Finding Your Voice: A Singing Lesson From Functional Imaging

Sarah J. Wilson,^{1,2*} David F. Abbott,² Dean Lusher,¹ Ellen C. Gentle,¹
and Graeme D. Jackson²

¹*Psychological Sciences, The University of Melbourne, Australia*

²*Brain Research Institute, Florey Neuroscience Institutes (Austin), Melbourne, Australia*

Abstract: Vocal singing (singing with lyrics) shares features common to music and language but it is not clear to what extent they use the same brain systems, particularly at the higher cortical level, and how this varies with expertise. Twenty-six participants of varying singing ability performed two functional imaging tasks. The first examined covert generative language using orthographic lexical retrieval while the second required covert vocal singing of a well-known song. The neural networks subserving covert vocal singing and language were found to be proximally located, and their extent of cortical overlap varied with singing expertise. Nonexpert singers showed greater engagement of their language network during vocal singing, likely accounting for their less tuneful performance. In contrast, expert singers showed a more unilateral pattern of activation associated with reduced engagement of the right frontal lobe. The findings indicate that singing expertise promotes independence from the language network with decoupling producing more tuneful performance. This means that the age-old singing practice of ‘finding your singing voice’ may be neurologically mediated by changing how strongly singing is coupled to the language system. *Hum Brain Mapp* 32:2115–2130, 2011. © 2010 Wiley Periodicals, Inc.

Key words: singing; music; language; musical expertise; fMRI

INTRODUCTION

There has been a long fascination with singing behavior in animals and humans. Neuroscience of singing in birds has led to fundamental insights into mechanisms of neural

plasticity, learning, and motor control that have informed our understanding of human behavior [Miller, 2003]. By comparison, there is a paucity of research directly examining singing behavior in humans, particularly as it relates to the organization of higher cortical functions [Brown et al., 2004].

Vocal singing is a universal human behavior that spontaneously emerges in infants at around 12 months of age [Peretz et al., 2004]. Singing to infants is biologically significant for modulating infant attention and arousal and facilitating parent-infant bonding through the conveyance of positive emotions [Trehub, 2001, 2003]. Singing by infants precedes and may benefit language development, with vocal singing sharing features common to music and language [Callan et al., 2006; Schön et al., 2005]. These shared features provide an opportunity to examine the organization of one higher cortical function relative to another.

Despite shared features, there are well-known differences in the roles of vocal singing and language for the individual and society [McDermott, 2008; Mithen, 2009]. For example a group can sing in unison, while talking does not serve

Additional Supporting Information may be found in the online version of this article.

Contract grant sponsor: Australian Research Council; Contract grant number: DP0449862; Contract grant sponsors: University of Auckland Research Committee; Operational Infrastructure Support Program of the State Government of Victoria, Australia.

*Correspondence to: Sarah J. Wilson, Psychological Sciences, The University of Melbourne, Victoria, 3010, Australia.

E-mail: sarahw@unimelb.edu.au

Received for publication 8 April 2009; Revised 16 August 2010; Accepted 27 August 2010

DOI: 10.1002/hbm.21173

Published online 15 December 2010 in Wiley Online Library (wileyonlinelibrary.com).

group functions in this way. Singing has an important role in social cohesion and motivation, and is important in ritual and group identity [Dalla Bella et al., 2007; Wallin et al., 2000]. Preservation of actual words is usually high in singing and low in story telling. Either the same cognitive system is serving different functions, or there is a distinct neurological network for singing that serves these functions.

Investigating higher cortical functions independently of each other may limit our understanding of how one function relates to another and whether changes in one, either through training or cerebral damage, impacts the other. Knowledge of interactions between higher cortical functions is likely to yield basic insights into the evolutionary design of human cognition and behavior, as well as have implications for the development of educational programs and clinically therapeutic interventions. As an example, melodic intonation therapy (MIT) uses vocal singing to facilitate language production in patients with nonfluent aphasia [Helm-Estebrooks, 1983; Racette et al., 2006; Schlaug et al., 2008; Wilson et al., 2006]. Recently, Schlaug et al. [2008] proposed that this leads to greater right hemisphere involvement in speech output due to priming of sensorimotor and premotor cortices.

Early neuroimaging reports and neuropsychological studies provided strong support for right-lateralized singing in contrast to left-lateralized language. In particular, patient lesion studies pointed to the prominent role of the right frontal lobe in singing [Henson, 1985; Jeffries et al., 2003; Peretz et al., 2004; Perry et al., 1999; Racette et al., 2006; Riecker et al., 2000; Wildgruber et al., 1996; Wilson et al., 2006; Yamadori et al., 1997]. More recent neuroimaging studies in neurologically intact individuals have reported bihemispheric activation associated with singing, and to a lesser extent language, with the latter thought due to the faster rate of speech production [Brown et al., 1999; Schlaug et al., 2008]. The same bihemispheric network is now considered to underpin singing individual notes, tunes, and harmonies [Zarate and Zatorre, 2008]. Across singing tasks, activation is common in auditory regions in the superior temporal gyrus, sensorimotor, premotor, and supplementary motor areas, the inferior frontal gyrus, cingulate, insula, and cerebellum [Brown et al., 2004, 2006; Callan et al., 2006; Hickok et al., 2003; Jeffries et al., 2003; Kleber et al., 2007, 2009; Koelsch et al., 2009; Özdemir et al., 2006; Perry et al., 1999; Riecker et al., 2000; Saito et al., 2006; Wildgruber et al., 1996; Zarate and Zatorre, 2008].

To date, there has been limited research investigating whether this network differs for expert and nonexpert singers [Formby et al., 1989; Kleber et al., 2009; Zarate and Zatorre, 2008]. Using a conjunction analysis, Zarate and Zatorre [2008] showed recruitment of similar areas when experts and nonexperts overtly sang individual pitches. They also noted differences (at $P < 0.001$ uncorrected), with activation of a more specialized network for audio-vocal integration in expert singers, contrasting with greater activation of premotor cortex in nonexpert singers that was thought to underpin general sensorimotor interactions. Consistent with this, Kleber et al. [2009] showed more focused (or task relevant) activation in experts compared with nonexperts when overtly singing an Italian aria. This activation

occurred in primary somatosensory cortex important for kinesthetic motor control, consistent with previous studies showing reduced premotor activation but circumscribed increases in sensorimotor regions associated with instrumental skills in expert performers [Haslinger et al., 2004; Hund-Georgiadis and von Cramon, 1999; Jänke et al., 2000; Krings et al., 2000; Lotze et al., 2003].

Studies using overt auditory and motor imaging paradigms have typically shown similar cortical activation patterns to covert paradigms, albeit with greater engagement of auditory and sensorimotor regions, and the insula [Gunji et al., 2007; Kleber et al., 2007; Langheim et al., 2002; Riecker et al., 2000; Shuster and Lemieux, 2005]. Zatorre and Halpern [2005] noted that covert paradigms provide a reliable method of assessing the neural structures underpinning musical abilities, including singing. Of particular benefit, covert paradigms minimize movement and respiratory artifacts and focus the analysis on higher cognitive representations rather than the primary sensory and motor aspects of singing and speech production [Formby et al., 1989; Gunji et al., 2007; Kleber et al., 2007; Koelsch et al., 2009; Langheim et al., 2002].

The shared correlates of singing and speech have also commonly included the auditory and sensorimotor cortices, with the majority of studies using overt paradigms [Brown et al., 2006; Jeffries et al., 2003; Özdemir et al., 2006; Riecker et al., 2000; Saito et al., 2006; Wildgruber et al., 1996]. With the exception of Brown et al. [2006], these studies used speech recitation or repetition rather than generative language tasks commonly used to localize language. This limits our understanding of the extent of cortical overlap for familiar vocal singing relative to generative language typical of everyday life. We know of only one study investigating shared correlates relative to singing expertise [Formby et al., 1989]. While no differences were found, this study was conducted before the availability of fMRI (^{133}Xe inhalation was used) and the authors noted several methodological limitations that might explain their negative result.

