.E. Kroodsma and E.H. Miller, Editors, *Ecology and Evolution of Acoustic Communication in Birds*, Cornell University Press, Ithaca (1996), pp. 39–60.
2. P. Belin, S. McAdams, B. Smith, S. Savel, L. Thivard, S. Samson and Y. Samson, The functional anatomy of sound intensity discrimination. *J. Neurosci.* 18 (1998), pp. 6388–6394.
3. P. Belin, R.J. Zatorre, P. Lafaille, P. Ahad and B. Pike, Voice-selective areas in human auditory cortex. *Nature* 403 (2000), pp. 309–312.
4. P. Belin, R.J. Zatorre and P. Ahad, Human temporal-lobe response to vocal sounds. *Cogn. Brain Res.* 13 (2002), pp. 17–26.
5. S. Brown, Evolutionary models of music: from sexual selection to group selection. In: F. Tonneau and N.S. Thompson, Editors, *Perspectives in Ethology: 13. Behavior, Evolution and Culture*, Plenum, New York (2000), pp. 231–281.

6. S. Brown, Contagious heterophony: a new theory about the origins of music, in: R. Tsurtsumia (Ed.), *Problems of Traditional Polyphony*, Tbilisi State Conservatory, Tbilisi, in press.
7. E. Brown and S.M. Farabaugh, Song sharing in a group-living songbird, the Australian magpie, *Gymnorhina tibicen*: Part III. Sex specificity and individual specificity of vocal parts in communal chorus and duet songs. *Behaviour* 118 (1991), pp. 244–274.
8. S. Brown, L.M. Parsons, M.J. Martinez, D.A. Hodges, C. Krumhansl, J. Xiong, P.T. Fox, The neural bases of producing, improvising, and perceiving music and language. Proceedings of the Annual Meeting of the Cognitive Neuroscience Society, *Journal of Cognitive Neuroscience*, in press.
9. S. Catrin Blank, S.K. Scott, K. Murphy, E. Warburton and R.J.S. Wise, Speech production: Wernicke, Broca and beyond. *Brain* 125 (2002), pp. 1829–1838.
10. E. Condillac, *An Essay on the Origin of Human Knowledge*. English translation by R.G. Weyant (1756). Reprinted in facsimile form, (1971). *Scholars' Facsimiles & Reprints*, Gainesville, 1746.
11. R.B. D'Agostino, A. Belatner and R.B. D'Agostino, Jr., A suggestion for using powerful and informative tests of normality. *Am. Stat.* 44 (1990), pp. 316–321.
12. C. Darwin, *The Descent of Man, and Selection in Relation to Sex.*, J. Murray, London (1871).
13. M.R. DeLong, The basal ganglia. In: E.R. Kandel, J.H. Schwartz and T.M. Jessell, Editors, *Principles of Neural Science*, McGraw-Hill, New York (2000), pp. 853–867.
14. C. Deng, G. Kaplan and L.J. Rogers, Similarity of the song nuclei of male and female Australian magpies (*Gymnorhina tibicen*). *Behav. Brain Res.* 123 (2001), pp. 89–102.
15. N.F. Dronkers, A new brain region for coordinating speech articulation. *Nature* 384 (1996), pp. 159–161.
16. M. Eens, R. Pinxten and R.F. Verheyen, No overlap in song repertoire between yearling and older starlings *Sturnus vulgaris*. *Ibis* 134 (1992), pp. 72–76.
17. P.T. Fox and M. Mintun, Noninvasive functional brain mapping by change-distribution analysis of averaged PET images of H₂¹⁵O tissue activity. *J. Nucl. Med.* 30 (1989), pp. 141–149.
18. P.T. Fox, M. Mintun, E. Reiman and M.E. Raichle, Enhanced detection of focal brain responses using inter-subject averaging and change-distribution analysis of subtracted PET images. *J. Cereb. Blood Flow Metab.* 8 (1988), pp. 642–653.
19. P.T. Fox, A. Huang, L.M. Parsons, J. Xiong, F. Zamarripa and J.L. Lancaster, Location-probability profiles for the mouth region of human primary motor-sensory cortex: model and validation. *NeuroImage* 13 (2001), pp. 196–209.
20. K.J. Friston, C.D. Frith, P.R. Liddle and R.S.J. Frackowiak, Comparing functional (PET) images: the assessment of significant change. *J. Cereb. Blood Flow Metab.* 11 (1991), pp. 690–699.
