BRAIN
A JOURNAL OF NEUROLOGY

Impaired pitch perception and memory in congenital amusia: the deficit starts in the auditory cortex

Philippe Albouy,[1,2,3] Jérémie Mattout,[2,3] Romain Bouet,[2,3] Emmanuel Maby,[2,3] Gaëtan Sanchez,[2,3] Pierre-Emmanuel Aguera,[2,3] Sébastien Daligault,[4] Claude Delpuech,[2,3,4] Olivier Bertrand,[2,3] Anne Caclin[2,3,]* and Barbara Tillmann[1,3,]*

1 CNRS, UMR5292; INSERM, U1028; Lyon Neuroscience Research Centre, Auditory Cognition and Psychoacoustics team, Lyon, F-69000, France
2 INSERM, U1028; CNRS, UMR5292; Lyon Neuroscience Research Centre, Brain Dynamics and Cognition Team, Lyon, F-69000, France
3 University Lyon 1, Lyon, F-69000, France
4 CERMEP, MEG Department, Bron, F-69000, France

*These authors contributed equally to this work.

Correspondence to: Philippe Albouy,
INSERM U1028 - CNRS UMR5292, Centre de Recherche en Neurosciences de Lyon,
Equipe Dynamique Cérébrale et Cognition,
Centre Hospitalier le Vinatier,
Batiment 452, 95 Bd Pinel,
Bron, F-69500,
France
E-mail: philippe.albouy@inserm.fr

Congenital amusia is a lifelong disorder of music perception and production. The present study investigated the cerebral bases of impaired pitch perception and memory in congenital amusia using behavioural measures, magnetoencephalography and voxel-based morphometry. Congenital amusics and matched control subjects performed two melodic tasks (a melodic contour task and an easier transposition task); they had to indicate whether sequences of six tones (presented in pairs) were the same or different. Behavioural data indicated that in comparison with control participants, amusics' short-term memory was impaired for the melodic contour task, but not for the transposition task. The major finding was that pitch processing and short-term memory deficits can be traced down to amusics' early brain responses during encoding of the melodic information. Temporal and frontal generators of the N100m evoked by each note of the melody were abnormally recruited in the amusic brain. Dynamic causal modelling of the N100m further revealed decreased intrinsic connectivity in both auditory cortices, increased lateral connectivity between auditory cortices as well as a decreased right fronto-temporal backward connectivity in amusics relative to control subjects. Abnormal functioning of this fronto-temporal network was also shown during the retention interval and the retrieval of melodic information. In particular, induced gamma oscillations in right frontal areas were decreased in amusics during the retention interval. Using voxel-based morphometry, we confirmed morphological brain anomalies in terms of white and grey matter concentration in the right inferior frontal gyrus and the right superior temporal gyrus in the amusic brain. The convergence between functional and structural brain differences strengthens the hypothesis of abnormalities in the fronto-temporal pathway of the amusic brain. Our data provide first evidence of altered functioning of the auditory cortices during pitch perception and memory in congenital amusia. They further support the hypothesis that in neurodevelopmental disorders impacting high-level functions (here musical abilities), abnormalities in cerebral processing can be observed in early brain responses.

Introduction

About 4% of the population is estimated to experience a lifelong deficit in music perception and production that cannot be explained by hearing loss, brain damage, or cognitive deficits (Ayotte *et al.*, 2002; Peretz *et al.*, 2002; Stewart 2006, 2008). Individuals afflicted with congenital amusia are unable to detect out-of-key tones, and are unaware when others (or themselves) sing out of tune. Thanks to the Montreal Battery of Evaluation of Amusia (Peretz *et al.*, 2003), this disorder has been systematically studied. Behavioural investigations have revealed that the impairment is linked to a deficit in pitch perception (Foxton *et al.*, 2004; Hyde and Peretz, 2004) and memory (Gosselin *et al.*, 2009; Tillmann *et al.*, 2009; Williamson *et al.*, 2010). However, only few studies have investigated the neural bases of congenital amusia from structural or functional cerebral perspectives.

Using voxel-based morphometry (VBM), cortical thickness and diffusion tensor imaging methods, anatomical abnormalities were reported in the right fronto-temporal pathway. In comparison with matched control subjects, amusics' brains showed decreased white matter concentration in the right inferior frontal gyrus associated with increased grey matter concentration in the same region (Hyde *et al.*, 2006), and in the right superior temporal gyrus (Hyde *et al.*, 2007). The hypothesis of an abnormal right fronto-temporal pathway in the amusic brain has received support by the observation of reduced fibre connectivity in the right arcuate fasciculus (Loui *et al.*, 2009).

To relate anatomical anomalies to behavioural expressions, functional investigations have explored the cerebral basis of pitch perception in congenital amusia. Electrophysiological (Peretz *et al.*, 2005, 2009; Moreau *et al.*, 2009) and functional MRI investigations (Hyde *et al.*, 2011) have shown that the auditory cortex of amusic individuals responds normally to pitch, and that the amusic brain can track small pitch changes (i.e. a quarter tone), suggesting near-normal cerebral processing of musical material (Moreau *et al.*, 2009; Peretz *et al.*, 2009). These findings were surprising in light of the observed behavioural deficits in amusia as the right auditory cortex has been shown to play a critical role in various aspects of pitch processing, such as melody perception and discrimination (Peretz, 1990; Johnsrude *et al.*, 2000).

Our study investigated for the first time the cerebral correlates of pitch perception and short-term memory in congenital amusia by combining behavioural, magnetoencephalography (MEG), and VBM approaches. During MEG recording, amusics and matched control subjects performed two melodic tasks, in which two six-tone sequences were compared (same/different paradigm); a contour task and an easier transposition task. We expected impaired short-term memory performance in amusic participants for the contour task, but not for the transposition task, which did not require melodic contour memorization. The easier transposition task was designed as a control condition, for which amusic participants perform as well as control subjects, to rule out the interpretation that brain activity differences between the two participant groups for the contour task might be mere correlates of motivation or attention-level differences between the groups.

