Restorative Neurology and Neuroscience 25 (2007) 323–334
IOS Press

Congenital amusia: An auditory-motor feedback disorder?

Jake Mandell, Katrin Schulze and Gottfried Schlaug*
Department of Neurology, Music and Neuroimaging Laboratory, Beth Israel Deaconess Medical Center and Harvard Medical School, 330 Brookline Avenue, Boston, MA 02215, USA

Abstract. *Purpose:* Congenital amusia (*tone deafness*) is a disorder in which those affected typically complain of or are identified by their inability to sing in tune. A psychophysical and possibly surrogate marker of this condition is the inability to recognize deviations in pitch that are one semitone (100 cents) or less. The aim of our study was to identify candidate brain regions that might be associated with this disorder.
Methods: We used Voxel-Based-Morphometry (VBM) to correlate performance on a commonly used assessment tool, the Montreal Battery for the Evaluation of Amusia (MBEA), with local inter-individual variations in gray matter volumes across a large group of individuals ($n = 51$) to identify brain regions potentially involved in the expression of this disorder.
Results: The analysis across the entire brain space revealed significant covariations between performance on the MBEA and inter-individual gray matter volume variations in the left superior temporal sulcus (BA 22) and the left inferior frontal gyrus (BA 47). The regression analyses identified subregions within the inferior frontal gyrus, and inferior portion of BA47 that correlated with performance on melodic subtests, while gray matter volume variations in a more superior subregion of BA47 correlated with performance on rhythmic subtests.
Conclusions: Our analyses demonstrate the existence of a left fronto-temporal network that appears to be involved in the melodic and rhythmic discrimination skills measured by the MBEA battery. These regions could also be part of a network that enable subjects to map motor actions to sounds including a feedback loop that allows for correction of motor actions (i.e., singing) based on perceptual feedback. Thus, it is conceivable that individuals with congenital amusia, or the inability to sing in tune, may actually have an impairment of the auditory-motor feedback loop and/or auditory-motor mapping system.

Keywords: Congenital amusia, tone deafness, voxel-based-morphometry (VBM), BA 22, BA 47, auditory-motor mapping, auditory-motor feedback loop

1. Introduction

Congenital amusia (CA), commonly known as *tone-deafness*, is defined as a developmental disorder affecting the perception and production of music in otherwise normal-functioning individuals (Ayotte et al., 2002; Peretz & Hyde, 2003). By definition, congenital amusia is not attributable to a lack of musical training, a macroscopically identifiable brain lesion (which differentiates congenital amusia from acquired amusia), low IQ or level of education, hearing impairment, or neurological/psychiatric disorder. It is estimated that approximately 4% of the general population may have this disorder (Kalmus & Fry, 1980), although it is not clear whether this group of individuals simply represent the lower extremes of an otherwise normal distribution, or comprise a distinct population that clearly differs from a normal population without any transition. It has been argued that individuals with congenital amusia may have been born with either insufficient or impaired neural correlates for the perception and/or production of certain aspects of music (Peretz et al., 2002), although the nature, location, and extent of the underlying neural correlates have not been de-

*Corresponding author: Gottfried Schlaug, MD, PhD, Department of Neurology, Music and Neuroimaging Laboratory, Palmer 127, Beth Israel Deaconess Medical Center and Harvard Medical School, 330 Brookline Avenue, Boston, Massachusetts 02215, USA. E-mail: gschlaug@bidmc.harvard.edu.

termined. Individuals with congenital amusia are typically identified by, or complain of an inability to sing in tune, and various psychophysical experiments have determined that these individuals also have an inability to detect pitch deviations of one semitone or less (Ayotte et al., 2002; Peretz et al., 2002; Hyde & Peretz, 2004). However, in what way this perceptual inability contributes to, is part of or poses as a surrogate marker for this disorder is not yet known since congenital amusics typically complain of their inability to sing in tune but not of their inability to discriminate between two tones that are very close in pitch height. Thus, congenital amusia may not be characterized a perceptual discrimination problem solely but by the more obvious production problem (singing in tune) or the ability to make a correction in the production based on auditory feedback which points to an auditory-motor integration or auditory-motor feedback loop problem as the possible underlying functional abnormality in congenital amusia.

Since no macroscopically visible lesions have been described in the brains of individuals presumed to have congenital amusia, it is possible that the neural abnormality, if it exists, is so subtle that it may not be detectable by standard visual inspection of brain images, but instead, requires more sophisticated computational methods to visualize an underlying microscopic abnormality. Such subtle abnormalities could be due to focal neuronal migration disorders, a regional neuronal dysfunction, or a regional disconnection syndrome (e.g., impaired auditory cortex connections to motor related regions in the frontal lobe). There have been speculations in the literature regarding possible candidate brain regions for such a disorder. Kleist reported a case with a lesion in the left superior portion of the temporal lobe, posterior to Heschl's gyrus that had characteristics of tone deafness (Kleist, 1959). The involvement of auditory association cortex would also be in agreement with a recent evoked potential study (Peretz et al., 2005) in which it was shown that amusic subjects had an enhanced response to large changes in pitch by eliciting an N2-P3 complex that was twice that seen in normal subjects. Since the N1 response was similar in amusic and normal subjects, it was assumed that the underlying neural abnormality might not involve primary or early secondary auditory cortex, but instead was more likely to be found in higher order auditory association cortex. An N1 response is typically mapped to early secondary auditory association cortex (e.g., planum temporale). Enhanced N1/P2 responses have been seen when subjects were instructed to discriminated complex instrumental tones (compared to the discrimination of simple sine wave tones) (Meyer et al., 2006).

The Montreal Battery for the Evaluation of Amusia (MBEA) was developed and standardized to identify subjects with congenital amusia (Peretz et al., 2002; Ayotte et al., 2002). The first three subtests of the MBEA assess melodic discrimination ability and the next two assess rhythmic discrimination ability. The diagnostic criteria for congenital amusia are still in flux, in particular, the cut-off levels that determine what is clearly abnormal, what constitutes a borderline performance, and what is normal have varied slightly over the years (Ayotte et al., 2002; Peretz et al., 2002; Peretz et al., 2003). In addition, pitch and rhythm processing may not be affected in the same way by this disorder. Furthermore, the normalized distribution of performance on the Montreal Battery (Peretz et al., 2002) suggests that there may be a range of severity of impairment in both melodic and rhythmic tasks.