21. J.M. Fuster, The prefrontal cortex—an update: time is of the essence. *Neuron* 30 (2001), pp. 319–333.
22. P.M. Gray, B. Krause, J. Atema, R. Payne, C. Krumhansl and L. Baptista, Biology and music: the music of nature. *Science* 291 (2001), pp. 52–54.
23. T.D. Griffiths, C. Büchel, R.S.J. Frackowiak and R.D. Patterson, Analysis of temporal structure in sound by the human brain. *Nat. Neurosci.* 1 (1998), pp. 422–427.
24. T. Griffiths, I. Johnsrude, J.L. Dean and G.G.R. Green, A common neural substrate for the analysis of pitch and duration pattern in segmented sound?. *NeuroReport* 10 (1999), pp. 3825–3830.
25. E.H. Hagen and G.A. Bryant, Music and dance as a coalition signaling system. *Hum. Nat.* 14 (2003), pp. 21–51.
26. A.R. Halpern and R.J. Zatorre, When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. *Cereb. Cortex* 9 (1999), pp. 697–704.
27. M.D. Hauser, N. Chomsky and W.T. Fitch, The faculty of language: what is it, who has it and how did it evolve?. *Science* 298 (2002), pp. 1569–1579.
28. S.E. Henschen, On the function of the right hemisphere of the brain in relation to the left in speech, music and calculation. *Brain* 49 (1926), pp. 110–123.
29. P. Janata, J.L. Birk, J.D. Van Horn, M. Leman, B. Tillmann and J. Bharucha, The cortical topography of tonal structures underlying western music. *Science* 298 (2002), pp. 2167–2170.
30. V.M. Janik and P.J.B. Slater, Vocal learning in mammals. *Adv. Study Behav.* 26 (1997), pp. 59–99. Abstract
31. V.M. Janik and P.J.B. Slater, The different roles of social learning in vocal communication. *Anim. Behav.* 60 (2000), pp. 1–11.

32. U. Jürgens, On the neurobiology of vocal communication. In: H. Papousek, U. Jürgens and M. Papousek, Editors, *Nonverbal Vocal Communication: Comparative and Developmental Approaches*, Cambridge Univ. Press, Cambridge (1992), pp. 31–42.
33. U. Jürgens, Neural pathways underlying vocal control. *Neurosci. Biobehav. Rev.* 26 (2002), pp. 235–258.
34. D. Klein, R.J. Zatorre, B. Milner, E. Meyer and A.C. Evans, Left putaminal activation when speaking a second language: evidence from PET. *NeuroReport* 21 (1994), pp. 2295–2297.
35. S. Koelsch, T.C. Gunter, D.Y. v Cramon, S. Zysset, G. Lohmann and A.D. Friederici, Bach speaks: a cortical “language-network” serves the processing of music. *NeuroImage* 17 (2002), pp. 956–966.
36. E. Kohler, C. Keysers, M.A. Umiltà, L. Fogassi, V. Gallese and G. Rizzolatti, Hearing sounds, understanding actions: action representation in mirror neurons. *Science* 297 (2002), pp. 846–848.
37. D.E. Kroodsma and M. Konishi, A suboscine bird (eastern phoebe *Sayornis phoebe*) develops normal song without auditory feedback. *Anim. Behav.* 42 (1991), pp. 477–487.
38. D.E. Kroodsma, W.-C. Liu, E. Goodwin and P.A. Bedell, The ecology of song improvisation as illustrated by North American sedge wrens. *Auk* 116 (1999), pp. 373–386.
39. P.K. Kuhl and A.N. Meltzoff, Infant vocalizations in response to speech: vocal imitation and developmental change. *J. Acoust. Soc. Am.* 100 (1996), pp. 2425–2438.
40. F.J.P. Langheim, J.H. Callicott, V.S. Mattay, J.H. Duyn and D.R. Weinberger, Cortical systems associated with covert musical rehearsal. *NeuroImage* 16 (2002), pp. 901–908.
41. S.A. MacDougall-Shackleton and G.F. Ball, Comparative studies of sex differences in the song-control system of songbirds. *Trends Neurosci.* 22 (1999), pp. 432–436.
42. S. MacDougall-Shackleton and S.H. Hulse, Concurrent absolute and relative pitch processing by European starlings (*Sturnus vulgaris*). *J. Comp. Psychol.* 110 (1996), pp. 139–146.
43. B. Maess, S. Koelsch, T.C. Gunter and A.D. Friederici, Musical syntax is processed in Broca's area: an MEG study. *Nat. Neurosci.* 4 (2001), pp. 540–545.