We analysed brain responses during encoding, retention and retrieval of melodic information using source modelling of MEG data to investigate whether altered functioning of the auditory and/or frontal cortices might underlie the disorder. We focused on (i) event-related responses for each note of the melodies for the encoding and retrieval parts of the tasks; (ii) sustained evoked responses for the encoding part of the tasks; and (iii) oscillatory gamma activities during the maintenance of musical information. Oscillations in the gamma frequency band have been reported to contribute to the neuronal underpinnings of working memory (Tallon-Baudry and Bertrand, 1999; Tallon-Baudry *et al.*, 1999; Kaiser *et al.*, 2003; Kaiser and Lutzenberger, 2005; Jensen *et al.*, 2007; Mainy *et al.*, 2007). Finally, using VBM in the same participants, we expected to observe anatomical abnormalities in amusic participants (compared with control subjects) in right frontal and temporal areas, as previously shown by Hyde *et al.* (2006, 2007), and we further aimed to compare the locations of these anatomical abnormalities with that of possible functional abnormalities revealed by MEG source analyses.

Material and methods

Participants and behavioural pretests

Nine amusic adults (five females; mean age, 31.55 ± 8.50 years, ranging from 20 to 44; mean education, 14.77 ± 1.71 years; mean musical education 1.16 years, ranging from 0 to 2.5) and nine matched non-musician control subjects (five females; mean age, 31.33 ± 7.31 years, ranging from 24 to 47; mean education, 16.11 ± 2.57 years; mean musical education 0.77 years ranging from 0 to 3) participated in the study. Each group was composed of seven right-handed and two left-handed participants. Severe peripheral hearing loss was excluded using standard audiometry and all participants reported normal hearing and no history of neurological or psychiatric disease. Participants gave their written informed consent, and were paid for their participation. Ethical approval was obtained from the French ethics committee on Human Research (CPP Sud-Est II, #2006-018/A-1).

To be considered as amusic, participants had to obtain an average score two standard deviations (SD) below the average of the normal population on the Montreal Battery of Evaluation of Amusia (i.e. a cut-off score of 23, maximum score = 30; Peretz *et al.*, 2003). In the Montreal Battery of Evaluation of Amusia, six subtests assess various components of music perception and memory by considering that musical material must be processed along a melodic dimension (i.e. sequential variations in pitch) and a temporal dimension (i.e. sequential variations in duration). All participants were tested with the Montreal Battery of Evaluation

of Amusia; the average scores of the amusic group (mean = 20.90, SD = 1.70, ranging from 18 to 22.83) differed significantly from the scores of the control group [mean = 27.61, SD = 0.83, ranging from 26.3 to 28.6, $t(16)$ = 10.59, $P < 0.0001$].

To determine pitch discrimination thresholds, a two-alternative forced-choice task was used with an adaptive tracking, two-down/one-up staircase procedure (see Tillmann *et al.*, 2009 for details). Observed pitch discrimination thresholds of the amusic group were higher (mean = 1.07, SD = 1.20, ranging from 0.14 to 4 semitones) than that of the control group [mean = 0.31, SD = 0.30, ranging from 0.07 to 0.95, $t(16)$ = -1.83, $P = 0.042$, one-tailed]. The observed overlap in pitch thresholds between amusic and control groups was in agreement with previous findings (Foxton *et al.*, 2004; Tillmann *et al.*, 2009).

Voxel-based morphometry

All participants, except one amusic participant (because of claustrophobia), underwent a 3D anatomical MPRAGE T_1-weighted MRI scan on a 1.5 T Siemens Magnetom scanner (Siemens AG) equipped with an 8-channel head coil (repetition time = 1970 ms; echo time = 3.93 ms; inversion time = 1100 ms; flip angle: 15°, field of view 256 × 256 mm, voxel size 1.0 × 1.0 × 1.0 mm). The anatomical volume consisted of 176 sagittal slices with 1 mm³ voxels, covering the whole brain.

Preprocessing and segmentation

All image preprocessing and voxel-by-voxel statistical analyses were performed using the VBM functions of SPM8 (Wellcome Trust Centre for Neuroimaging, http://www.fil.ion.ucl.ac.uk/spm/, London, UK). Before preprocessing, all images were checked for artefacts and automatically aligned so that the origin of the coordinate system was located at the anterior commissure. Using the unified segmentation procedure implemented in SPM8 (Ashburner and Friston, 2005), the images were segmented into grey matter, white matter and CSF. For each participant, this resulted in a set of three images in the same space as the original T_1-weighted image, in which each voxel was assigned a probability of being grey matter, white matter and CSF, respectively, as well as a normalized version of these images (using the T_1-template from the Montreal Neurological Institute, provided by SPM8). Finally, normalized-segmented images were spatially smoothed with an 8-mm full-width at half-maximum isotropic Gaussian kernel.

Voxel-based morphometry statistical analyses

Group comparisons with two-sample t-tests at each voxel were performed separately for grey matter and white matter. Grey and white matter concentrations were compared for each brain voxel between amusics ($n = 8$) and control subjects ($n = 8$) (the control participant matched to the amusic participant without T_1-weighted MRI was excluded from the analysis). Given the limited sample size, an exploratory whole-brain analysis was not possible, and we focused our analyses on bilateral inferior frontal gyrus and superior temporal gyrus, where abnormalities have been reported previously in congenital amusia (Hyde *et al.*, 2006, 2007; Mandell *et al.*, 2007; Loui *et al.*, 2009), and which are key regions in

pitch processing and memorizing. We thus adopted a lenient statistical threshold of $P < 0.05$ (uncorrected), and only report group differences that were found <1 cm away (in any direction) from the coordinates in the right inferior frontal gyrus and right superior temporal gyrus observed in previous studies, which directly compared amusics and control subjects with VBM (Hyde *et al.*, 2006) or cortical thickness measures (Hyde *et al.*, 2007). We also assessed symmetric locations in the left hemisphere.