The aim of our study was to investigate the relationship between music and language functions and their interaction with expertise using covert vocal singing and language generation tasks. We employed fMRI to compare patterns of cortical activation associated with singing a familiar tune with lyrics (vocal singing) to generative language. We hypothesized that (i) there would be overlap in the neural structures underpinning covert vocal singing and generative language, (ii) expert singers would use partially different singing networks to nonexpert singers, and (iii) the overlap between singing and language networks would vary according to individual differences in singing expertise.

METHODS

Participants

Twenty-six adults of varying singing expertise were recruited from the Faculty of Music and the Department of

TABLE I. The singing groups derived from out-of-scanner assessment

Characteristic	High pitch accuracy (<i>n</i> = 10)	Mid-pitch accuracy (<i>n</i> = 7)	Low pitch accuracy (<i>n</i> = 9)
Number of females	8	4	3
Mean years of age ($\pm$ SD); Range	30.6 $\pm$ 9.9; 19–50	31.6 $\pm$ 14.3; 18–52	33.1 $\pm$ 14.5; 19–52
Mean years of education ($\pm$ SD) ^a ; range	17.2 $\pm$ 3.5; 12–25	15.3 $\pm$ 2.1; 12–18	17.0 $\pm$ 5.2; 11–29
Laterality Index ($\pm$ SD) ^{a, b} ; Range	84.4 $\pm$ 15.9; 50–100	79.3 $\pm$ 16.4; 50–100	73.9 $\pm$ 42.9; –20–100
Early exposure to singing at home ($\pm$ SD) ^c	3.8 $\pm$ 1.0	3.7 $\pm$ 0.8	3.0 $\pm$ 1.2
Mean years of singing experience ($\pm$ SD) ^{a, d} ; Range	10.4 $\pm$ 9.6; 0–30	1.6 $\pm$ 2.7; 0–6	0
Percent pitch accuracy ($\pm$ SD); Range	81.2 $\pm$ 8.9; 70.5–98.1	60.3 $\pm$ 3.1; 57.6–65.2	34.5 $\pm$ 9.7; 14.8–46.2

The table shows the sociodemographic and musical characteristics of the three singing groups.

^a1 case of missing data from the high pitch accuracy group.

^bThe Laterality Index was calculated from the *Edinburgh Handedness Inventory* [Oldfield, 1971]. Raw scores were converted to an index within the range of –100 to +100, with positive scores representing right-hand predominance.

^cEarly exposure to singing in the home was rated by participants on a 5-point Likert type scale, ranging from 1 = Never to 5 = Every-day. There were two cases of missing data from the high pitch accuracy group and one case from the mid-pitch accuracy group.

^dSinging experience was based on the total number of years participants reported they had been actively engaged in singing practice, such as public performances, choral singing, and formal singing training. No weighting was applied to these different activities, which in many individuals occurred simultaneously.

Psychology, The University of Melbourne, the School of Music, Victorian College of the Arts, University-related choirs, and community-based volunteers. The study received approval from the relevant Human Research Ethics Committees and all participants gave written informed consent in accordance with the Declaration of Helsinki.

All participants underwent a medical screen for significant neurological, psychiatric, and hearing impairments. They completed a detailed history of their music training and background using the *Survey of Musical Experience* [Wilson et al., 1999], with additional questions focusing on singing experience. This allowed years of singing experience to be estimated from their accumulated amount of deliberate singing practice from starting age to the present [Ericsson, 1997], including any formal singing training, choral singing, and public performances. Duration of singing experience was used rather than duration of formal singing training, as singers often show variability in the amount of formal training they have received, and many begin training later in life after their voice has matured.

The final sample comprised 15 females and 11 males with a mean age of 31.7 years (SD = 12.4), and an average of 16.6 years education (SD = 3.8; see Table I). All participants were predominantly right-handed, as indicated by a mean Laterality Index of 79.2 (SD = 28) on the *Edinburgh Handedness Inventory* [Oldfield, 1971]. No participant had absolute or quasi-absolute pitch as based on self-report.

Out-of-Scanner Assessment of Singing Expertise

Prior to scanning, a pitch accuracy score was derived by asking each participant to sing the melody of the main theme of the Finale of the *William Tell Overture* by Rossini into a microphone as if performing. Singing without lyrics

was used to avoid pitch errors associated with the production of varying phonemes and was performed using the syllable “deh”. The main theme of the Finale was chosen for a range of reasons, including its similar length, and pitch and rhythmic structure to the in-scanner vocal singing task, its high familiarity in the Australian population, and its suitable level of difficulty to challenge expert singers. The melody of the theme was played to the participants in the key of C Major using a synthesized “grand piano” timbre of a Yamaha S80 keyboard. Their performance was then recorded using ProTools LE (DIGI001 hardware, ProTools LE software, version 5.1; www.digidesign.com), and imported into Praat (version 4.3.22 for Mac OSX; www.praat.org) to analyze the accuracy of their pitch production.

The average of the fundamental frequencies of the first three notes (with the same pitch) was used to estimate each participant’s starting pitch. The accuracy of the fundamental frequencies of subsequent pitches was assessed relative to this starting pitch in accordance with the melodic template played to the participants. In total, 42 notes were each assigned a score out of five based on the percent variation of pitch ratios from the expected ratios using a 5-point ordinal scale (5 = $\pm$ 2.5%, 4 = $\pm$ 2.5–5.0%, 3 = $\pm$ 5.0–7.5%, 2 = $\pm$ 7.5–10.0%, 1 = $\pm$ 10.0–12.5%). This produced a total score out of 210 for each participant that was then converted to a pitch accuracy score ranging from 0 (low) to 100 (high). We employed a 5-point ordinal scale to address missed (unsung) notes in the performance (scored 0). This ensured that each sung note was equally valued and had a constrained range, and avoided the need to specify a subjective cut-off point for missing values. It also had the effect of minimising the influence of large pitch deviations for particular notes on the pitch accuracy scores, producing scores that were normally distributed for use in statistical analyses.

Cognitive Activation Paradigms

To ensure that the in-scanner vocal singing and language tasks were adequately completed they were performed before scanning to familiarize participants with the cognitive paradigms and to obtain behavioral measures of performance accuracy. For the vocal singing task, a template of the song was played to each participant out-of-scanner and prior to the relevant task-blocks in-scanner to promote a consistent rate of performance. After scanning, detailed questioning was undertaken to ensure that each participant adequately completed the in-scanner tasks. They also rated the level of familiarity of the vocal singing task, including the familiarity of the lyrics and the tune using a 5-point Likert-type scale (1 = Not at all familiar, 5 = Totally familiar).

The cognitive paradigms during functional imaging were of block design. For both tasks the stimuli were repeatedly performed “in mind” as a first-person action during task-blocks without mouthing the words. The vocal singing task alternated singing a familiar tune with lyrics with rest (sing-rest). The song comprised the first two lines of the chorus of a well-known Australian folk song, *Waltzing Matilda*. The played song template had a duration of 10 s, allowing almost two complete repetitions of the song in each task-block. During rest-blocks, participants were asked to “relax and try not to think about anything in particular.” The duration of each task- and each rest-block was 18 s. The visual signals “sing” and “rest” were used to control the commencement and cessation of imagined singing and rest in each block.

Language activation was assessed using an orthographically cued lexical retrieval (OLR) task in accordance with our previously published methods [Wood et al., 2001]. In short, this is a well-established generative language task that is routinely used by our group to reliably activate the language network. It comprised alternating 36 s rest-blocks (described above) and 36 s task-blocks (speech-rest). During task-blocks, participants were visually presented with a letter of the alphabet and asked to covertly generate as many words as possible beginning with that letter. Rules included avoiding proper nouns and derivatives or extensions of a given word (e.g., run, running). A new letter appeared after 18 s resulting in two letters per task-block. Out-of-scanner assessment of generative language was based on performance of the Controlled Oral Word Association Test (COWAT), undertaken before the scanning session. This is the overt equivalent of the covert, in-scanner task [Wood et al., 2001].