In order to ascertain the neural correlates of congenital amusia, we used an analysis technique called voxel-based-morphometry (VBM) that allows whole-brain analysis without requiring the delineation of predetermined regions of interest (Ashburner & Friston, 2000). VBM studies have typically been used to examine covariations or changes in gray matter volume and/or density either between groups or within groups over time (Ashburner & Friston, 2000; Maguire et al., 2000; Sluming et al., 2002; Watkins et al., 2002; Gaser & Schlaug, 2003). Furthermore, we showed in one study that VBM findings were similar to those of region-based morphometric studies, which cross-validates the VBM methods (Luders et al., 2004). Although VBM studies examining gray matter volume or density have been numerous in the past few years, VBM studies focused on white matter differences are rare, mostly because signal intensity differences seen in white matter either between groups or within subjects over time, are not very pronounced, and thus, making it more difficult to find VBM effects in white matter (Ashburner & Friston, 2000). Nevertheless, recently Hyde et al. (2006) reported white matter differences comparing a group of middle-aged (mean age = mid-fifties) amusic subjects with a group of normal controls. These between-group differences not only mapped to the white matter of the right inferior frontal gyrus, but also uncovered correlations between the inter-individual white matter signal intensity and performance on a pitch-based task.

The aim of our study was to determine the neural correlates of congenital amusia using a voxel-based morphometric technique. Assuming that subjects with

congenital amusia represent the lower extremes of an otherwise normal distribution, we examined covariations between performance on a musical assessment test (MBEA) and inter-individual variations in gray matter volume on a voxel-by-voxel basis across the entire brain space. Our subjects consisted of a large number of young individuals with varying levels of performance on the MBEA. Gray matter analysis was used, since previous studies have shown that VBM is particularly sensitive for detecting inter-individual variations in gray matter density and volume. Our overall aim was to identify candidate brain regions that are related to the phenotypic expression of congenital amusia. These brain regions could then become the basis of further exploration to examine their precise role in the expression of this disorder.

2. Subjects and methods

2.1. Subject recruitment and profiles

The study group consisted of 51 healthy, right-handed individuals who either responded to a newspaper advertisement asking for volunteers for a study on tone-deafness ($n = 37$) or were recruited as normal controls ($n = 14$) for other VBM studies in our laboratory that were going on at the same time. In all, 30 females and 21 males with a mean age of 25.5 (SD 4.6; age range: 18–40) were included in the analysis. All volunteers gave signed, informed consent and the study was approved by the Institutional Review Board of Beth Israel Deaconess Medical Center, Boston, MA.

2.2. Behavioral testing

All subjects were screened for neurological and psychiatric disorders before being enrolled, and subsequently underwent the Shipley/Hartford vocabulary and abstraction tests (Shipley, 1940; this test correlates highly with the Wechsler Adult Intelligence Scale full-scale IQ (Paulson & Lin, 1970)), standard audiometric testing, and subtests of the Montreal Battery of Evaluation of Amusia (MBEA). No significant differences were found in two-sample t-tests comparing amusics (using a criterion of 2SD below the mean MBEA score as a cutoff for amusia) to normal controls with respect to age, Shipley abstract and verbal scores, years of education, and years of playing a musical instrument.

Table 1
Profile of Subjects' MBEA total scores (average of first five subtests), number of subjects within each group, gender distribution, and mean ages (SD)

	N	MBEA Total %	Age
All Subjects	**51**	**82.3% (8.6)**	**25.5 (4.6)**
Males	21	82.6% (9.1)	27.0 (6.0)
Females	30	82.0% (8.4)	24.5 (3.1)
Amusic*	**13**	**71.7% (6.1)**	**24.5 (4.7)**
Males	6	71.9% (7.6)	25.7 (6.4)
Females	7	71.5% (5.2)	23.4 (2.5)
Non-Amusic	**38**	**86.4% (5.6)**	**26.1 (4.7)**
Males	15	86.9% (5.5)	27.5 (5.9)
Females	23	86.1% (5.9)	25.0 (3.3)

*The cutoff for Amusia is defined as 2 standard deviations below the mean, with 10 local controls determining the mean. Using this method, the amusic cutoff is 76.7% based on the average of the first five subtests of the MBEA.

2.3. MRI image acquisition and data analysis

A high-resolution (voxel size: 1 mm^3), strongly T1-weighted MR data set was acquired for each subject on a 1.5T Siemens Vision MR scanner (Erlangen, Germany). In addition, each subject underwent routine T2-weighted and Proton-density (PD)-weighted imaging to rule out the possibility of acquired lesions being the cause of amusia. None of our subjects had any obvious lesions on the T2 or PD images. Image pre-processing and VBM analyses were performed on a Linux workstation using MATLAB 6.0 (Mathworks Inc., Natick, MA, USA) and SPM2 (Wellcome Department of Cognitive Neurology, London, UK). Additional image viewing and Region of Interest (ROI) creation was performed in MRIcro (http://people.cas.sc.edu/rorden/). Further statistical analyses were done in Graphpad Prism (http://www.graphpad.com/).

2.4. Image preprocessing: template creation and segmentation

All image preprocessing and voxel-by-voxel statistical analyses were performed using the built-in functions of SPM2. Preprocessing of the data involved spatial normalization, segmentation, modulation and spatial smoothing with a 12 mm Gaussian kernel (Ashburner & Friston, 2000; Good et al., 2001). Customized gray matter, white matter, and CSF templates were created from the group of subjects in order to reduce scanner-specific bias. To facilitate optimal segmentation, we estimated normalization parameters while removing non-brain voxels (skull, sinus) using an optimized protocol (Good et al., 2001). The optimized parameters, estimated while normalizing extracted GM images to the

Table 2
Summary of the two subgroups that were used for the two-sample t-test comparisons in the VBM analyses

t-test	Normal Group	Amusic Group	Normal Cutoff %	Amusic Cutoff %
Melodic Average t-test	$n = 11$	$n = 16$	$>= 89\%$	$<= 76\%$
Rhythmic Average t-test	$n = 16$	$n = 12$	$>= 93\%$	$<= 75\%$
Total Score t-test	$n = 10$	$n = 13$	$>= 91\%$	$<= 77\%$