Short-term memory tasks for tone sequences

Material

Participants performed two melodic tasks, both requesting to compare two six-tone sequences (S1, S2) separated by a silent retention period of 2000 ms. All sequences were composed of six 250 ms piano tones presented successively without interstimulus interval. The two tone sequences could be either the same or different. For 'different' trials, the second sequence differed by a single tone in the contour task and by all tones [i.e. transposed an interval of seven semi-tones (a fifth) up or down] in the transposition task (Fig. 1). One hundred and ninety-two different melodies (sequences) were created using eight piano tones differing in pitch height (Cubase software); all used tones belonged to the key of C Major (C3, D3, E3, F3, G3, A3, B3, C4). These 192 sequences were used as S1 and were the same for the contour and transposition tasks. In order to strengthen the tonality, the tone C occurred twice in each sequence, and half of the melodies ended on the tone C and the other half on the tone G (i.e. two functionally important tones in the key of C Major). Identical tones were not repeated consecutively in a sequence. For S2 in different trials of the contour and transposition tasks, variants of these sequences were created. For the contour task, one tone (in positions 2 to 5) was replaced by a different tone of the set to create a contour-violation in the melody. For the transposition task (i.e. the transposed sequences), two supplementary tone sets were created for the pitch-up condition (i.e. D4, E4, F4#, G4), and the pitch-down condition (i.e. F2, G2, A2, B2b), respectively.
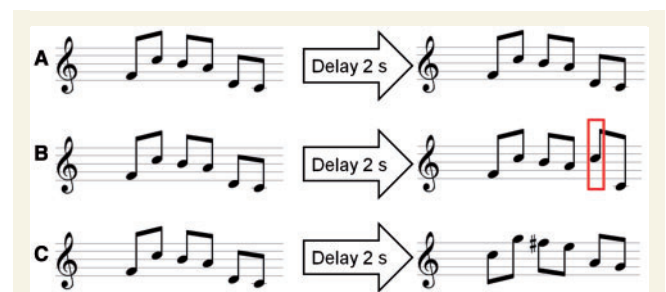
Figure 1 Examples of the musical stimuli used. (**A**) 'Same' trial in the contour task or the transposition task. S1 was repeated as the second melody of the pair (S2) after a 2 s delay. (**B** and **C**) For 'different' trials, the second melody of the pair changed only for one tone in contour task (**B**, red square) and was a transposition of S1 in transposition task (example of the pitch-up condition in **C**).

Procedure

Participants performed the contour and transposition tasks during the MEG recording. Presentation software (Neurobehavioral systems) was used to run the experiment and to record button presses. For each trial, participants listened binaurally to the first six-tone sequence with a total duration of 1.5 s (S1), followed by a silent retention period of 2 s, and then the second sequence (S2, 1.5 s). Participants had to decide whether S2 was identical to S1 or different from S1. There were six blocks of each of the two tasks. The blocks were separated by a break of 2–3 min. The two tasks were presented in alternation (counterbalanced across participants). Participants were informed of task order and asked to indicate their answers by pressing one of two keys with their right hand after the end of S2. They had 2 s to respond before the next trial, which occurred 2.5–3 s after the end of S2. No feedback was given during the experiment. In each block, 32 trials were presented (16 same pairs, 16 different pairs), resulting in 192 trials for each task. Within each block, the trials were presented in a pseudo-randomized order, with several constraints: the same trial type (i.e. same, different), or melodies ending by the same tone (C or G), could not be repeated more than three times in a row. Before the MEG recording, participants performed for each task a set of 15 practice trials without feedback.

Magnetoencephalography

Recordings

The recordings were carried out using a 275-channel whole-head MEG system (CTF-275 by VSM Medtech Inc.) with continuous sampling at a rate of 600 Hz, a 0–150 Hz filter bandwidth, and first-order spatial gradient noise cancellation. Horizontal and vertical electrooculograms and electrocardiogram were acquired with bipolar montages. Head position was determined with coils fixated at the nasion and the preauricular points (fiducial points). Head position was acquired continuously (continuous sampling at a rate of 150 Hz) and checked at the beginning of each block to ensure that head movements did not exceed 0.5 cm (this was confirmed by additional offline checking before the data analyses). Participants were seated upright in a sound-attenuated, magnetically-shielded recording room, and listened to the sounds presented binaurally through air-conducting tubes with foam ear tips. Before the MEG recording, participants' sound detection thresholds (using G3, the tone in the centre of the tone set for S1) were determined for each ear, and the level was adjusted so that the sounds were presented at ~50–55 dB sensation level with a central position (stereo) with respect to the participant's head. After the MEG session, participants' subjective reports regarding their strategies were collected.

Outline of the data analyses

The analyses reported here focused on event-related fields evoked by S1 (i.e. during the encoding and memorization of the melodic pattern); induced gamma oscillations during the delay period, corresponding to the retention in memory of the melodic information; and event-related fields evoked by the changed tone in S2 for the contour task (i.e. during the retrieval of melodic information).

MEG data were first analysed in sensor space using CTF tools (VSM Medtech Inc.) and the ELAN software package developed in the Brain Dynamics and Cognition team (Lyon Neuroscience Research Center, http://elan.lyon.inserm.fr/; Aguera *et al.*, 2011). Source reconstruction was performed with SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK; Friston *et al.*, 2008; Litvak *et al.*, 2011) for event-related fields, and the FieldTrip toolbox (Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands, http://www.ru.nl/fcdonders/fieldtrip) for oscillatory activities, both using MATLAB 7.6 (Mathworks Inc.).

Data preprocessing

For two amusic participants, the original MEG recordings were contaminated by ferromagnetic artefacts caused by metallic dental prostheses, which created a temporally stationary artefact at the participant's respiratory frequency (artefact frequency; mean = 0.24 Hz, ranging from 0.22 to 0.29 Hz, SD = 0.02 Hz for one amusic participant; and mean = 0.31 Hz, ranging from 0.27 to 0.39, SD = 0.05 Hz for the second amusic participant). In addition, the MEG raw data of another amusic participant were contaminated by heartbeat signal (mean artefact frequency = 1.41 Hz, ranging from 1.38 to 1.44; SD = 0.01 Hz). As congenital amusics represent a limited participant pool, we applied a signal correction model with a signal space projection method provided by CTF (Tesche *et al.*, 1995; Ramirez *et al.*, 2011). This method allows uncovering a signal-space vector that describes the global topography of the artefact in sensor space (magnitude varied between 2000 fT and 2500 fT for the dental prosthesis artefacts, and between 20 000 fT and 22 000 fT for the heartbeat artefact). A projection of the MEG data onto the space orthogonal to this vector was then applied to remove the artefact. The efficiency of this correction to faithfully recover the cerebral signals was further assessed with additional recordings of auditory event-related fields in a control participant with and without an artefact-generating piece of metal attached to his head (not shown).