The singing and language tasks constituted one functional run each. The functional run for singing contained eight task-blocks and nine rest-blocks in total, while the language run comprised four task-blocks and five rest-blocks in total. With three exceptions, functional runs were performed on the same day with speech-rest followed by sing-rest. Scanner difficulties meant that two participants performed speech-rest after sing-rest on the same day,

while a further participant undertook the language study on a different day.

Behavioral Data Analysis

The behavioral data met assumptions for parametric analysis that was performed using SPSS (version 13.0.0 for Mac OS X; www.spss.com), with $P < 0.05$ (two-tailed) set as the criterion of statistical significance. To identify groups of singers of varying expertise, a Ward’s hierarchical cluster analysis with squared Euclidian distances was used to explore the existence of natural groupings in the pitch accuracy data derived from the out-of-scanner singing task. This analysis progressively groups individuals with the most similar scores, minimising the variability within clusters and maximizing the variance between clusters [Ward, 1963]. Independent samples t -tests and analysis of variance (ANOVA) with planned contrasts were then used for group comparisons on behavioral measures, including demographic and background singing variables, familiarity of the singing task, and out-of-scanner assessment of language.

fMRI Acquisition

Functional MRI was performed using a 3.0 tesla GE Signa LX whole body scanner (General Electric, Milwaukee, WI) with standard birdcage quadrature transmit-receive head coil. Participants were fixed using a Velcro strap over the forehead. Functional images were acquired using a multi-slice Echo Planar Imaging (EPI) sequence (single shot gradient recalled echo) providing T2*-weighted blood-oxygenation-level dependent (BOLD) contrast [Ogawa et al., 1990]. Functional MRI acquisition parameters were as follows: repetition time (TR) = 3.0 s, echo-time (TE) = 40 ms, flip-angle = 60°, 25 axial oblique slices 4 mm thick +1 mm gap, field of view (FOV) = 24 cm, 128 × 128 matrix, 1.88 × 1.88 mm in-plane. Thus an image volume consisting of a series of slices covering the whole brain was acquired each TR. The first four image volumes acquired in each run were automatically discarded to allow magnetization to reach a steady-state.

The collected images were converted to Analyze format using iBrain^{TM} [Abbott and Jackson, 2001] and then pre-processed using Statistical Parametric Mapping software (SPM8 release 3408; Wellcome Department of Imaging Neuroscience, London, UK) with the aid of iBrain^{TM}. Images were first slice-time corrected, using a temporal interpolation scheme to estimate the response at the time of commencement of each acquisition volume. Images were then realigned to a single target image within the time series to minimize the effects of participant motion between scans. The target selected by iBrain^{TM} was the image whose within-brain centre-of-mass was located closest to the median of all images in the time series. Slice-time-corrected, realigned images were then spatially normalized to a common space by coregistering to the

standard EPI template supplied with SPM8, which is approximately in the space of the 152 brain template of the Montreal Neurological Institute.

Because we wished to compare different runs within each participant, and distortions inherent in EPI images can differ between runs, we undertook a multistage spatial normalization procedure that improved intrasubject interrater registration. We also incorporated an explicit intensity non-uniformity (bias) correction step, as without it, we found that some participant images were not adequately normalized to the standard template. Specifically, for each run the mean of the slice-time-corrected, realigned images was obtained, and each run mean was approximately registered (affine normalization) to the standard EPI template, then bias corrected using the Segment module of SPM8. An iterative procedure was then employed to coregister (nonlinear normalization) all the bias corrected images for each participant. In the first iteration the “target” was the bias corrected image from the speech-rest functional run. In subsequent iterations the target was the mean across runs of the result of the previous iteration. After three iterations, the resultant mean image for the participant was then nonlinearly spatially normalized to the standard EPI template. For each run, the relevant spatial transformation matrices were combined in the deformations toolbox of SPM8 and applied to the original slice-time-corrected, realigned images. Finally, images were smoothed with an isotropic Gaussian kernel of full-width-at-half-maximum (FWHM) = 8.0 mm.

fMRI Analysis

Statistical analysis was conducted in SPM8 using a general linear model. The first rest-block of each scanning run was disregarded from the imaging analysis to allow participants plenty of time to settle when each run commenced. In addition to the effects of interest (task and rest), effects of no interest consisting of the six rigid-body transformation parameters estimated during image realignment preprocessing were included in the model as estimates of residual variability in the fMRI signal due to participant motion. Prior to estimation, the fMRI data and design matrix were high-pass filtered (cutoff = 128 s) to remove slow drifts in the signal due to scanner instability and slow physiological effects, and prewhitened to correct for autocorrelation in the data, modeled as a first-order autoregressive process [Friston et al., 2002]. The BOLD response of the task compared to rest state was modeled assuming the SPM canonical hemodynamic response function (HRF) and assessed using unpaired *t*-tests. Individual maps were first thresholded at $P < 0.001$ (uncorrected) and voxels considered significant if they belonged to a cluster size that survived a threshold of $P < 0.05$, corrected for topological false discovery rate (FDR) [Chumbley and Friston 2009].

Resultant statistical parametric maps are displayed in color and in radiological convention (left is participant's right), superimposed onto the mean of the unsmoothed spatially normalized EPI images. We also constructed pen-

etrance maps to summarize individual differences in activation missed by group averaging [Berl et al., 2006]. These display in color activated voxels where two or more participants were significant, corrected at topological FDR ($P < 0.05$). This provides a view of the spatial consistency of activation among participants [Fox et al., 1996].

Voxel-based group analyses were undertaken using mixed (random) effects analysis to permit valid population inference. Resultant maps were first thresholded at $P < 0.001$ (uncorrected) and voxels considered significant if they belonged to a cluster size that survived a threshold of $P < 0.05$ corrected for multiple comparisons (family-wise error) [Worsley et al., 1996]. To address hypothesis 1, that there would be overlap in the neural structures underpinning covert vocal singing and generative language, a whole group analysis of each task was first conducted. We then classified a voxel as overlapping if it was significant in both group maps. In addition, a mixed effects regional analysis was undertaken by counting, for each individual, the number of overlapping voxels activated for sing-rest and speech-rest (each $P < 0.001$ uncorrected) in particular regions of interest (ROI). These regions targeted frontal and temporo-parietal areas known to be involved in singing and language [Everts et al., 2010; Wood et al., 2001; Zarate and Zatorre, 2008] and were defined in accordance with our previously published methods [Sveller et al., 2006]. All regions were bilaterally assessed, including the middle and inferior frontal gyri, the planum temporale, and the angular gyri. In each hemisphere, the region with the smallest total count was contrasted with the remaining regions to identify intrahemispheric differences in overlap across participants using repeated measures ANOVA with planned orthogonal contrasts.

Hypothesis 2, that expert and nonexpert singers would use partially different singing networks, was assessed for sing-rest by contrasting activation of low and high pitch accuracy groups identified from the hierarchical cluster analysis to maximize differences in singing expertise. We also compared the number of activated voxels for sing-rest ($P < 0.001$ uncorrected) in the identified ROIs using *a posteriori* comparisons of the pitch accuracy groups, with Tukey's Honestly Significant Difference test to adjust for multiple comparisons or Tamhane's T2 where equal variance was not assumed. Where data violated the Kolmogorov-Smirnov test of normality we used the nonparametric Kruskal Wallis test.

Finally to test hypothesis 3, that overlap between the singing and language networks would vary with singing expertise, we classified a voxel as overlapping if it was significant in both the group difference map (low-high pitch accuracy) of sing-rest and in the whole group language map (speech-rest). We also quantified the number of overlapping voxels (each $P < 0.001$ uncorrected) from individual sing-rest and speech-rest maps in targeted ROIs. Similar to hypothesis 2, we used *a posteriori* comparisons of the pitch accuracy groups, as described above.