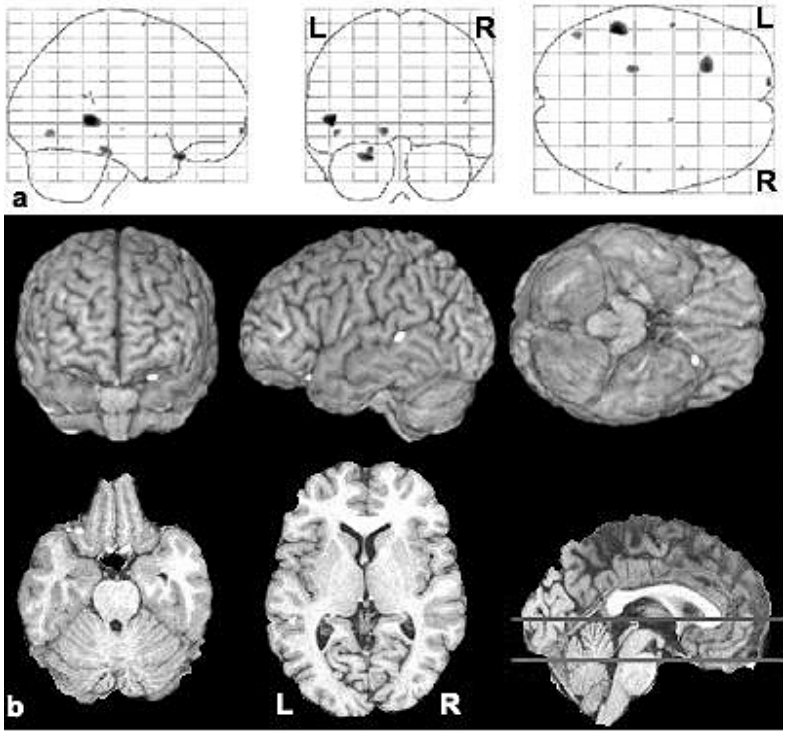

Fig. 1. Map generated by the regression analysis between all subjects's melodic subtests and the individual gray matter concentrations ($P < 0.005$ uncorrected). The results of this regression analysis were converted to a binary mask using the SPM's ImCalc function. Figure 1b: Significant group differences in gray matter concentration between the *true amusic subgroup* and the *normal control group* (p< 0.05, FWE corrected) after applying the binary mask from Fig. 1a overlaid onto the surface reconstruction of a single spatially standardized brain. Positions of the two axial slices in the bottom row are marked with red in a midsagittal slice (L = left hemisphere; R = right hemisphere).

customized GM template, were reapplied to the original whole brain images. All images were spatially normalized with the stereotactic space defined by the Montreal Neurological Institute (MNI) using a 12-parameter affine transformation, corrected for non-uniformities in signal intensity, and then partitioned into gray matter (GM), white matter (WM), cerebrospinal fluid (CSF), and background using a modified mixture model cluster analysis. In addition, we performed a correction for volume changes (modulation) by modulating each voxel by the Jacobian determinants derived from the spatial normalization, allowing us to also test for regional gray matter volume differences (Ashburner & Friston, 2000; Good et al., 2001). Only the smoothed gray matter images were used in the statistical analyses.

2.5. *VBM statistical analyses*

Three simple regression analyses were performed across the entire gray matter space regressing the 51 preprocessed gray matter images with their corresponding MBEA scores using an average of the melodic subtests (subtests #1–3), an average of the rhythmic subtests (subtests #4–5), as well as the total score across these 5 subtests on a voxel-by-voxel basis. The results of these regression analyses (at $p < 0.005$ uncorrected) were converted to a binary mask using SPM's ImCalc function. Next, we compared two subsets of subjects with each other using the binary maps of the 3 regression analyses as masks to restrict the analysis to only those brain regions included in the mask. We opted to choose a more liberal, uncorrected threshold to cre-

Table 3
Summary of the regions found as significant (FEW correct) in all VBM analyses

ROI Found in which t-test?	Where in Brain?	Local maxima: MNI (SPM) Coordinates	Corrected ($p < 0.05$; FWE) cluster size
Melodic	Left Superior Temporal Sulcus (BA 22)	−48, −40, 3	177 voxels
Melodic	Left Inferior Frontal (BA 47)	−24, 24, −26	129 voxels
Rhythmic	Left Superior Temporal Sulcus (BA 22)	−49, −42, 4	101 voxels
Rhythmic	Left Inferior Frontal (BA 47)	−33, 30, 6	174 voxels
Total	Left Superior Temporal Sulcus (BA 22)	−48, −40, 3	284 voxels

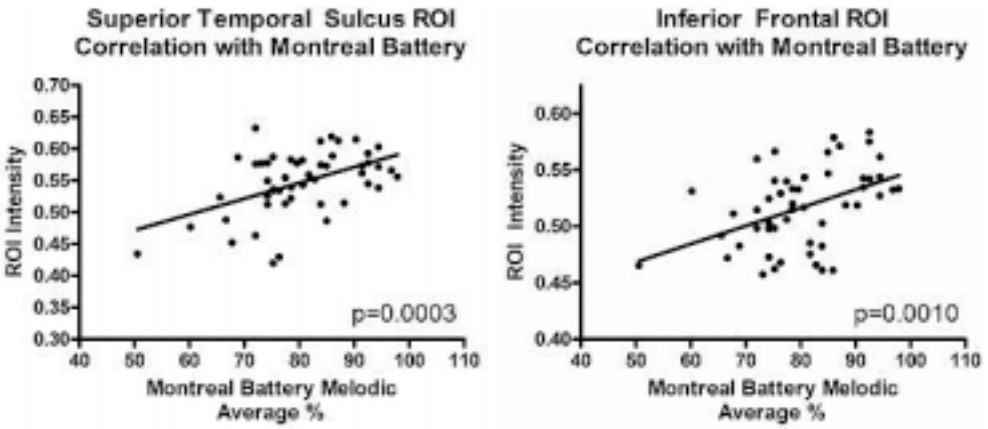

Fig. 2. Correlation analyses comparing the intensity of two regions of interest with the average melodic scores on the MBEA.

ate masks that would include as many brain regions as possible for this second step in the analysis. Subjects with an MBEA score of two standard deviations below the mean (as determined by a local group of 10 control subjects, rounded to the nearest integer) were included in the "true amusic subjects" group. Subjects scoring at the mean score (as determined by the local control group, rounded to the nearest integer) or higher constituted the "non-amusic subjects" control group. Only those voxel-by-voxel t-test results that survived Family Wise Error (FWE) corrections at $p < 0.05$ were considered statistically significant. The mean GM voxel intensity of each suprathreshold cluster of voxels was regressed against the MBEA performance for both melodic and rhythmic subtests and the total score.