Individual MEG trials were automatically inspected from −100 ms to 5500 ms with respect to the onset of the first S1 tone (i.e. a time window covering S1, the delay, S2 and an additional 500 ms after S2). Trials with ranges of values exceeding ± 3000 fT within a 1000 ms sliding time-window at any sensor site (±100 µV at electrooculograms channels) were excluded from the analysis; as a result, between 90 and 165 trials were kept for each participant and condition.

Event-related fields

Averaging was done separately for the two experimental conditions (contour task and transposition task) and a − 100 to 0 ms interval before the first tone in S1 was used for baseline correction. Grand-average event-related fields were plotted at the sensor level using ELAN and were examined to determine topographical patterns reflecting sources in the auditory cortices (Fig. 4D). Accordingly, S1 analyses focused on (i) transient evoked responses: P50m and N100m (the magnetic counterparts of P50 and N100 waves observed with EEG) for the first tone, and N100m for subsequent tones (an average of the responses to tones 2 to 6 was created for that purpose); and (ii) the sustained

evoked response during the entire melody. The two types of responses were dissociated using two different second-order Butterworth filters (12 dB/octave slope): for transient evoked responses, a band-pass filter (cut-off frequencies at 2 and 30 Hz) was used to eliminate the sustained evoked responses and high-frequency noise (Fig. 4B); for the sustained evoked response a low-pass filter with a 2 Hz cut-off frequency was applied (Fig. 4C).

To analyse event-related responses following the 'changed' tone in S2 of the contour task, two additional averages were performed for each participant (note that we always kept the baseline in the −100 to 0 ms interval before S1); firstly, an average of all correctly detected changed tones, in a −100 to 700 ms time-window around the onset of the change (this event-related field thus combined data for differences in all possible positions in S2), and secondly, an average of tones from correctly classified same trials in the contour task with, for each participant, the same number of tones in position 2, 3, 4 or 5 was used for the event-related fields of the different trials. The change-specific response was then assessed as the difference between these two event-related fields. Note that because control participants had more correct responses than amusic participants, we used for each control participant the same number of trials as his/her matched amusic participant (selected randomly from the entire set of correct response trials). For the analysis of the change-specific responses, filtering was done with a band-pass filter between 0.5 and 30 Hz. As for S1, the frequency band of interest for S2 was chosen after observation of grand-average event-related responses.

Analyses were performed at the source level for the transient evoked responses to determine and characterize the generators involved in the encoding of S1 and the retrieval of melodic information in S2. For the sustained evoked responses, analyses were performed at the sensor-level (see below).

Source reconstruction of transient responses (S1 and S2)

For these analyses, we used the original MEG signals for all participants but one, for whom artefact-corrected data were used because of the heartbeat artefact. Distributed cortical source reconstruction of event-related fields were performed for the brain responses to the first tone and the average responses to the five subsequent tones of S1 separately, in a post-stimulus time window from 25 to 175 ms. The source reconstruction for responses to the changed tone in S2 (contour task) was performed for the difference waveform for correct responses (different trial – same trial, see above) in a −100 to 700 ms time window relative to the onset of the changed tone. For each participant (except the amusic participant without individual MRI), we created individual head meshes (20 484 vertices or dipolar sources) describing the boundaries of different head compartments (scalp, inner skull and cortical sheet) based on the participant's own structural brain scans (note that the same set of vertices were used for all participants, but the mesh was deformed to account for each participant anatomy). For the amusic participant who did not undergo the 3D MRI scan, we used the MNI template provided with SPM8. We then performed landmark-based co-registration of MEG data and MRIs using the locations of nasion and pre-auricular points (Mattout *et al.*, 2007). Individual inverse solutions were obtained using the empirical Bayesian approach

implemented in SPM8 (Mattout *et al.*, 2006). For each participant, and using the multiple sparse prior model, the two conditions for S1 (contour task, transposition task) were inverted together to increase the sensitivity of the ensuing statistical comparison between them (Friston *et al.*, 2008). For S2, only the contour task was inverted as we were interested in the changed tone of the melodic contour task.

Definition of the regions of interest for transient responses in S1 and S2

Having completed the inversion for S1, we divided the source-reconstructed data in two separate time windows of interest, corresponding to the P50m peak (40–70 ms after the onset of the first tone); and to the N100m components (70–160 ms after the onset of each tone), respectively. For S2, the time window of interest was the 0–600 ms period after the onset of the changed tone. For each participant and time sample in each time window of interest, we extracted the cortical points (region of interest) that were significantly different from baseline according to the posterior probability maps provided by the Bayesian inversion. A posterior probability map provides the probability or confidence that the activation in a particular cortical vertex exceeds some specified threshold, given the data (Friston and Penny, 2003). We considered a zero-value threshold. As the data had been baseline-corrected prior to the inversion, this analysis resulted in identifying vertices with activity that significantly differ from baseline. This approach has two main advantages: (i) it provides inference at the 3D cortical surface of each individual, as a consequence of the Bayesian inversion; and (ii) it is not submitted to the multiple-comparison problem because the probability that an activation has occurred, at any particular cortical vertex, is the same, irrespective of whether only that vertex or the entire brain is analysed (Friston and Penny, 2003). Moreover, as each individual mesh is mapping in a one-to-one fashion to the MNI–template mesh (Mattout *et al.*, 2007), individual statistics can be combined to perform group-level inference. For each participant and time sample in each time window of interest, we used posterior probability maps with a 99.98% threshold (i.e. 5% of false positive at most, Bonferroni-corrected across participants). This step provided for each participant and each time sample a list of cortical dipoles (vertices) where activity emerged significantly from baseline. We only analysed vertices with activity emerging for at least one-third of the duration of the time window of interest, in at least three participants (without considering the group factor). This provided separate sets of activated cortical vertices for the P50m for the first tone, the N100m for the first tone, the N100m for the tones 2 to 6, and the responses elicited by the changed tone in S2.