Talairach co-ordinates of significant regions of activation [Talairach and Tournoux, 1988] were obtained by

translating co-ordinates from MNI space to Talairach space using the transform of Lancaster et al. [2007]. This was implemented in *icbm2tal*, *BrainMap* *GingerALE* (version 2.0, Research Imaging Center, University of Texas Health Science Center, San Antonio) with labels for regions determined using the Talairach Client version 2.4.2 [Lancaster et al., 2000]. This approach allowed observed regions of activation to be related to previous results published in Talairach space, particularly relevant for interpreting the location of activation peaks in premotor cortex [Chen et al., 2008, 2009; Rizzolatti et al., 2002; Schubotz and von Cramon, 2003]. Because spatial normalization has limited accuracy, the gyral location of the activation peaks was also reviewed on the MR scans of individual participants by a neuroanatomical expert (author GDJ).

RESULTS

Behavioral Assessment of Vocal Singing and Language

The out-of-scanner pitch accuracy scores were significantly correlated with years of singing experience (Spearman's $\rho = 0.771$, $P < 0.001$) indicating that pitch accuracy provided a reliable marker of singing expertise. In particular, it supports its use as a more rigorous behavioral marker of out-of-scanner ability [Dalla Bella et al., 2007] that we then used to cluster the singers for group fMRI analyses.

The hierarchical cluster analysis of pitch accuracy scores revealed three distinct, empirically derived clusters of singers: (i) High pitch accuracy, whose average sung pitch varied from the target ratios within a range of 2–5%. (ii) Mid pitch accuracy, whose average sung pitch varied from the target ratios within a range of 5–7.5%. (iii) Low pitch accuracy, whose average sung pitch varied substantially (>10%) from target ratios (Table I).

The high pitch accuracy singers largely comprised expert singers with professional operatic careers or a history of regular public performances in theatrical productions such as musicals, including solo work. On average, these individuals had commenced vocal training at the age of 12.7 years ($SD = 7.1$), and had significantly more years of singing experience ($F(2,22) = 7.851$, $P < 0.01$) than the mid pitch accuracy (contrast estimate = -8.873 , $P < 0.01$), and low pitch accuracy groups (contrast estimate = -10.444 , $P = 0.001$; Table I). Two individuals in the mid pitch accuracy group had received formal vocal training (mean onset age = 15.5 ± 2.1 years) and had performed publicly in amateur musical productions or choirs, including solo work. The remaining individuals had no history of vocal training or performance but had naturally good pitch production ability. The low pitch accuracy group constituted nontuneful (nonexpert) singers, with no history of vocal training or musical performance.

The three groups did not differ in their report of exposure to singing in the home as a child, and there were no significant differences between the groups for sex, age,

years of education, or handedness ($P > 0.05$ for all comparisons; Table I). All groups rated the in-scanner song as equally familiar, including the tune ($F(2,20) = 0.463$, $P > 0.05$) and the lyrics ($F(2,20) = 2.233$, $P > 0.05$). Out-of-scanner assessment of generative language using the COWAT showed no significant difference between the groups ($F(2,19) = 0.794$, $P > 0.05$).

Cortical Overlap for Covert Vocal Singing and Generative Language

To assess hypothesis 1, mixed effects analysis of speech-rest showed significant regions of activation in the left medial frontal and precentral gyri (BA 6), the left precuneus (BA 31), left fusiform gyrus (BA 36), and right insula. Significant left subcortical activation was also observed in the claustrum, lentiform nucleus, and cerebellum (see Table II). For the mixed effects analysis of sing-rest, we partialled out singing expertise by including pitch accuracy scores as a covariate of no interest. The resultant statistical parametric map showed significant left-sided activation in the medial frontal gyrus (BA 6) and the superior temporal gyrus (BA 22; Table II), consistent with previously reported regions of activation during covert singing [Callan et al., 2006; Gunji et al., 2007; Kleber et al., 2007].

Figure 1A displays the overlap of the group mixed effects activation maps for sing-rest and speech-rest. A medial frontal region (left > right) was identified as common to both tasks in this analysis. Figure 1C depicts an example of the activation maps created for each individual to assess activation specific to speech-rest, sing-rest, and common to both tasks (see also Supporting Information Fig. 1). The individual maps reveal that in addition to the medial region of overlap, overlap often occurred in lateral regions; however, the location of the overlapping voxels varied considerably between participants. Group penetrance maps shown in Figure 1D summarize individual activation results by displaying voxels where two or more participants were significant (at topological FDR, $P < 0.05$) for both speech-rest and sing-rest within an individual. Across individuals, activation specific to speech-rest (Fig. 1D) was predominantly left-lateralized involving frontal and temporo-parietal language areas identified in our previous research and that of others [Wood et al., 2001]. Commonly activated regions specific to sing-rest lay proximal to language specific regions bilaterally (Fig. 1D). In particular, singing activation lay posterior to language activation in the frontal lobes, but more anterior to language activation in the temporal lobes. Cortical overlap occurred in regions where covert generative language and vocal singing were proximally located (Fig 1D).

Table III contains the results of the ROI analysis of the number of activated voxels common to sing-rest and speech-rest. This generally indicates greater cortical overlap in the left compared with the right hemisphere, and in anterior compared with posterior ROIs. The greatest cortical overlap was observed in the middle frontal gyrus,

TABLE II. Activation peaks for the covert singing and language tasks

Contrast	Activated region (Brodmann Area)	Talairach coordinates			P-value (SPM)
Speech-rest ^a					
All participants	Left medial frontal gyrus (BA 6)	−7	0	58	0.000
	Left precentral gyrus (BA 6)	−51	2	29	0.000
		−26	−16	55	0.004
	Left precuneus (BA 31)	−24	−72	26	0.001
	Left fusiform gyrus (BA 36)	−42	−37	−22	0.000
	Right insula	34	17	1	0.000
	Left claustrum	−32	13	3	0.000
	Left lentiform nucleus (putamen)	−18	−6	14	0.000
	Left cerebellum (tonsil)	−1	−44	−41	0.019
	−38	−46	−39	0.022	
Sing-rest					
All participants ^b	Left medial frontal gyrus (BA 6)	−4	−12	70	0.004
	Left superior temporal gyrus (BA 22)	−55	−50	15	0.021

The table shows the results of the cluster analysis of statistical parametric maps ($P < 0.05$ corrected for multiple comparisons). Talairach co-ordinates were converted from MNI co-ordinates using BrainMap GingerALE 2.0.

^aAll regions were also significant at a voxel level after correction for family-wise error ($P < 0.05$).

^bThe out-of-scanner pitch accuracy score was included in the contrast analysis as a covariate of no interest.

particularly on the left, while the least overlap was observed in the angular gyrus, most striking on the right. Considerable variability in overlap was evident across participants as indicated by large standard deviations for the total mean counts (see Table III). Intrahemispheric differences were also observed, particularly on the right, with repeated measures ANOVA indicating that across participants, the extent of overlap was greater for the right middle and inferior frontal gyri and the planum temporale relative to the right angular gyrus ($F(1.3,32.5) = 8.186$, $P = 0.004$). In contrast, only the left middle frontal gyrus showed greater overlap relative to the left angular gyrus across participants ($F(1.7,42.4) = 14.084$, $P < 0.001$; Table III). Taken together, the findings support the hypothesis of overlap in the neural structures underpinning covert vocal singing and generative language, with variability in the data pointing to differences associated with singing expertise.

Expertise and the Singing Network

To assess hypothesis 2, mixed effects analysis of sing-rest with a nonexpert-expert contrast (low-high pitch accuracy groups) showed a sole region of significant difference in the right middle frontal gyrus (rMFG), with a peak in mid premotor cortex (BA6; $x = 43$ mm, $y = 5$ mm, $z = 47$ mm, $P < 0.05$; Fig. 2A). The penetrance maps in Figure 2B reveal how individuals fared, with nonexperts showing greater bilateral frontal activation for sing-rest compared with experts. Of note, the mixed effects analysis of speech-rest for the low-high contrast showed no significant differences ($P > 0.5$ corrected and $P > 0.05$ uncorrected), as evident from the penetrance maps of speech-rest for the expert and nonexpert singers (Fig. 2B).