3. Results

3.1. Melodic subtests VBM

The regression analysis of gray matter density and performance on the MBEA melodic subtest showed significant correlations within the temporal and inferior frontal lobe (Fig. 1a). Using the regression map as a template, we compared subjects who performed below a cutoff (mean – 2SD) – *the true amusic subjects* – with a group of normal controls in a two sample t-test. Two regions with local maxima within the left superior temporal sulcus (−48, −39, 5 (all coordinates are in MNI space); BA 22) and left inferior frontal gyrus (−24, 22, −23; BA47) showed significant differences between these two groups ($p < 0.05$, FWE corrected) (Fig. 1b). By extracting the mean regional gray matter intensity within these two regions we found a significant correlation between the MBEA melodic subtests and the individual gray matter concentrations for the ROI located in BA22 ($p = 0.0003$) and for the ROI located in BA47 ($p = 0.001$) (Fig. 2).

3.2. Rhythmic subtests VBM

Regressing gray matter density with performance on the MBEA rhythmic subtests showed several significant ($P < 0.005$) clusters (Fig. 3a). Using these regression maps as a template, a two-sample t-test showed significant differences in gray matter concentrations between the subgroup of "*true amusic subjects*" and a normal control group in the left STS (−49, −41, 6; BA 22) and the left IFG (−33, 29, 4; BA47) (Fig. 3b). We found strong correlations between regional mean gray matter concentrations and each subject's MBEA rhyth-

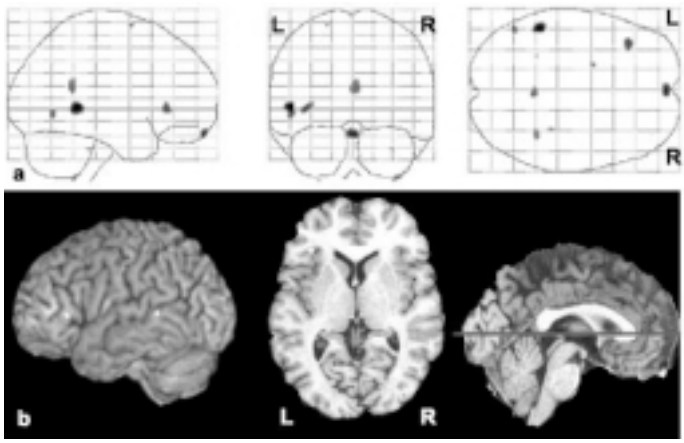

Fig. 3. Map generated by the regression analysis between all subjects' average rhythmic score and the individual gray matter concentrations ($P < 0.005$ uncorrected). This map was transformed into a binary mask which was then used as a template for the subsequent two-sample t-test. Figure 3b: Significant group differences in gray matter concentration between the *true amusic subgroup* and the *normal control group* ($p < 0.05$, FWE corrected) after applying the binary mask from Fig. 3a overlaid onto the surface reconstruction of a single spatially standardized brain. The position of the axial slice is marked with red in a midsagittal slice.

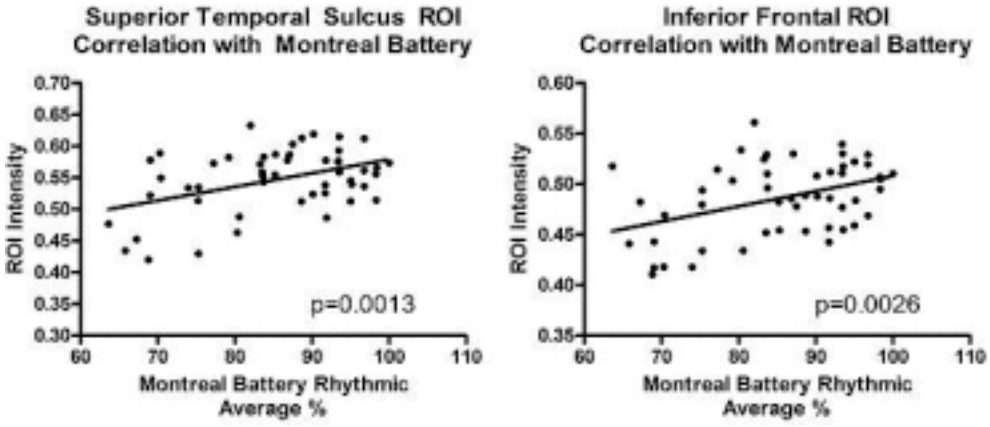

Fig. 4. Correlation analyses comparing the intensity of the three regions of interest produced by the rhythmic VBM with the rhythmic average on the MBEA.

mic subtest average for the ROI located in BA 22 ($p = 0.0013$) and the ROI located in BA 47 ($p = 0.0026$) (Fig. 4).

Both the rhythmic and melodic VBM analyses showed a correlation in BA47, but interestingly, the significant correlation for the rhythmic subtests was found in the superior aspect of BA47, while the significant correlation for the melodic subtests was in the most inferior subregion of BA 47. Both the rhythmic and melodic subtests showed correlations within subregions of BA22. Despite the fact that the local maxima produced by the rhythmic regression analysis ($-49, -42, 4$) is a few voxels off from the local maxima produced by the melodic regression ($-48, -40, 3$), the resolution limits imposed by the 12 mm smoothing inherent in the VBM preprocessing makes it extremely likely that the two BA22 regions are the same.