44. P. Marler and R. Pickert, Species-universal microstructure in the learned song of the swamp sparrow (*Melospiza georgiana*). *Anim. Behav.* 32 (1984), pp. 679–689.
45. V. Menon, D.J. Levitin, B.K. Smith, A. Lembke, B.D. Krasnow, D. Glazer, G.H. Glover and S. McAdams, Neural correlates of timbre change in harmonic sounds. *NeuroImage* 17 (2002), pp. 1742–1754.
46. G.S. Miller, *The Mating Mind: How Sexual Selection Shaped the Evolution of Human Nature.*, Doubleday, New York (2000).
47. M. Mintun, P.T. Fox and M.E. Raichle, A highly accurate method of localizing regions of neuronal activity in the human brain with PET. *J. Cereb. Blood Flow Metab.* 9 (1989), pp. 96–103.
48. M.A. Moran, E.J. Mufson and M.M. Mesulam, Neural inputs into the temporopolar cortex of the rhesus monkey. *J. Comp. Neurol.* 256 (1987), pp. 88–103.
49. R. Oldfield, The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9 (1971), pp. 97–113.
50. M. Papousek, Intuitive parenting: a hidden source of musical stimulation in infancy. In: I. Deliège and J. Sloboda, Editors, *Musical Beginnings: Origins and Development of Musical Competence*, Oxford Univ. Press, Oxford (1996), pp. 88–112.
51. L.M. Parsons, Exploring the functional neuroanatomy of music performance, perception and comprehension. In: R.J. Zatorre and I. Peretz, Editors, *The Cognitive Neuroscience of Music*, Oxford Univ. Press, Oxford, UK (2003), pp. 247–268.
52. L.M. Parsons and D. Osherson, New evidence for distinct right and left brain systems for deductive versus probabilistic reasoning. *Cereb. Cortex* 11 (2001), pp. 954–965.
53. D.W. Perry, R.J. Zatorre, M. Petrides, B. Alivisatos, E. Meyer and A.C. Evans, Localization of cerebral activity during simple singing. *NeuroReport* 10 (1999), pp. 3979–3984.
54. A. Poremba, R.C. Saunders, A.M. Crane, M. Cook, L. Solokoff and M. Mishkin, Functional mapping of primate auditory cortex. *Science* 299 (2003), pp. 568–572.
55. C.L. Poulson, E. Kymissis, K.F. Reeve, M. Andreators and L. Reeve, Generalized vocal imitation in infants. *J. Exp. Child Psychol.* 51 (1991), pp. 267–279.

56. J. Rademacher, P. Morosan, T. Schormann, A. Schleicher, C. Werner, H.J. Freund and K. Zilles, Probabilistic mapping and volume measurement of human primary auditory cortex. *NeuroImage* 13 (2001), pp. 669–683.
57. M.E. Raichle, M.R.W. Martin, P. Herskovitch, M.A. Mintun and J. Markham, Brain blood flow measured with intravenous H₂¹⁵O: II. Implementation and validation. *J. Nucl. Med.* 24 (1983), pp. 790–798.
58. J.-P. Rameau, Treatise on Harmony. Philip Gossett, translator. , Dover Publications, New York (1722/1971).
59. S.M. Rao, A.R. Mayer and D.L. Harrington, The evolution of brain activation during temporal processing. *Nat. Neurosci.* 4 (2001), pp. 317–323.
60. A. Riecker, H. Ackermann, D. Wildgruber, G. Dogil and W. Grodd, Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *NeuroReport* 11 (2000), pp. 1997–2000.
61. G. Rizzolatti and M.A. Arbib, Language within our grasp. *Trends Neurosci.* 21 (1998), pp. 188–194.
62. G. Rizzolatti, L. Fadiga, L. Fogassi and V. Gallese, Resonance behaviors and mirror neurons. *Arch. Ital. Biol.* 137 (1999), pp. 85–100.
63. J.-J. Rousseau, Essay on the origin of languages. In: J.T. Scott, Editor, *Essay on the Origin of Languages and Writings Related to Music*, University Press of New England, Hanover (1781/1998), pp. 289–332.
64. K. Sakai, O. Hikosaki, S. Miyauchi, R. Takino, T. Tamada, N.K. Iwata and M. Nielsen, Neural representations of a rhythm depends on its interval ratio. *J. Neurosci.* 19 (1999), pp. 10074–10081.
65. S. Samson and R.J. Zatorre, Melodic and harmonic discrimination following unilateral cerebral excision. *Brain Cogn.* 7 (1988), pp. 348–360.