Amplitude and latency analyses of source data for transient responses in S1 and S2

Source activities were analysed for each region of the cortical mesh where activity was significantly different from baseline (Tables 1 and 2), an average of the activity of all vertices was created for each region in each hemisphere. For the amplitude analysis of the source data for S1, we conducted an ANOVA on the source time course at each time sample and for each

Table 1 Frontal and Temporal Generators of the N100m for tone 1 and tones 2–6

Tone(s)	Lobe	Region	Hemisphere	x	y	z	mm^2	Number of vertices
Tone 1	Frontal	Inferior frontal gyrus, pars opercularis	Right	54	8	6	38	4
			Left	−51	4	5	14	2
	Temporal	Heschl's gyrus/superior temporal gyrus/planum temporale	Right	45	−19	6	44	8
			Left	−44	−18	5	85	11
Tones 2 to 6	Frontal	Inferior frontal gyrus, pars opercularis	Right	54	8	6	38	4
			Left	−51	4	5	14	4
	Temporal	Heschl's gyrus/superior temporal gyrus/planum temporale/ planum polare	Right	49	−8	5	20	5
			Left	−49	−12	5	81	11

We report regions where activity was significantly different from baseline ($P < 0.05$ Bonferroni corrected across participants), for at least one-third of the time window of interest, for at least three participants, as assessed from the obtained posterior probability maps (see 'Materials and methods' section for details). Coordinates correspond to the vertex with maximal amplitude within each region (coordinates are in MNI space). For each vertex, amplitude data of the window of interest were averaged across all participants and tasks to determine the coordinates of the cortical vertex showing the highest peak amplitude.

region with 'Group' as between-participants factor and 'Task' as within-participant factor. To correct for multiple comparisons, only effects lasting for >15 ms (i.e. 9 consecutive samples) were considered significant (Guthrie and Buchwald, 1991; Caclin *et al.*, 2008). For the S2 analyses, two sample *t*-tests were performed to compare amusic and control groups with the same duration criterion as for S1 (i.e. effects lasting >15 ms).

To estimate the peak latency of the P50m (for the first tone) and N100m (for the first tone and for the tones 2 to 6), we extracted the latency of the maximum of absolute amplitude of the source activities in the 25–70 ms and 90–160 ms time windows, respectively. For the P50m of the first tone, the latency values of the maximum amplitude were analysed with a $2 \times 2 \times 2$ ANOVA, with Group as between-participants factor (amusics, control subjects), and Task (contour task, transposition task) and Hemisphere (right, left) as within-participant factors. For the N100m, latency values were analysed with a $2 \times 2 \times 2 \times 2 \times 2$ ANOVA with Tone Rank (tone 1, tones 2 to 6), region of interest (Heschl's gyrus or inferior frontal gyrus), as additional within-participant factors. The latency analysis focused on the vertices that were significantly different from baseline for both the first tone and the tones 2 to 6 (resulting in two sets for each hemisphere: a set of three vertices with a peak at coordinates in Heschl's gyrus; and a set of two vertices at coordinates in the pars opercularis of the inferior frontal gyrus). We did not perform latency analyses of the brain responses associated with the changed tone of S2 (contour task) as they tended to present several local maxima within the time window of interest.

Dynamic causal modelling of transient responses in S1

To test whether group differences in event-related responses associated with the encoding of tones in S1 could be explained by changes in effective connectivity between sources, we used a dynamic causal modelling approach as implemented in SPM8 (David and Friston, 2003; David *et al.*, 2006; Garrido *et al.*, 2007; Kiebel *et al.*, 2009). Dynamic causal modelling attempts to explain event-related potentials using a network of interacting cortical sources and waveform differences in terms of coupling changes among sources. We focused our analyses on the N100m component in response to tones 2 to 6. We excluded tone 1 for which no major

difference could be observed between control subjects and amusics in the first steps of the analyses.

To characterize the two participant groups with a high signal-to-noise ratio and to investigate the putative differences in effective connectivity between them, we compared and modelled the difference between the grand average data of the control subjects (average over the nine participants and the two tasks) and the grand average data of the amusics (average over the nine participants and the two tasks).

All the models we compared were based on the same network architecture, which was motivated by (i) the results of our classical source reconstruction analysis of the N100m component revealing sources in a bilateral fronto-temporal network (see Table 1 for regions of interest that were significantly different from baseline); and (ii) the hypothesis of impaired fronto-temporal connectivity and interhemispheric connectivity observed in amusia with both functional (Hyde *et al.*, 2011) and anatomical approaches (Loui *et al.*, 2009). We thus assumed four sources, modelled as equivalent current dipoles, over left and right primary auditory cortices (A1), left and right pars opercularis of the inferior frontal gyrus (Table 1). Using these sources, we constructed the following dynamic causal model: an extrinsic input entered bilaterally to the primary auditory cortices (A1), which were connected to their ipsilateral inferior frontal gyrus. Interhemispheric (lateral) connections were placed between left and right A1. All connections were reciprocal (i.e. connected with forward and backward connections or with bilateral connections; Fig. 8B).

Given this network architecture, we used a factorial design and a family inference approach to assess various types of connection modulations to explain the group difference in auditory evoked responses (control subjects defined the baseline). Factor 1 refers to the modulation of intrinsic connectivity in bilateral auditory cortices [and includes two families (or levels), corresponding to models where these intrinsic connections were modulated or not between the two groups]. Factor 2 refers to the modulation of lateral connections between the two auditory cortices (two families: models which included modulation or not). Factor 3 refers to the type of connections between auditory and frontal areas that are modulated (namely forward, backward, or both forward and backward connections) or not (four families). Factor

4 refers to the hemispheric location of the above modulated connections, either in the right hemisphere, the left hemisphere, or both (three families). We thus fitted and compared 48 models. Assuming equal prior probabilities over models, we used Bayesian model selection to compare model families (Penny et al., 2010). This rests upon the free energy (or approximate marginal likelihood or evidence) for each model, and yields a posterior probability associated with each model family.