3.3. Total score VBM

The VBM regression analysis of the total MBEA score (average of subtests #1–5) showed a single large region located in the left temporal lobe (Fig. 5a). Using this regression map as a template, we found significant differences in gray matter density centered in the superior temporal sulcus ($-48, -39, 5$) (Fig. 5b) when comparing the subgroup of *"true amusic subjects"* with the group of normal controls. The mean regional gray matter concentrations from this region significantly correlated with the total MBEA score. Furthermore, gray

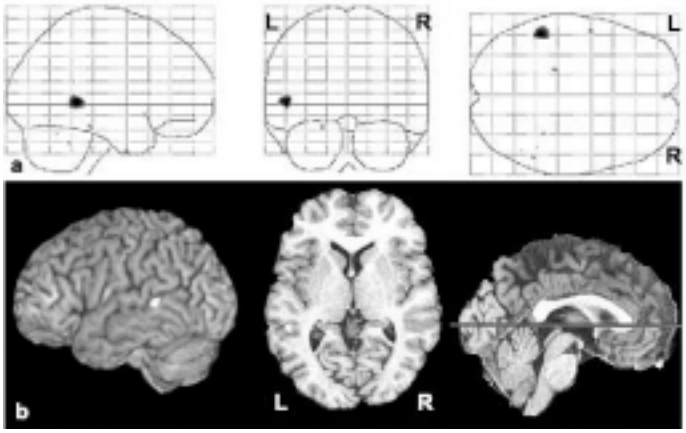

Fig. 5. Map generated by the regression analysis between all subjects' total scores and the individual gray matter concentrations ($P < 0.005$ uncorrected). This map was transformed into a binary mask which was then used as a template for the subsequent two-sample t-test. Figure 5b: Significant group differences in gray matter concentration between the *true amusic subgroup* and the *normal control group* ($p< 0.05$, FWE corrected) after applying the binary mask from Fig. 5a overlaid onto the surface reconstruction of a single spatially standardized brain. The position of the axial slice is marked with red in a midsagittal slice.

matter concentrations from the two regions identified in the melodic and rhythmic VBM analyses also showed significant correlations with the total MBEA performance (Fig. 6).

4. Discussion

Our results showed positive correlations between gray matter density variations between two regions in the brain (the left superior temporal sulcus and the posterior inferior frontal gyrus) and the averaged total MBEA scores as well as averaged melodic and rhythmic subtests scores. Although gray matter density in the left inferior frontal gyrus correlated with performance on both melodic and rhythmic subtests, a difference did emerge between the two: an inferior part of BA47 correlated with melodic performance, while a more superior part of BA47 correlated with rhythmic performance. In comparisons between subjects that were categorized as amusic (according to their performance on the MBEA) and subjects performing within the normal range of this test battery, the amusic subjects had significantly less gray matter volume than the normal controls in these regions.

In interpreting these findings, it is important to keep in mind that the MBEA requires subjects to make discriminations as they compare two musical phrases (with a short silence in between) in a forced, alternate choice design. It's interesting to ponder what strategies subjects are using when they take the MBEA. The incoming stream of pitches/sounds must be remembered in a temporally coherent way, discriminations and categorical decisions must be made. A developmental defect in any of these processes might result in a low performance on the MBEA which could explain some of the anomalies in amusic subjects that have been identified by various other test batteries. Previous studies have shown that amusics have a problem with frequency discrimination (Foxton et al., 2004; Peretz et al., 2002). Amusic subjects appear unable to discriminate between two frequencies that are less than 1 semitone apart. Although data from prior studies have supported a role for the primary auditory cortex in frequency discrimination (Menning et al., 2000; Tramo et al., 2002; Griffiths, 2003), neither our analysis using VBM of gray matter density images, nor analyses by other groups using either gray or white matter VBM (Hyde et al., 2004; Hyde et al., 2006) have found any structural anomalies involving primary auditory cortex. This might suggest that the underlying functional abnormality in congenital amusic subjects is not just a pitch discrimination problem but might include higher auditory processing or an auditory-motor integration problem as we are speculating further down.

It is interesting that we found these strong behavioral-anatomical covariations in the left hemisphere. There is extensive and sometimes conflicting literature on the lateralization of perceptual music tasks. Some general agreements seem to be that spectral processing involves more right-hemisphere regions while temporal processing involves more regions in the left hemisphere (Zatorre & Belin, 2001), although functional brain imaging studies still show activations in both hemispheres

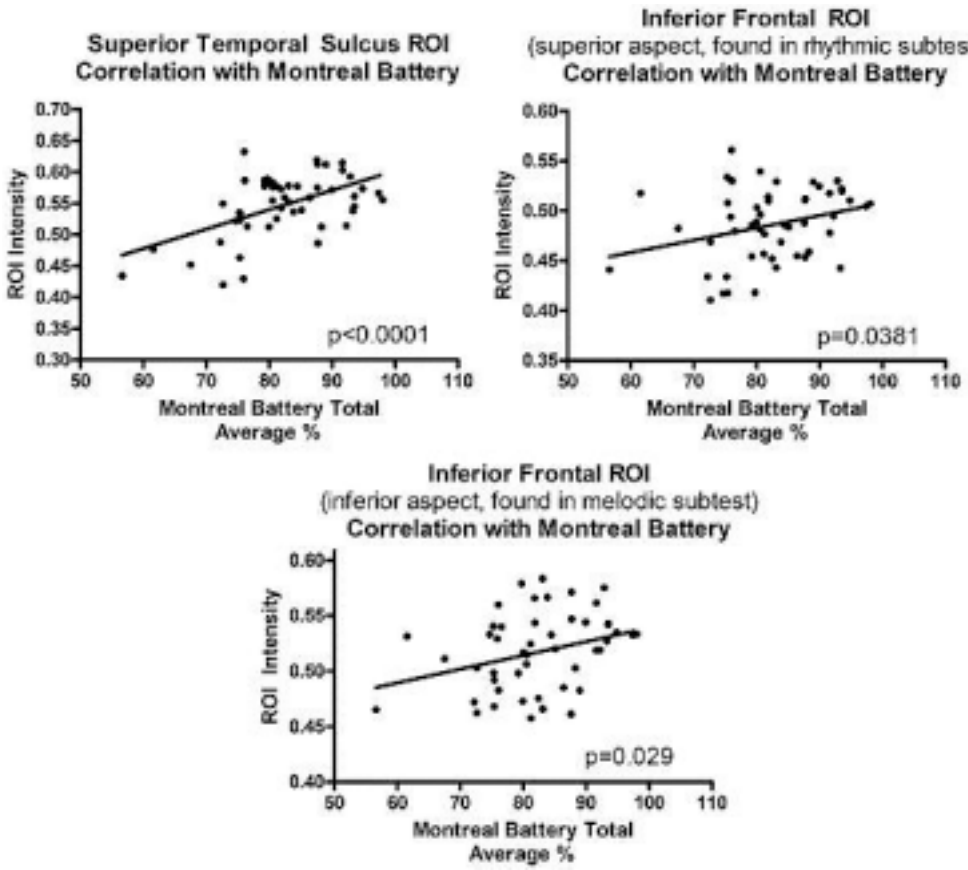

Fig. 6. Correlation analyses comparing the intensity of all three regions of interest produced by the melodic, rhythmic, and total score VBMs regressed with the total average of the MBEA.

even if one hemisphere is more activated than the other. Similarly, it has been found that tasks that require local processing (rhythm and pitch tasks) might show a left-hemisphere advantage while tasks that require global processing strategies (meter and melodic tasks) might show a right hemisphere advantage (Schuppert et al., 2000).