Figure 2C displays the mean voxel counts for sing-rest in the identified ROIs for the high, mid, and low pitch accuracy groups. Consistent with the findings of the mixed effects analysis, voxel counts were greater bilaterally in the middle frontal gyrus of nonexperts (Right: $F(2,23) = 6.438$, $P < 0.01$; Left: $F(2,23) = 3.467$, $P < 0.05$). This was significant on the right when compared with both the mid pitch accuracy (mean difference = 656.794, $P = 0.019$) and high pitch accuracy groups (mean difference = 654.722, $P = 0.01$), and significant on the left for the high (expert) compared with the low pitch accuracy (nonexpert) singers (mean difference = 613.056, $P = 0.045$). A significant effect was also observed in the right inferior frontal gyrus (Kruskal Wallis ($\chi^2(2) = 8.9$, $P = 0.012$), with mean ranks indicating a greater number of activated voxels in the low compared with the high and mid pitch accuracy groups. Combined, these findings support the hypothesis that expert and nonexpert singers use partially different singing networks. They are consistent with previous research showing involvement of the right frontal lobe during singing [Henson, 1985; Yamadori et al., 1997; Wilson et al., 2006], and indicate that this pattern is more typical of nonexperts.

Expertise and Patterns of Cortical Overlap

To assess hypothesis 3, the mixed effects group difference map (low-high pitch accuracy) of sing-rest and the whole group language map (speech-rest) were assessed; however, no overlapping significant voxels were found. The mixed effects group analysis of the sing-rest task failed to significantly activate the rMFG, despite individual maps (summarized in penetrance maps in Fig. 3A) suggesting that at least some individuals appeared to differ in an area of overlap in this region. Table III displays a

Figure 1.

TABLE III. ROI voxel counts for the covert singing and language tasks

Region of Interest (ROI)	Overlapping voxels for speech-rest and sing-rest for the pitch accuracy groups				P-value ^a
	High	Mid	Low	Total	
Right hemisphere					
Middle frontal gyrus	7.9 (20.3)*	48.6 (57.4)	202 (166.6)*	86 (132.3)	0.003
Inferior frontal gyrus	1 (2.3)	19.3 (37.4)	55.8 (68.5)	24.9 (49.2)	0.016
Planum temporale	14.1 (26.5)	39.4 (68.5)	19 (52.9)	22.6 (48.9)	0.028
Angular gyrus	0	0.6 (1.5)	0.4 (1.3)	0.3 (1.1)	
Left hemisphere					
Middle frontal gyrus	203.5 (120)	261 (271.6)	485.9 (502.5)	316.7 (346.4)	0.001
Inferior frontal gyrus	20.4 (24.4)	22.9 (36.2)	166.9 (202.6)	71.8 (136.6)	0.294
Planum temporale	15 (16.9)	45.7 (82.5)	83.2 (240.1)	46.9 (145.1)	0.926
Angular gyrus	28.5 (49.1)	41.3 (59.2)	63.3 (105.4)	44.0 (74.1)	

The table shows the results of the region of interest (ROI) analysis for the mean number of overlapping voxels ($\pm$ SD) for the covert singing and language tasks for all participants (total), and the high, mid, and low pitch accuracy groups.

^aSignificance value for the mean number of overlapping voxels collapsed across participants, analyzed separately for hemisphere and comparing each ROI with the mean voxel count of the angular gyrus.

* $P = 0.024$ for the high compared to the low pitch accuracy group.

general tendency for the number of overlapping voxels for sing-rest and speech-rest to be greater in the low compared with the mid and high pitch accuracy groups in the identified ROIs across individuals. This effect was significant in the rMFG ($F(2,23) = 8.989$, $P = 0.001$), with the nonexperts showing greater overlap than the experts (mean difference = 194.1, $P < 0.05$). While similar trends were evident in other regions, these failed to reach significance.

These findings point to the pivotal role of the rMFG, particularly the mid region of the premotor cortex (BA6) in covert vocal singing, that is shared with covert language especially in nonexpert singers. We have previously observed significant rMFG activity associated with a related verbal fluency task (noun-verb generation) also routinely used by our group to localise language in-scanner [Sveller et al., 2006]. To further probe the importance of this region we established another ROI, being the area of significant rMFG activity in response to a covert noun-verb generation task in 30 independent healthy controls that were previously established as left lateralized for language [Abbott

et al., 2010]. We examined the location of the significant nonexpert-expert difference in vocal singing in the present participants in relation to this rMFG noun-verb ROI.

The analysis of voxel counts in this ROI in individuals revealed a significantly greater number of overlapping voxels in the rMFG of nonexperts compared with experts for sing-rest and the whole group language map of noun-verb generation in the independent controls (Mann-Whitney $U = 13.0$, $z = -2.827$, $P = 0.005$; see Fig. 3B). Overlap of the mixed effect group analyses is shown in Figure 3C. It demonstrates that the location of the sole region of significantly increased activation observed on the difference map (low-high pitch accuracy) of sing-rest (see also Fig. 2A) is highly concordant with the rMFG noun-verb ROI. This highlights the importance of this region in both covert singing with words and generative language in nonexperts. Taken together, the findings support the hypothesis that the extent of cortical overlap for covert vocal singing and generative language varies according to individual differences in singing expertise, with partial overlap of

Figure 1.

Brain activity for covert vocal singing (sing-rest) and generative language (speech-rest) in all participants. **A:** Overlap of the group mixed effects activation maps for sing-rest and speech-rest indicating, in yellow, activated voxels in the medial frontal region (left > right) significant in both maps. Sagittal slice (0 mm). Coronal slice (0 mm). Axial slice (+72 mm). **B:** The three singing groups derived from cluster analysis of the out-of-scanner assessment of singing expertise based on pitch accuracy scores. Red diamond = high pitch accuracy (expert); Green triangle = mid pitch accuracy; Blue circle = low pitch accuracy (nonexpert). **C:** Individual imaging analysis of a participant in the nonexpert group showing activation specific to speech-rest only

(red), sing-rest only (yellow), and common to both tasks (white). Sagittal slices; left (+56 mm), medial (0 mm), right (−56 mm). Axial slices (+48 mm, +40 mm, +32 mm, +24 mm, +16 mm, +8 mm, 0 mm, −8 mm, −16 mm, −24 mm). **D:** Summary of individual activation results using group penetrance maps to indicate, in colour, voxels where two or more participants were significant (at topological FDR, $P < 0.05$) for both speech-rest and sing-rest within an individual (top panel) and for speech-rest only (middle panel) and sing-rest only (bottom panel). Sagittal slices; left (+56 mm), right (−56 mm). Axial slices (+48 mm, +16 mm, +8 mm, −24 mm).

Figure 2.

covert vocal singing and generative language particularly evident in nonexperts, whereas expert singers showed greater differentiation of the singing and language networks. This difference in overlap may partially account for the less tuneful performance of nonexpert singers.

DISCUSSION

By comparing patterns of cortical activation associated with covert vocal singing and generative language we have demonstrated that these higher cortical functions are supported by proximally located but interacting neural networks. We have further shown that regions underpinning covert vocal singing in nonexperts partially differ from experts, which in turn alters the interaction between the networks. In particular, the extent of overlap decreased with singing expertise, evident from decreased bilateral frontal activation and an associated reduction in cortical overlap with language-identified regions in the right frontal lobe [Abbott et al., 2010; Sveller et al., 2006; Wood et al., 2001]. In other words, expert vocal singing appears less dependent on the language network which when engaged, may produce less tuneful, nonexpert performance.