Sustained evoked response analyses (S1)

For the analyses of sustained evoked responses, we used artefact-corrected data for the three amusic participants with respiratory or heartbeat artefacts. These analyses were performed within a 1000 ms time window during S1 (500–1500 ms after the onset of the first tone) as this window covered the maximum amplitude of the responses in all participants and avoided the rising slope and the descending slope of the sustained response. Sensor topography indicated a distribution that was consistent with bilateral sources in the auditory cortices during S1. Analyses were performed at the sensor level because of complications for source reconstruction of the artefact-corrected data of the two amusic participants with artefacts at the respiratory frequency (because of residual artefacts in the most frontal sensors). For each hemisphere, the three sensors with the highest positive amplitude and the three sensors with the highest negative amplitude within the 500 to 1500 ms time window were selected. Thereby, negative responses in the right hemisphere corresponded with the right temporal site, and positive responses to the right frontal site, and conversely for the left hemisphere (negative/frontal site; positive/temporal site) (Fig. 9). To determine whether there were group differences in the amplitude of the sustained responses during S1, we computed the absolute values of the mean of these event-related fields over the 500 to 1500 ms time window. These amplitude values were analysed with a $2 \times 2 \times 2 \times 2$ ANOVA, with Group as between-participants factor (amusics, control subjects), and Task (contour task, transposition task), Hemisphere (right, left) and Sensor site (frontal, temporal) as within-participant factors.

Sensor-level analysis and source reconstruction of induced gamma activity (retention period)

We were interested in investigating the potential group and task differences in oscillatory activities in the gamma-band induced by the musical information during memory maintenance. Because neither latency nor frequency of these oscillatory bursts was known a priori, we first assessed sensor-level data using a method that preserved both types of information: the time-frequency representation based on a wavelet transform of the signals (Tallon-Baudry et al., 1996). The MEG signal was convoluted by complex Morlet's wavelets having a Gaussian shape both in the time domain (SD σt) and in the frequency domain (SD σf) around its central frequency f0. The wavelet family was defined by (f0/σf) = 10, with f0 ranging from 24 to 80 Hz in 2 Hz steps. The time–frequency wavelet transform was applied to each trial after subtraction of the evoked response at each MEG sensor and then averaged across trials, resulting in an estimate of oscillatory power at each time sample and at each frequency between 24

and 80 Hz. For all participants, except the one amusic participant with the heartbeat artefact, original MEG signals were used for this analysis (corrected data were used for the amusic with the heartbeat artefact). We investigated oscillatory gamma activities during the retention period, with respect to a prestimulus baseline, by subtracting for each frequency the mean power computed over the period defined as -1000 to 0 ms preceding the first tone of S1.

To identify the maximal increase of induced power within the gamma band relative to baseline, Wilcoxon tests were computed for each frequency and each time sample across all participants to compare the mean power (average of both tasks) with the mean power in the prestimulus baseline (defined between -1000 and 0 ms before the onset of the first tone of S1). This analysis allowed us to select the 30–40 Hz frequency band over a 1000 ms time window covering the central part of the retention period (2000 to 3000 ms).

To localize the neural sources in the frequency band of interest and explore the potential differences between tasks and groups within the retention period, we applied an adaptive spatial filtering or beamforming technique [Dynamic Imaging of Coherent Sources (DICS) as implemented in Fieldtrip] (Gross et al., 2001; Bouet et al., 2012). Based on individual MRIs (except for the amusic participant who did no undergo the T_1-MRI scan), a grid of $5 \times 5 \times 5 \text{ mm}^3$ volume elements covering the whole brain was constructed, and the source power was computed for each of those elements. For each source location, a linear spatial filter was computed that passes activity from that location with unit gain while maximally suppressing activity from other sources. This filter depended on the signal spectral properties extracted for each participant through a fast Fourier transformation.

For each trial, we estimated the power in the 30–40 Hz frequency band at each volumetric element for both active (retention) and baseline periods, and we computed the log ratio between these two power values to determine the extent of the emerging gamma oscillations. After normalization of the brain in the MNI space, statistical comparisons were performed to compare groups for each task, and to compare tasks for each group with two sided t-tests corrected for multiple comparison using cluster-level statistics, as implemented in Fieldtrip.

Correlations between behavioural, functional and anatomical data

Five types of analyses were performed: (i) correlation between the behavioural data in the short-term memory task (Hit-False-Alarms, reaction times) and data from the pretests (Montreal Battery of Evaluation of Amusia, pitch discrimination threshold); (ii) correlation between the different measures of the behavioural data in the short-term memory task; (iii) correlation between the different measures (white matter, grey matter concentrations) of the VBM data; (iv) correlation between the different measures of the MEG data (event-related fields amplitude and latency, gamma oscillations; connectivity values); and (v) correlations between behavioural, functional and anatomical data. To explore differences explained not only by the group differences defined by the Montreal Battery of Evaluation of Amusia (Peretz et al., 2003), we present here only correlations that were significant over all

participants and for at least one of the two groups separately (amusic group or control group). Note that, except for the VBM data, none of the correlations met these criteria.

Results

## Behavioural data

Performance was significantly above chance (i.e. 0% of Hits–False Alarms), for each condition in each group (*t*-tests, all *P*'s < 0.0001). Hits-False Alarms (Fig. 2A) were analysed with a 2 × 2 ANOVA with Group (amusics, control subjects) as between-participants factor and Task (contour task, transposition task) as within-participant factor. The main effects of Group [$F(1,16) = 74.82, P < 0.0001$], and Task [$F(1,16) = 300.08, P < 0.0001$] were significant, as well as their interaction [$F(1,16) = 119.15, P < 0.0001$]. Fischer's LSD *post hoc* tests indicated that the performance of the amusic group was lower than that of the control group for the contour task ($P < 0.0001$), but not for the transposition task ($P = 0.28$). All amusic participants exhibited a deficit in the contour task in comparison with the control participants (see individual data in Fig. 2A). This is in line with previous results (Gosselin *et al*. 2009; Tillmann *et al*., 2009), suggesting a short-term memory deficit for tone sequences in congenital amusia. It is noteworthy that there was no overlap in performance between the two groups in the contour task, despite a substantial overlap between groups in pitch discrimination thresholds (see 'Materials and methods' section).