It is highly likely that subjects employ different cognitive strategies when listening to music, or for that matter, when taking the MBEA. Not only is music processing dependent on specific neural correlates relating to music (and amusia), but global cognitive processes such as memory, attention, and frontal processes (Schuppert et al., 2000) also come into play. Although our results support the existence of a leftward-dominance for the neural correlates that underlie congenital amusia, the involvement of such diverse and global cognitive processes as working memory, comparisons between two samples, categorical decisions, and focused attention could actually mask the more fundamental musical processing such as frequency discrimination, contour and pitch classification, and pitch memory that must take place underneath these global processes. Overall, music processing seems to rely on a bihemispheric network including (but not limited to) the superior temporal gyrus and sulcus, the inferior and superior parietal lobule as well as the inferior frontal gyrus and other parts of the premotor cortex in the middle frontal gyrus region (Schuppert et al., 2000; Zatorre & Belin, 2001; Gaab et al., 2003a,b; Patel, 2005). In contrast to the bihemispheric aspects of musical processing, language processing seems to be more strongly lateralized. The strong leftward lateralization of the anatomical-behavioral correlations in the present study could suggest that there might be more similarities between the underlying abnormalities in congenital amusia and language functions or language dysfunctions. This notion is supported by the growing literature suggesting that musical tasks and/or musical stimuli activate brain regions that are either identical or overlap

with brain regions that are active during language tasks (Koelsch et al., 2002, 2005; Gaab et al., 2003a,b; Patel et al., 1998, 2003; Guenther et al., 2006; Ozdemir et al., 2006).

Most studies have used perceptual tasks to examine the neural correlates of music and language processing. Production or expressive tasks have only rarely been used in functional imaging experiments, mostly because of the problems that overt expressive tasks might create in the functional imaging environment (e.g., movement artifacts). Nevertheless, published studies using positron emission tomography (PET) and fMRI methods have supported a bi-hemispheric role for the execution and sensorimotor control of vocal production both in speaking and in singing (Guenther, 1998; Jeffries et al., 2003; Brown et al., 2004; Okada & Hickock, 2006; Guenther et al., 2006; Ozdemir et al., 2006), but with a greater left-lateralization for speaking under normal physiological conditions. It is possible that the actual motor processes and sensorimotor control for speaking and singing are shared, but that the sensory representations of spoken and sung elements are separate or in different locations with a lesser degree of overlap than the expressive functions. The possible sharing of motor processes and sensorimotor control for expressive functions is important, since there is already some existing theoretical work and functional imaging work on the components of an articulatory network. This network might be important not only for articulatory problems in speaking but also for expressive problems while singing and it could potentially lead to the identification of key brain regions that might be altered in congenital amusia. Furthermore, we will show below that some of the regions that we have identified in our voxel-based morphometric analysis, are actually part of an articulatory network consisting of auditory regions that receive feedback and regions that map motor actions to the appropriate sound.

Based on imaging and cell recording studies, Guenther and colleagues (2006) proposed that three interacting subsystems control speech production: an auditory and a somatosensory feedback subsystem, and a feedforward control subsystem. In this model, the superior part of the temporal lobe (either STG or STS) receives projections from the frontal motor cortical areas that predict the sound of one's own voice and compare them with the auditory feedback (this is the function of the auditory error cells). The somatosensory feedback subsystem consists of primary and higher-order somatosensory areas that encode tactile and proprioceptive information for the sound being produced. As the third component of the model, the feed-forward control subsystem involves cortico-cortical projections from premotor to motor cortex (Guenther et al., 2006). The critical components of this network are in the superior part of the temporal lobe and the inferior part of the frontal gyrus which receive auditory feedback of the vocal output and use this information to make adjustments to the speech-sound map (or the auditory-motor map). It is most likely that singing requires a similar network of feedback regions and sound–motor action mapping regions. If one of these network regions or the connections between the nodal points in this network are impaired, then a subject won't be able to sing in tune or receive feedback to make the necessary adjustments to the singing output. The regions that play a critical role in the network, the superior part of the temporal lobe and the inferior part of the frontal lobe, were regions in which we found a significant decrease in gray-matter volume between the true amusic subjects and the control subjects.

The most significant gray matter differences were seen in the superior temporal sulcus on the left. Although the precise functional of this part of the STS is not know, it is thought that the STS might be involved in the categorization and recognition of sounds based on their elementary properties (Belin & Zatorre, 2003; Warren et al., 2003). Thus, the STS could be the perfect place for assessing whether a perceived sound (from auditory feedback) was congruent with the intended sound; this particular role could be a function of the auditory error cells which play an important role in the articulatory network. Several other studies have associated the STS with the identification or categorization of a variety of sounds (Engelien et al., 1995; Binder et al., 2000; Warren et al., 2003; Liebenthal et al., 2005; Mottonen et al., 2006).

The second region that showed pronounced gray matter differences between amusic and non-amusic subjects is the inferior frontal gyrus. Hyde et al. (2006) already identified the inferior frontal region as a potential area of abnormality in amusic subjects. Hyde et al. (2006) found white matter concentration differences in the white matter underlying the right inferior frontal gyrus. How can this be related our findings of less gray matter in the left IFG? One explanation might be that our current study and the one of Hyde et al. (2006) are looking at two sides of the same coin. A gray matter variation on one side of the brain could indirectly affect white matter composition on the homologue region of the other hemisphere through changing the composition of transcallosal fibers. Another explanation might

be that the underlying abnormality is bihemispheric but affects gray and white matter differently and depending on the sample size and the specific image analysis technique used, one investigator might find more abnormalities on the left and in the left IFG in particular (the current study), while another investigator might find more abnormalities on the right, such as the subcortical region of the right IFG (Hyde et al., 2006).