Our findings are consistent with previous neuroimaging studies that show singing is a bihemispheric task that is proximally located to language regions of the brain [Brown et al., 2004, 2006; Callan et al., 2006; Gunji et al., 2007; Özdemir et al., 2006]. We found bihemispheric singing activation in both expert and nonexpert singers in previously identified regions of the singing network, including the planum temporale (BA22), and premotor and supplementary motor areas (BA6) [Brown et al., 2004; Brown et al., 2006; Callan et al., 2006; Hickok et al., 2003; Jeffries et al., 2003; Kleber et al., 2007; Koelsch et al., 2009; Özdemir et al., 2006; Perry et al., 1999; Riecker et al., 2000; Saito et al., 2006; Wildgruber et al., 1996; Zarate and Zatorre, 2008]. We also found that prominent right frontal lobe involvement in singing is more typical of nonexperts, converging with early neuroimaging and lesion-based research as well as compelling accounts of fluent word production during singing in patients with aphasia [Helm-Estebrooks, 1983; Henson, 1985; Jeffries et al., 2003; Peretz

et al., 2004; Perry et al., 1999; Racette et al., 2006; Riecker et al., 2000; Schlaug et al., 2008; Wildgruber et al., 1996; Wilson et al., 2006; Yamadori et al., 1997].

Notably, both expert and nonexpert singers showed typical patterns of language activation and behavioral performance of the verbal fluency task [Mummery et al., 1996; Wallentin, 2009; Wood et al., 2001], suggesting normal language organization in both groups. Activation unique to language lay anterior to singing-specific activation in the frontal lobes but more posterior in the temporal lobes, with substantive overlap between regions. In other words, although vocal singing may serve distinct individual and social functions it shares features with language that are mediated by a partially shared neural network, consistent with findings of previous research [Brown et al., 2004; Brown et al., 2006; Callan et al., 2006; Hickok et al., 2003; Jeffries et al., 2003; Kleber et al., 2007; Koelsch et al., 2009; Özdemir et al., 2006; Platel et al., 2003].

The few previous neuroimaging studies comparing expert and nonexpert singers [Formby et al., 1989; Kleber et al., 2009; Zarate and Zatorre, 2008] have pointed to differences. Of relevance to our findings, Zarate and Zatorre [2008] suggested that expert singers rely less on dorsal premotor cortex than nonexperts for basic sensorimotor integration of audio-vocal feedback. Of note, they observed greater right than left-sided activation when participants were instructed to ignore pitch-shifted feedback while performing a simple singing task that was associated with poorer behavioral performance in nonexperts. Chen et al. [2008, 2009] noted bilateral activation in a proximal region they termed mid premotor cortex that is similar in location to the peak activation we observed in our nonexpert singers during sing-rest. Of note, this region was significantly recruited when participants passively listened to musical rhythms. More generally, increased activation of premotor cortex has been reported for performance of covert compared with overt tasks, including imagined singing and instrumental performance [Hickok et al., 2003; Kleber et al., 2007; Koelsch et al., 2009; Kristeva et al., 2003; Langheim et al., 2002; Lotze et al., 2003; Riecker et al., 2000], and during perceptual music imagery tasks [Halpern and Zatorre, 1999; Halpern et al., 2004; Zatorre et al., 1996]. Taken together, these findings

Figure 2.

Brain activity for covert vocal singing (sing-rest) and generative language (speech-rest) in expert and nonexpert singers. **A:** A comparison of the nonexpert and expert singers for sing-rest using mixed effects group analysis ($P < 0.05$ corrected for multiple comparisons). This shows the only area of significantly increased activity in nonexperts compared to experts, located in the right middle frontal gyrus (BA6; $x = 43$ mm, $y = 5$ mm, $z = 47$ mm). There were no areas where activity in experts was significantly higher than nonexperts. Axial slices ($+56$ mm, $+48$ mm, $+40$ mm, $+32$ mm, $+24$ mm). **(B)** Group penetrance maps for experts and nonexperts

showing voxels where two or more participants were significant (at topological FDR, $P < 0.05$) for sing-rest or speech-rest only. Nonexpert singers show greater bilateral frontal involvement for sing-rest compared with experts. Axial slices ($+48$ mm, $+40$ mm, $+8$ mm, 0 mm). **C:** Bar graph of the mean number of voxels (sized $2 \times 2 \times 2$ mm³) activated during sing-rest in each region of interest (ROI) for the three singing groups, shown separately for the left and right cerebral hemispheres. Red = high pitch accuracy (expert); Green = mid pitch accuracy; Blue = low pitch accuracy (nonexpert). L, left; R, right. Error bar = 95% confidence interval.

A Overlap of sing-rest and speech-rest **Figure 3.**

are consistent with our observation of increased activation of right premotor cortex in nonexperts compared with experts during covert singing.

It is well-established that activation of premotor and supplementary motor areas is not specific to covert singing, but has been associated with a range of tasks, notably the planning and initiation of motor actions [Chen et al., 2008, 2009; Chouinard and Paus, 2006; Rizzolatti et al., 2002; Schubotz and von Cramon, 2003; Zatorre et al., 2007] including speech production [Hickok and Poeppel, 2007; Shuster and Lemieux, 2005]. Language neuroimaging studies have shown bilateral involvement of premotor and supplementary motor areas during planning and initiation of serial word production, phonological rehearsal, generating inner speech, verbal fluency, and auditory-verbal imagery tasks [Bohland and Guenther, 2006; Hickok et al., 2003; Koelsch et al., 2009; Mummery et al., 1996; Richardson and Price, 2009; Shergill et al., 2001; Shuster and Lemieux, 2005]. Combined with our findings, this work points to a shared region for audio-vocal integration that may be routinely engaged in generating vocal singing and language. In keeping with this, the premotor and supplementary motor areas are typically implicated in integrating motor and music-auditory maps during music performance [Chen et al., 2008, 2009; Chouinard and Paus, 2006; Koelsch et al., 2009; Kristeva et al., 2003; Langheim et al., 2002; Saito et al., 2006; Zatorre et al., 2007]. In the case of our expert singers, decreased reliance on this shared audio-vocal interface may reflect a more refined [Hund-Georgiadis and von Cramon, 1999; Jänke et al., 2000; Kleber et al., 2009; Krings et al., 2000; Langheim et al., 2002; Lotze et al., 2003; Münte et al., 2002; Zarate and Zatorre, 2008] and less “language-driven” network compared with nonexperts, with decoupling producing more tuneful, skilled vocal singing performance.

Limitations

The tasks used in this study successfully activated higher cortical regions that typically support familiar vocal

singing and generative language in everyday life. Vocal singing, however, produces a potential source of variability relating to task difficulty. Because most of us practice our use of language everyday, we assume more homogeneity through proficiency and skill in contrast to tuneful vocal singing which some of us practice rarely! This issue of proficiency was overcome by examining individuals with both expert and nonexpert singing ability. Dalla Bella et al. [2007] noted a lack of consensus for obtaining objective measures of singing proficiency for melodies, and promoted acoustic-based analyses of sung performance over traditional measures derived from expert ratings. While they found that nonexpert singers were more accurate for well-known tunes than isolated pitches, they showed differences between experts and nonexperts for the majority of pitch variables they measured. In our study we used pitch accuracy as a marker of singing proficiency reflecting the conclusion that pitch accuracy (particularly with regulated tempo) produces an optimal estimate of singing ability in the normal population [Dalla Bella et al., 2007]. Our findings point to the potential benefits of examining the extent to which regular singing practice might alter pitch accuracy and associated functional activation patterns in individuals of varying singing ability.

Brown et al. [2006] showed similar regions of activation to those reported in our study for novel sentence and melody generation tasks in amateur musicians. What remains to be shown, however, is the extent to which language generativity might alter the interaction between vocal singing and language networks relative to expertise. Sex differences might also play a role, given the assumption that females show more bilateral patterns of language activation than males, and perform superiorly on tasks like verbal fluency [Wallentin, 2009]. Wallentin [2009], however, has recently challenged the evidence base of this assumption, and concluded that sex should not be considered a confounding factor in language neuroimaging studies. Of the few studies investigating sex differences in pitch abilities, minimal effects have been shown [Brown et al., 1999; McRoberts and Sanders, 1992].

Figure 3.