For the contour task, to assess the potential effect of pitch interval size in the changed tone of S2 (relative to the corresponding tone in S1) on performance, percentages of correct responses for different trials were extracted as a function of interval sizes. Within 96 different trials of the contour task, there were 36.46% of trials with a change of small interval size (with equal proportions of 1.5 tone, 2 tones and 2.5 tones); 35.41% of trials with a

change of medium interval size (3 tones, 3.5 tones and 4 tones) and 28.13% of trials with a change of large interval size (4.5 tones, 5 tones, 5.5 tones, 6 tones). These data were analysed with a 2 × 3 ANOVA with Group (amusics, control subjects) as between-participants factor and Interval (small, medium, large) as within-participant factor. The main effects of Group [$F(1,16) = 75.37, P < 0.0001$], and Interval [$F(2,32) = 32.38, P < 0.0001$] were significant, as well as their interaction [$F(2,32) = 4,36, P = 0.02$]. *Post hoc* tests indicated that the performance of the amusic group was lower than that of the control group for all interval sizes (all *P*'s < 0.01). Moreover, in amusics, increasing performance was observed with increasing interval changes (all *P*'s < 0.01 for small/medium, medium/large, and large/small comparisons). This increased performance was also observed in control subjects (all *P*'s < 0.05), except for the medium/large comparison ($P = 0.47$), which was probably due to ceiling performance.

Reaction times for correct responses (relative to the end of S2) were analysed by a 2 × 2 × 2 ANOVA with Group (amusics, control subjects) as between-participants factor, and Task (contour task, transposition task) and type of trial (same, different) as within-participant factors. The ANOVA revealed a significant main effect of Task [$F(1,16) = 19.63, P < 0.001$], with shorter reaction time for the transposition task (mean = 479.18 ms, SD = 49.19 ms) than for the contour task (mean = 585.37 ms, SD = 61.21 ms), and a significant interaction between task and group [$F(1,16) = 7.60; P = 0.01$]. Fischer's LSD *post hoc* tests revealed that only for the amusic group ($P < 0.001$), reaction times were longer for the contour task (mean = 629.84 ms, SD = 39.94 ms) than for the transposition task (mean = 479.71 ms, SD = 25.55 ms). This was not the case for control subjects (mean = 529.85 ms, SD = 77.84 ms for the contour task and mean = 489.74 ms, SD = 66.80 for the transposition task, $P = 0.25$). Moreover, the analysis revealed a significant interaction between task (contour task, transposition task) and type of trial [$F(1,16) = 5.36; P = 0.01$]; only for the transposition task
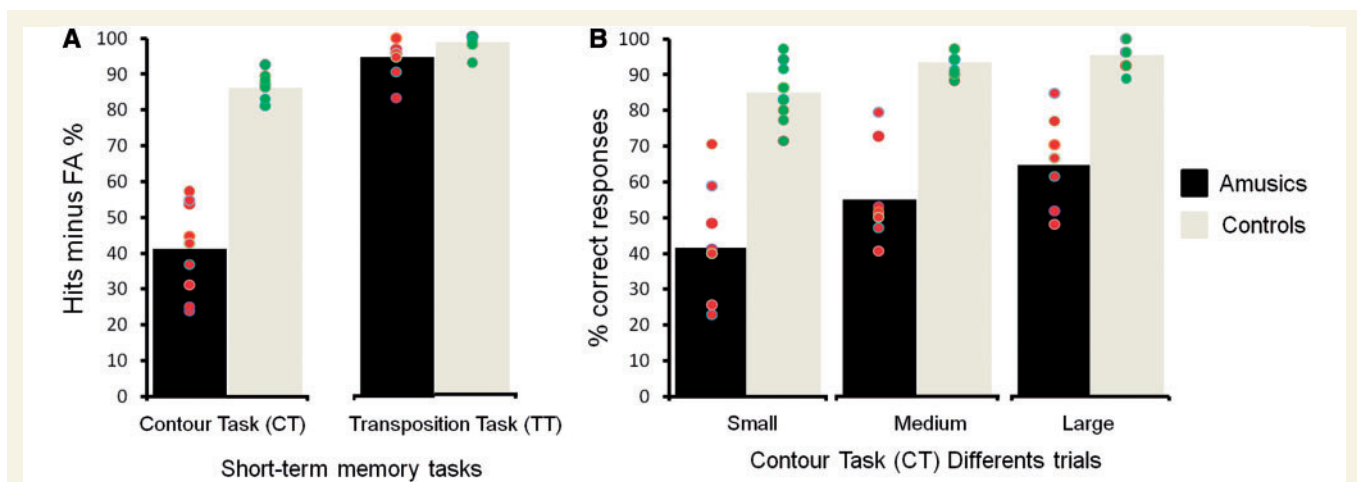
Figure 2 (**A**) Performance of amusic and control groups (grey = control subjects; black = amusics) in terms of Hits minus False Alarms (FA), presented as a function of the task (contour task, transposition task) and group (amusics *n* = 9; control subjects *n* = 9). (**B**) Performance of amusic and control groups for the contour task in terms of per cent of correct responses for different trials as a function of interval size (small, medium, large). Green circles: control subjects' individual performance; red circles: amusics' individual performance.

($P = 0.04$), reaction times were shorter (mean = 451.51 ms, SD = 46.76 ms) for different trials than for same trials (mean = 506.59 ms, SD = 51.93 ms). Analyses of reaction times for correct responses as a function of interval size for the contour task [with a 2×3 ANOVA with Group as between-participants factor and interval size (small, medium, large) as within-participant factor] revealed no significant effects of group ($P = 0.91$) or interval size ($P = 0.14$) and no interaction between them ($P = 0.35$).

Subjective reports

Differences were observed in participants' reports after the session: while all control subjects consistently reported to repeat the melody back in their mind during the delay (eight control participants by internal singing and one by visual mental imagery), amusics did not report to have elaborated any particular strategy (three amusic participants reported to have tried at the beginning to repeat S1 in their mind during the delay, but stopped doing so because of its inefficiency).