There has been increased interest in the function of the inferior frontal gyrus, since more and more studies found activations in this area with various fMRI tasks. It has been suggested that the IFG might play a role in simulating or integrating sequential (auditory) events or actions (Platel et al., 1997; Gaab et al., 2003a,b; Levitin et al., 2003; Nishitani et al., 2005), in the recognition of alterations in sequential auditory-perceptual events (Maess et al., 2001; Iacoboni et al., 2005), and in mapping sounds with motor actions (Bangert & Altenmueller, 2003; Baumann et al., 2005; Bangert et al., 2006; Lahav et al., 2007). Two independent voxel-based morphometric studies found more gray matter volume in the inferior frontal gyrus in musicians compared with non-musicians (Sluming et al., 2002; Gaser & Schlaug, 2003). Recent work suggests that the function of Broca's area (typically thought to consist of BA $44 + 45$) extends into BA 47 (Thompson-Schill, 2003), and BA47 is coactivated with Broca's area during language tasks (Sahin et al., 2004). This larger "Broca's Complex" includes two regions of interest found in our VBM analyses. Data supporting the idea of Broca's area as a general sequencer of actions (Nishitani et al., 2005; Fiebach & Schubotz, 2006) also support our finding that gray matter density variation within this region covaries with performance on melodic and rhythmic discrimination tasks. Various studies have found activations in the left inferior frontal gyrus with musical tasks, most typically with tasks requiring sequencing of musical stimuli (Platel et al, 1997; Maess et al., 2001; Gaab et al., 2003a,b). Similarly, data that have linked Broca's region to mapping of actions with sounds (Bangert & Altenmueller, 2003; Lahav et al., 2007) would support our hypothesis that the underlying dysfunction in congenital amusia might be that of a disorder of auditory-motor feedback or impairment mapping sounds with corresponding actions(i.e., the sounds of singing to the motor actions of singing). Amusic subjects seem to lack the ability to use the auditory feedback that they receive to evaluate and make corrections/adjustments as they sing. This suggests that the regions identified in the STS and the IFG may actually constitute a network of regions that enable the mapping of actions to sounds and create a feedback loop that allows for corrections of the motor action (i.e., singing) based on that perceptual feedback. Thus, the question that arises is whether congenital amusia (tone-deafness/the inability to sing in tune) is a disorder of the auditory-motor feedback loop or of auditory-motor integration. Our analysis suggests some candidate regions for further exploration of these hypotheses in future studies.

Acknowledgements

The authors wish to thank Dr. Isabelle Peretz for kindly providing us the Montreal Battery for the Evaluation of Amusia which was given to all of our subjects in order to determine their performance scores on these tests. We greatly appreciate the financial support of the Dana Foundation, the International Foundation for Music Research, and the GRAMMY Foundation that partially supported this research in addition to it's ongoing support of Dr. Schlaug's laboratory. We are also grateful for the many fruitful discussions that we had with colleagues in our group, in particular, Drs. Katie Overy, Amir Lahav, Marc Bangert, Nadine Gaab, and Andrea Norton about this research and its findings. And we thank Dr. Christian Gaser for his invaluable help with voxel-based morphometric methods.

References

Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry–the methods. *Neuroimage, 11*, 805-821.

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain, 125*, 238-251.

Bangert, M., & Altenmuller, E. O. (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neurosci, 4*, 26.

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

Baumann, S., Koeneke, S., Meyer, M., Lutz, K., & Jancke, L. (2005). A Network for Sensory-Motor Integration: What Happens in the Auditory Cortex during Piano Playing without Acoustic Feedback? *Ann N Y Acad Sci, 1060*, 186-188.

Belin, P., & Zatorre, R. J. (2003). Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport, 14*, 2105-2109.

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

Brown, S., Martinez, M. J., Hodges, D. A., Fox, P. T., & Parsons, L. M. (2004). The song system of the human brain. Brain Res. *Cogn Brain Res, 20*, 363-375.

Engelien, A., Silbersweig, D., Stern, E., Huber, W., Doring, W., Frith, C., & Frackowiak R. S. (1995). The functional anatomy of recovery from auditory agnosia. A PET study of sound categorization in a neurological patient and normal controls. *Brain, 118*, 1395-1409.

Fiebach, C. J., & Schubotz, R. I. (2006). Dynamic anticipatory processing of hierarchical sequential events: a common role for Broca's area and ventral premotor cortex across domains? *Cortex, 42*, 499-502.

Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., & Griffiths, T. D. (2004). Characterization of deficits in pitch perception underlying 'tone deafness'. *Brain, 127*, 801-810.

Gaab, N., Gaser, C., Zaehle, T., Jancke, L., & Schlaug, G. (2003a). Functional anatomy of pitch memory–an fMRI study with sparse temporal sampling, *Neuroimage, 19*, 1417-1426.

Gaab, N., Gaser, C., Zaehle, T., Chen, Y., & Schlaug, G. (2003b). Functional anatomy of pitch memory – a fMRI study with sparse temporal sampling. *Neuroimage, 19*, 1417-1426.

Gaab, N., & Schlaug, G. (2003). The Effect of musicianship on pitch memory in performance matched groups. *Neuroreport, 14*, 2291-2295.

Gaser, C., & Schlaug, G. (2003). Gray matter differences between musicians and nonmusicians. *Ann NY Acad Sci, 999*, 514-517.

Good, C. D., Ashburner, J., & Frackowiak, R. S. (2001). Computational neuroanatomy: new perspectives for neuroradiology. *Rev Neurol* (*Paris*)*, 157*, 797-806.

Griffiths, T. D. (2003). Functional imaging of pitch analysis. *Ann NY Acad Sci, 999*, 40-49.

Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychol Rev, 105*, 611-633.

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang, 96*, 280-301.

Hyde, K. L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychol Sci, 15*, 356-360.

Hyde, K. L., Zatorre, R. J., Griffiths, T. D., Lerch, J. P., & Peretz, I. (2006). Morphometry of the amusic brain: a two-site study. *Brain, 128*, 2562-2570.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol, 3*, e79.

Kalmus, H., & Fry, D. B. (1980). On tune deafness (dysmelodia): frequency, development, genetics and musical background. *Ann Hum Genet, 43*, 369-382.

Kleist, K. (1959). Sensorische Aphasien und Amusien auf myeloarchitectonischer Grundlage (Thieme Verlag, Stuttgart), pp. 32-43.