Cortical overlap for covert vocal singing (sing-rest) and generative language (speech-rest) relative to singing expertise. **A:** Group penetration maps for experts and nonexperts showing voxels where two or more participants were significant (at topological FDR, $P < 0.05$) for both speech-rest and sing-rest within an individual. Nonexperts show increased cortical overlap, particularly in the right frontal lobe. Sagittal slices; left (+56 mm), right (−56 mm). Axial slices (+48 mm, +40 mm, +8 mm, 0 mm). **B:** Dotplot of the number of voxels (sized $2 \times 2 \times 2 \text{ mm}^3$) activated for sing-rest in the right middle frontal gyrus of experts and nonexperts that were also significant in a group mixed effects analysis of

noun-verb generation in 30 independent left-lateralized healthy controls [Abbott et al., 2010]. Nonexperts show significantly greater overlap than experts. **C:** Region of Interest (ROI) analysis in the right middle frontal gyrus showing activation due to noun-verb generation (red and white) in 30 independent left-lateralised healthy controls, and a region where activity specific to sing-rest in nonexperts is greater than expert singers (yellow and white); the portion common to both is shown in white. This analysis highlights the role of the right middle frontal gyrus in verbal fluency and, depending upon expertise, singing. Axial slices (+56 mm, +48 mm, +40 mm, +32 mm, +24 mm).

CONCLUSIONS

This study found greater activation of the right frontal homologue of language in nonexperts during vocal singing, with increased expertise associated with less extensive activation in the right hemisphere and less overlap with traditional language areas. These findings suggest that expert singers have less right frontal involvement, more left lateralized activation, and less overlap with generative language. This means that taking account of expertise in singing is important as it may contribute to variability in previous research findings.

Brown et al. [2004] speculated that the song system of the human brain evolved from a more primitive system supporting vocal imitation [see also Mithen, 2009]. We can further speculate that the evolution of language may represent an expansion of the singing network, as based on its proximal but more anterior frontal representation, with language now dominating singing behavior unless practice or training is undertaken. In other words, these findings may reflect an aspect of the evolutionary design of human cognition and behavior: for humans the language system dominates, and to sing well, training may assist by facilitating modularization [McMullen and Saffran, 2004; Peretz, 2009]. The age-old singing adage of ‘finding your singing voice’ through regular practice and training may be neurologically mediated by changing how strongly singing is coupled to the language system, with decoupling producing more tuneful, expert performance.

ACKNOWLEDGMENTS

The authors sincerely thank R. S. Briellmann, A. B. Waites, M. Harvey, R. Bos, L. Hassaram and L. Bird for their help with manuscript preparation, data collection and analysis. They would also like to acknowledge M.M. Saling for development of the orthographically cued lexical retrieval task in our group, and J. A. Ogden for helpful early discussions and support.

REFERENCES

- Abbott D, Jackson G (2001): iBrain® Software for analysis and visualisation of functional MR images. *NeuroImage* 13:S59.
- Abbott DF, Waites AB, Lillywhite LM, Jackson GD (2010): fMRI assessment of language lateralization: An objective approach. *NeuroImage* 50:1446–1455.
- Berl MM, Vaidya CJ, Gaillard WD (2006): Functional imaging of developmental and adaptive changes in neurocognition. *NeuroImage* 30:679–691.
- Bohland JW, Guenther FH (2006): An fMRI investigation of syllable sequence production. *NeuroImage* 32:821–841.
- Brown CP, Fitch RH, Tallal P (1999): Sex and hemispheric differences for rapid auditory processing in normal adults. *Lateral* 4:39–50.
- Brown S, Martinez MJ, Hodges DA, Fox PT, Parsons LM (2004): The song system of the human brain. *Cogn Brain Res* 20:363–375.
- Brown S, Martinez MJ, Parsons LM (2006): Music and language side by side in the brain: A PET study of the generation of melodies and sentences. *Eur J Neurosci* 23:2791–2803.
- Callan DE, Tsytsarev V, Hanakawa T, Callan AM, Katsuhara M, Fukuyama H, Turner R (2006): Song and speech: Brain regions involved with perception and covert production. *NeuroImage* 31:1327–1342.
- Chen JL, Penhune VB, Zatorre RJ (2008): Listening to musical rhythms recruits motor regions of the brain. *Cereb Cortex* 18:2844–2854.
- Chen JL, Penhune VB, Zatorre RJ (2009): The role of auditory and premotor cortex in sensorimotor transformations. In: Dalla Bella, S, Kraus N, Overy K, Pantev C, Snyder JS, Tervaniemi M, Tillmann B, Schlaug G, editors. *The Neurosciences and Music III - Disorders and Plasticity*. Ann NY Acad Sci 1169:15–34.
- Chouinard PA, Paus T (2006): The primary motor and premotor areas of the human cerebral cortex. *The Neurosci* 12:143–152.
- Chumbley JR, Friston KJ (2009): False discovery rate revisited: FDR and topological inference using Gaussian random fields. *NeuroImage* 44:62–70.
- Dalla Bella S, Giguère J, Peretz I (2007): Singing proficiency in the general population. *J Acoust Soc Am* 121:1182–1189.
- Ericsson KA (1997): Deliberate practice and the acquisition of expert performance: An overview. In: Jorgensen H, Lehmann AC, editors. *Does Practice Make Perfect?* Norway: Norges Musikkhøgskole.
- Formby C, Thomas RG, Halsey JH (1989): Regional cerebral blood flow for singers and nonsingers while speaking, singing, and humming a rote passage. *Brain Lang* 36:690–698.
- Fox PT, Ingham RJ, Ingham JC, Hirsch TB, Downs JH, Martin C, Jerabek P, Glass T, Lancaster JL (1996): A PET study of the neural systems of stuttering. *Nature* 382:158–161.
- Friston KJ, Glaser DE, Henson RN, Kiebel S, Phillips C, Ashburner J (2002): Classical and Bayesian inference in neuroimaging: Applications. *NeuroImage* 16:484–512.
- Gunji A, Ishii R, Chau W, Kakigi R, Pantev C (2007): Rhythmic brain activities related to singing in humans. *NeuroImage* 34:426–434.
- Halpern AR, Zatorre RJ (1999): When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cereb Cortex* 9:697–704.
- Halpern AR, Zatorre RJ, Bouffard M, Johnson JA (2004): Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia* 42:1281–1292.
- Haslinger B, Erhard P, Altenmüller E, Hennenlotter A, Schwaiger M, Gräfin von Einsiedel H, Rummeny E, Conrad B, Ceballos-Baumann AO (2004): Reduced recruitment of motor association areas during bimanual coordination in concert pianists. *Hum Brain Mapp* 22:206–215.
- Helm-Estebrooks N (1983): Exploiting the right hemisphere for language rehabilitation: Melodic intonation therapy. In: Perelman E, editor. *Cognitive Processing in the Right Hemisphere*. New York: Academic Press.
- Henson RA (1985): Amusia. In: Frederiks JAM, editor. *Handbook of Clinical Neurology*. New York: Elsevier. pp 483–490.
- Hickok G, Buchsbaum B, Humphries C, Muftuler T (2003): Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area SPT. *J Cog Neurosci* 15:673–682.
- Hickok G, Poeppel D (2007): The cortical organization of speech processing. *Nature Rev Neurosci* 8:393–402.
- Hund-Georgiadis M, von Cramon DY (1999): Motor learning related changes in piano players and non-musicians revealed