66. M. Satoh, K. Takeda, K. Nagata, J. Hatazawa and S. Kuzuhara, Activated brain regions in musicians during an ensemble: a PET study. *Cogn. Brain Res.* 12 (2001), pp. 101–108.
67. J.D. Schmahmann, J.A. Doyon, D. McDonald, C. Holmes, K. Lavoie, A. Hurwitz, N. Kabani, A. Toga, E. Evans and M. Petrides, Three-dimensional MRI atlas of the human cerebellum in proportional stereotaxic space. *NeuroImage* 10 (1999), pp. 233–260.
68. W.A. Searcy, S. Nowicki and S. Peters, Song types as fundamental units in vocal repertoires. *Anim. Behav.* 58 (1999), pp. 37–44.
69. P.J.B. Slater, Birdsong repertoires: their origins and use. In: N.L. Wallin, B. Merker and S. Brown, Editors, *The Origins of Music*, MIT Press, Cambridge, MA (2000), pp. 49–63.
70. R. Stripling, L. Milewski, A.A. Kruse and D.F. Clayton, Rapidly learned song-discrimination without behavioral reinforcement in adult male zebra finches (*Taeniopygia guttata*). *Neurobiol. Learn Mem.* 79 (2003), pp. 49–50.
71. S.C. Strother, N. Lang, J.R. Anderson, K.A. Schaper, K. Rehm, L.K. Hansen and D.A. Rottenberg, Activation pattern reproducibility: measuring the effects of group size and data analysis models. *Hum. Brain Mapp.* 5 (1997), pp. 312–316.
72. B. Tian, D. Reser, A. Durham, A. Kustov and J.P. Rauschecker, Functional specialization in rhesus monkey auditory cortex. *Science* 292 (2001), pp. 290–293.
73. B. Tillmann, P. Janata and J.J. Bharucha, Activation of the inferior frontal cortex in musical priming. *Cogn. Brain Res.* 16 (2003), pp. 145–161.
74. S.E. Trehub, Musical predispositions in infancy. In: R.J. Zatorre and I. Peretz, Editors, *The Biological Foundations of Music*, New York Academy of Sciences, New York (2001), pp. 1–16.
75. P.E. Turkeltaub, G.F. Eden, K.M. Jones and T.A. Zeffiro, Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *NeuroImage* 16 (2002), pp. 765–780.
76. N.L. Wallin, B. Merker and S. Brown, Editors, *The Origins of Music*, MIT Press, Cambridge, MA (2000).
77. J.D. Warren, J.E. Warren, N.C. Fox and E.K. Warrington, Nothing to say, something to sing: primary progressive dynamic aphasia. *Neurocase* 9 (2003), pp. 140–155.
78. D. Wildgruber, H. Ackermann and W. Grodd, Differential contributions of motor cortex, basal ganglia, and cerebellum to speech motor control: effects of syllable repetition rate evaluated by fMRI. *NeuroImage* 13 (2001), pp. 101–109.
79. A. Yamadori, Y. Osumi, S. Masuhara and M. Okubo, Preservation of singing in Broca's aphasia. *J. Neurol. Neurosurg. Psychiatry* 40 (1977), pp. 221–224.

80. R.J. Zatorre, Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychologia* 23 (1985), pp. 31–41.
81. R.J. Zatorre and J.R. Binder, Functional and structural imaging of the human auditory cortex. In: A.W. Toga and J.C. Mazziotta, Editors, *Brain Mapping: The Systems*, Academic Press, San Diego (2000), pp. 365–402.
82. R.J. Zatorre and P. Belin, Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11 (2001), pp. 946–953.
83. R.J. Zatorre, A.C. Evans, E. Meyer and A. Gjedde, Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256 (1992), pp. 846–849.
84. R.J. Zatorre, A.C. Evans and E. Meyer, Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14 (1994), pp. 1908–1919.
85. R.J. Zatorre, E. Meyer, A. Gjedde and A.C. Evans, PET studies of phonetic processing of speech: review, replication, and reanalysis. *Cereb. Cortex* 6 (1996), pp. 21–30.
86. W. Ziegler, B. Kilian and K. Deger, The role of the left mesial frontal cortex in fluent speech: evidence from a case of left supplementary motor area hemorrhage. *Neuropsychologia* 35 (1997), pp. 1197–1208.
87. R.J. Zatorre, P. Belin and V.B. Penhune, Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci.* 6 (2002), pp. 37–46.