Koelsch, S., Gunter, T. C., v Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: a cortical "language-network" serves the processing of music. *Neuroimage, 17*, 956-966.

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

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

Levitin, D. J., & Menon, V. (2003). Musical structure is processed in "language" areas of the brain: a possible role for Brodmann Area 47 in temporal coherence. *Neuroimage, 20*, 2142-2152.

Liebenthal, E., Binder, J. R., Spitzer, S. M., Possing, E. T., & Medler, D. A. (2005). Neural substrates of phonemic perception, *Cereb Cortex, 15*, 1621-1631.

Luders, E., Gaser, C., & Schlaug, G. (2004). A voxel-based approach to gray matter asymmetries. *Neuroimage, 22*, 656-664.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nat Neurosci, 4*, 540-545.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., & Frith, C.D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proc Natl Acad Sci USA, 97*, 4398-4403.

McChesney-Atkins, S., Davies, K. G., Montouris, G. D., Silver, J. T., & Menkes, D. L. (2003). Amusia after right frontal resection for epilepsy with singing seizures: case report and review of the literature. *Epilepsy Behav, 4*, 343-347.

Menning, H., Roberts, L. E., & Pantev, C. (2000). Plastic changes in the auditory cortex induced by intensive frequency discrimination training. *Neuroreport, 11*, 817-822.

Meyer, M., Baumann, S., & Jancke, L. (2006). Electrical brain imaging reveals spatio-temporal dynamics of timbre perception in humans. *Neuroimage, 32*, 1510-1523.

Mottonen, R., Calvert, G. A., Jaaskelainen, I. P., Matthews, P. M., Thesen. T., Tuomainen, J., & Sams, M. (2006). Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *Neuroimage, 30*, 563-569.

Nishitani, N., Schürmann, M., Amunts, K., & Hari, R. (2005). Broca's Region: From Action to Language. *Physiology, 20*, 60-69.

Okada, K., & Hickok, G. (2006). Left posterior auditory-related cortices participate both in speech perception and speech production: Neural overlap revealed by fMRI. *Brain Lang, 98*, 112-117.

Ozdemir, E., Norton, A., & Schlaug, G. (2006). Shared and distinct neural correlates of singing and speaking. *Neuroimage, 33*, 628-635.

Patel, A. D., Peretz, I., Tramo, M., & Labreque, R. (1998). Processing prosodic and musical patterns: A neuropsychological investigation. *Brain Lang, 61*, 123-144.

Patel, A. D. (2005). The relationship of music to the melody of speech and to syntactic processing disorders in aphasia. *Ann NY Acad Sci, 1060*, 59-70.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *J Cogn Neurosci, 10*, 717-733.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nat Neurosci, 6*, 674-681.

Paulson, M. J., & Lin T. (1970). Predicting WAIS IQ from shipley-hartford scores. *J Clin Psychol, 26*, 453-461.

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., et al. (2002). Congenital amusia: a disorder of fine-grained pitch discrimination. *Neuron, 17;33*, 185-191.

Peretz, I., & Hyde, K. L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends Cogn Sci, 7*, 362-367.

Peretz, I., Brattico, E., & Tervaniemi, M. (2005). Abnormal electrical brain responses to pitch in congenital amusia. *Ann Neurol, 58*, 478-482.

Platel, H., Price, C., Baron, J. C., Wise, R., Lambert, J., Frackowiak, R. S., Lechevalier, B., & Eustache, F. (1997). The structural components of music perception. A functional anatomical study. *Brain, 120*, 229-243.

Sahin, N. T., Halgren, E., Ulbert, I., Dale, A., Schomer, D., Wu, J., & Pinker, S., Abstract grammatical processing in Broca's Area: Evidence from human *in-vivo* electrophysiology and fMRI. Poster presented at the Organization for Human Brain Mapping Annual Meeting, June 14–18, 2004, Budapest, Hungary. Poster number MO 144.

Satoh. M., Takeda, K., Murakami, Y., Onouchi, K., Inoue, K., & Kuzuhara, S. (2005). Cortex. A case of amusia caused by the infarction of anterior portion of bilateral temporal lobes. *Cortex, 41*, 77-83.

Schlaug, G. (2001). The brain of musicians. *Ann N Y Acad Sci, 930*, 281-299.

Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat Neurosci, 5*, 688-694.

Schuppert, M., Münte, T., Wieringa, B., & Altenmüller, E. (2000). Receptive amusia: evidence for cross-hemispheric neural networks underlying music processing strategies. *Brain, 123*, 546-559.

Shipley, W. C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *J Psychol, 9*, 371-377.

Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes. A., & Roberts, N. (2002). Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *Neuroimage, 17*, 1613-1622.

Terao, Y., Mizuno, T., Shindoh, M., Sakurai, Y., Ugawa, Y., Kobayashi, S., Nagai, C., Furubayashi, T., Arai, N., Okabe, S., Mochizuki, H., Hanajima, R., & Tsuji, S. (2006). Vocal amusia in a professional tango singer due to a right superior temporal cortex infarction. *Neuropsychologia, 44*, 479-8.

Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: inferring "how" from "where". *Neuropsychologia, 41*, 280-292.

Tramo, M. J., Shah, G. D., & Braida, L. D. (2002). Functional role of auditory cortex in frequency processing and pitch perception. *J Neurophysiol, 87*, 122-139.

Tzortzis. C., Goldblum, M. C., Dang, M., Forette, F., & Boller, F. (2000). Absence of amusia and preserved naming of musical instruments in an aphasic composer. *Cortex, 36*, 227-242.

Warren, J. D., Uppenkamp, S., Patterson, R. D., & Griffiths, T. D. (2003). Separating pitch chroma and pitch height in the human brain. *Proc Natl Acad Sci USA, 100*, 10038-10042.

Watkins, K. E., Vargha-Khadem, F., Ashburner, J., Passingham, R. E., Connelly, A., Friston, K. J., et al. (2002). MRI analysis of an inherited speech and language disorder: structural brain abnormalities. *Brain, 125*, 465-478.

Yoo, S. S., O'leary, H. M., Dickey, C. C., Wei, X. C., Guttmann, C. R., Park, H. W., & Panych, L. P. (2005). Functional asymmetry in human primary auditory cortex: Identified from longitudinal fMRI study. *Neurosci Lett, 383*, 1-6.

Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb Cortex, 11*, 946-953.