- by functional magnetic-resonance imaging. *Exp Brain Res* 125:417–425.
- Jäncke L, Shah NJ, Peters M (2000): Cortical activation in primary and secondary motor areas for complex bimanual movements in professional pianists. *Cogn Brain Res* 10:177–183.
- Jeffries KJ, Fritz JB, Braun AR (2003): Words in melody: An $H_2^{15}O$ PET study of brain activation during singing and speaking. *NeuroReport* 14:749–754.
- Kleber B, Birbaumer N, Veit R, Trevorrow T, Lotze M (2007): Overt and imagined singing of an Italian aria. *NeuroImage* 36:889–900.
- Kleber B, Veit R, Birbaumer N, Gruzelier J, Lotze M (2010): The brain of opera singers: Experience-dependent changes in functional activation. *Cereb Cortex* 20:1144–1152.
- Koelsch S, Schulze K, Sammler D, Fritz T, Müller K, Gruber O (2009): Functional architecture of verbal and tonal working memory: An fMRI study. *Hum Brain Mapp* 30:859–873.
- Krings T, Toepper R, Foltys H, Erberich S, Sparing R, Willmes K, Thron A (2000): Cortical activation patterns during complex motor tasks in piano players and control subjects. A functional magnetic resonance imaging study. *Neurosci Lett* 278:189–193.
- Kristeva R, Chakarov V, Schulte-Monting J, Spreer J (2003): Activation of cortical areas in music execution and imagining: A high-resolution EEG study. *NeuroImage* 20:1872–1883.
- Lancaster JL, Tordesillas-Gutierrez D, Martinez M, Salinas F, Evans A, Zilles K, Mazziotta JC, Fox PT (2007): Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Hum Brain Mapp* 28:1194–1205.
- Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov PV, Nickerson D, Mikiten SA, Fox PT (2000): Automated Talairach atlas labels for functional brain mapping. *Hum Brain Mapp* 10:120–131.
- Langheim F, Callicott JH, Mattay VS, Duyn JH, Weinberger DR (2002): Cortical systems associated with covert music rehearsal. *NeuroImage* 16:901–908.
- Lotze M, Scheler G, Tan H-RM, Braun C, Birbaumer N (2003): The musician's brain: Functional imaging of amateurs and professionals during performance and imagery. *NeuroImage* 20:1817–1829.
- McDermott J (2008): The evolution of music. *Nature* 453:287–288.
- McMullen E, Saffran JR (2004): Music and language: A developmental comparison. *Music Percept* 21:289–311.
- McRoberts GW, Sanders B (1992): Sex differences in performance and hemispheric organization for a nonverbal auditory task. *Percept Psychophys* 51:118–122.
- Miller G (2003): Singing in the brain. *Science* 299:646–648.
- Mithen S (2009): The music instinct: The evolutionary basis of musicality. In: Dalla Bella, S, Kraus N, Overy K, Pantev C, Snyder JS, Tervaniemi M, Tillmann B, Schlaug G, editors. *The Neurosciences and Music III - Disorders and Plasticity*. Ann NY Acad Sci 1169:3–12.
- Mummary CJ, Patterson K, Hodges JR, Wise RJS (1996): Generating 'Tiger' as an animal name or a word beginning with T: Differences in brain activation. *Proc R Soc Lond B* 263:989–995.
- Münste TF, Altenmüller E, Jäncke L (2002): The musician's brain as a model of neuroplasticity. *Nature Rev Neurosci* 3:473–478.
- Özdemir E, Norton A, Schlaug G (2006): Shared and distinct neural correlates of singing and speaking. *NeuroImage* 33:628–635.
- Ogawa S, Lee TM, Kay AR, Tank DW (1990): Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci USA* 87:9868–9872.
- Oldfield RC (1971): The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 9:97–113.
- Peretz I (2009): Music, language and modularity framed in action. *Psychol Belgica* 49:157–175.
- Peretz I, Gagnon L, Hébert S, Macoir J (2004): Singing in the brain: Insights from cognitive neuropsychology. *Mus Percept* 21:373–390.
- Perry DW, Zatorre RJ, Petrides M, Alivisatos B, Meyer E, Evans AC (1999): Localization of cerebral activity during simple singing. *NeuroReport* 10:3979–3984.
- Platel H, Baron JC, Desgranges B, Bernard F, Eustache F (2003): Semantic and episodic memory of music are subserved by distinct neural networks. *NeuroImage* 20:244–256.
- Racette A, Bard C, Peretz I (2006): Making non-fluent aphasics speak: Sing along! *Brain* 129:2571–2584.
- Richardson FM, Price CJ (2009): Structural MRI studies of language function in the undamaged brain. *Brain Struct Funct* 213:511–523.
- Riecker A, Ackermann H, Wildgruber D, Dogil G, Grodd W (2000): Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *NeuroReport* 11:1997–2000.
- Rizzolatti G, Fogassi L, Gallese V (2002): Motor and cognitive functions of the ventral premotor cortex. *Curr Opin Neurobiol* 12:149–154.
- Saito Y, Ishii K, Yagi K, Tatsumi IF, Mizusawa H (2006): Cerebral networks for spontaneous and synchronized singing and speaking. *NeuroReport* 17:1893–1897.
- Schlaug G, Marchina S, Norton A (2008): From singing to speaking: Why singing may lead to recovery of expressive language function in patients with Broca's aphasia. *Mus Percept* 25:315–323.
- Schön D, Gordon RL, Besson M (2005): Musical and linguistic processing in song perception. *Ann NY Acad Sci* 1060:71–81.
- Schubotz RI, von Cramon DY (2003): Functional-anatomical concepts of human premotor cortex: Evidence from fMRI and PET studies. *NeuroImage* 20:S120–S131.
- Shergill SS, Bullmore ET, Brammer MJ, Williams SCR, Murray RM, McGuire PK (2001): A functional study of auditory verbal imagery. *Psychol Med* 31:241–253.
- Shuster LI, Lemieux SK (2005): An fMRI investigation of covertly and overtly produced mono and multisyllabic words. *Brain Lang* 93:20–31.
- Sveller C, Briellmann RS, Saling MM, Lillywhite L, Abbott DF, Masterton RAJ, Jackson GD (2006): Relationship between language lateralization and handedness in left-hemispheric partial epilepsy. *Neurol* 67:1813–1817.
- Talairach J, Tournoux P (1988): *Co-Planar Stereotaxic Atlas of the Human Brain*. New York: Thieme Medical.
- Trehub SE (2001): Musical predispositions in infancy. *Ann NY Acad Sci* 930:1–16.
- Trehub SE (2003): The developmental origins of musicality. *Nat Neurosci* 6:669–673.
- Wallin N, Merker B, Brown S, editors (2000): *The Origins of Music*. Cambridge: MIT Press.
- Wallentin M (2009): Putative sex differences in verbal abilities and language cortex: A critical review. *Brain Lang* 108:175–183.
- Ward JH (1963): Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 58:236–244.
- Wildgruber D, Ackermann H, Klose U, Kardatzki B, Grodd W (1996): Functional lateralization of speech production at primary motor cortex: A fMRI study. *NeuroReport* 7:2791–2795.

- Wilson SJ, Pressing J, Wales RJ, Pattison P (1999): Cognitive models of music psychology and the lateralisation of musical function within the brain. *Aust J Psychol* 5:125–139.
- Wilson SJ, Parsons K, Reutens DC (2006): Preserved singing in aphasia: A case study of the efficacy of Melodic Intonation Therapy. *Mus Percept* 24:23–36.
- Wood AG, Saling MM, Abbott DF, Jackson GD (2001): A neurocognitive account of frontal lobe involvement in orthographic lexical retrieval: Preliminary insights with fMRI. *NeuroImage* 14:162–169.
- Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC (1996): A unified statistical approach for determining significant signals in images of cerebral activation. *Hum Brain Mapp* 4:58–73.
- Yamadori A, Osumi Y, Masuhara S, Okubo M (1997): Preservation of singing in Broca's aphasia. *J Neurol Neurosurg Psychiatry* 40:221–224.
- Zarate JM, Zatorre RJ (2008): Experience-dependent neural substrates involved in vocal pitch regulation during singing. *NeuroImage* 40:1871–1887.
- 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.
- Zatorre RJ, Halpern AR (2005): Mental concerts: Musical imagery and auditory cortex. *Neuron* 47:9–12.
- Zatorre RJ, Halpern AR, Perry DW, Meyer E, Evans EC (1996): Hearing in the mind's ear: A PET investigation of musical imagery and perception. *J Cogn Neurosci* 8:29–46.