Voxel-based morphometry

In the right inferior frontal gyrus, in line with previous findings (Hyde *et al.*, 2006, 2007), between-group differences were found in both grey and white matter (Fig. 3). The analyses of grey matter showed larger concentrations in the right inferior frontal gyrus (bordering the right middle frontal gyrus) in amusics than in control subjects [maximum peak at $x = 36$, $y = 46$, $z = 6$; $t(14) = 2.68$, $P = 0.009$; K_E (cluster extent) = 192 mm^3]. The maximal difference was <1 cm away from that observed with VBM in the two independent samples of amusic and control participants in Hyde *et al.* (2006) ($x = 37$, $y = 42$, $z = -2$ and $x = 40$, $y = 43$, $z = -2$, respectively). Conversely, less white matter concentration in amusics than in control subjects was observed in the right inferior frontal gyrus [in particular in a cluster with a maximum peak at $x = 40$, $y = 30$, $z = 28$; $t(14) = 4.37$, $P < 0.0001$; $K_E = 1328$ mm^3, which encompassed an additional maximum at $x = 36$, $y = 44$,

$z = -6$, close to the coordinates previously reported in Hyde *et al.* (2006): $x = 37$, $y = 42$, $z = -2$ and $x = 42$, $y = 44$, $z = -2$]. This inverse relationship between white and grey matter concentrations in amusics' right inferior frontal gyrus is in line with Hyde *et al.* (2006, 2007). In symmetrical regions of the inferior frontal gyrus in the left hemisphere, between-group differences were observed for white matter only [maximum peak at $x = -34$, $y = 40$, $z = 2$; $t(14) = 3.71$, $P = 0.001$; $K_E = 5344$ mm^3], with less white matter concentration for amusic participants.

The right superior temporal gyrus showed less grey matter concentration in amusic participants than in control subjects (maximum peak of the cluster at $x = 44$, $y = -16$, $z = -18$; $t(14) = 3.86$, $P = 0.001$; $K_E = 17\,936$ mm^3; additional peak in the right superior temporal gyrus at $x = 60$, $y = -14$, $z = 6$). This result contrasts with cortical thickness data (Hyde *et al.*, 2007), showing more grey matter concentration in the right auditory cortex of amusics than of control subjects at approximately the same location (they reported a maximal difference at $x = 60$, $y = -12$, $z = 5$). In the vicinity of the grey matter difference [maximum peak at $x = 44$, $y = -8$, $z = 16$; $t(14) = 3.19$, $P = 0.003$; $K_E = 5520$ mm^3], we also observed less white matter concentration in amusic participants than in control subjects. In symmetrical regions of the left hemisphere, we observed only small between-group differences in the white matter (in two clusters of 464 and 760 mm^3, respectively). Note that there were no previous reports of white matter anomaly in the superior temporal gyrus for congenital amusia.

The correlation between grey matter concentration and white matter concentration was not significant for the right inferior frontal gyrus, [$r(14) = -0.01$, $P > 0.69$], but was significant for the right superior temporal gyrus for all participants [$r(14) = 0.94$, $P < 0.01$], as well as for amusic participants [$r(6) = 0.87$, $P < 0.01$] and control participants [$r(6) = 0.96$, $P < 0.01$], separately. In line with the abnormal lack of fibre connectivity along the right arcuate fasciculus (connecting right inferior frontal gyrus and
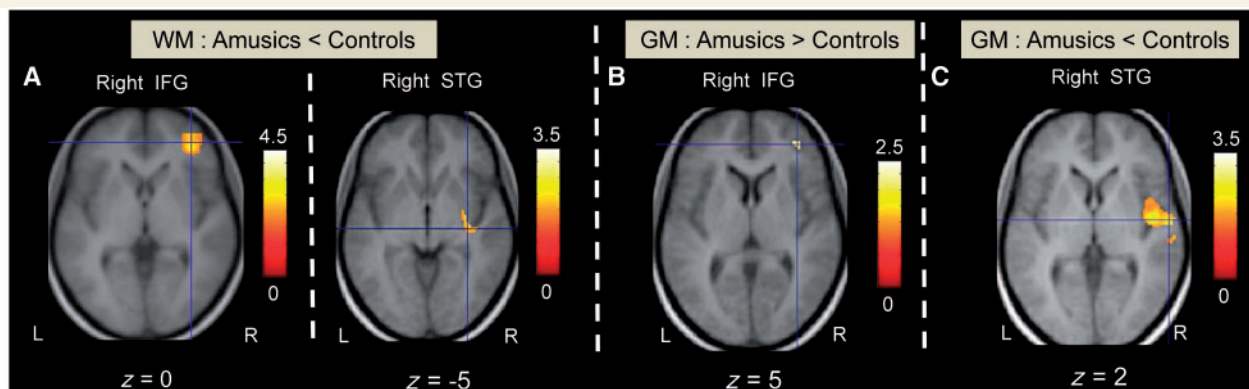
Figure 3 VBM group comparison of white matter (WM) and grey matter (GM) concentration differences: Each brain image corresponds to a statistical map ($P < 0.05$, uncorrected, T-threshold = 1.76) superimposed on the average anatomical MRI of all participants (R = right; L = left). (**A**) Control participants showed more white matter concentration relative to amusics participants in the right inferior frontal gyrus (IFG) and right superior temporal gyrus (STG). (**B**) Amusic participants showed more grey matter concentration relative to control subjects in the right inferior frontal gyrus. (**C**) Control participants showed more grey matter concentration relative to amusics participants in the right superior temporal gyrus.

right superior temporal gyrus) described by Loui *et al.* (2009) in the amusic brain, we observed a positive correlation between white matter concentration of right inferior frontal gyrus and white matter concentration of right superior temporal gyrus. This correlation was observed over all participants [$r(14) = 0.79$, $P < 0.01$] but was significant only for the amusic participants [$r(6) = 0.81$, $P < 0.01$] and marginally significant for the control participants [$r(6) = 0.62$, $P = 0.09$].

Magnetoencephalography

Auditory evoked responses were observed after each tone in all participants. The sensor plots indicated clear P50m, N100m and sustained evoked responses during S1 and S2. The topography of these responses was consistent with bilateral sources in the auditory cortices (Fig. 4).

Encoding of S1 melodies: transient evoked responses (P50m and N100m)

P50m and N100m represent the activity of generators located in the auditory and frontal cortices (Näätänen and Picton, 1987; Alcaini *et al.*, 1994; Giard *et al.*, 1994; Pantev *et al.*, 1995;
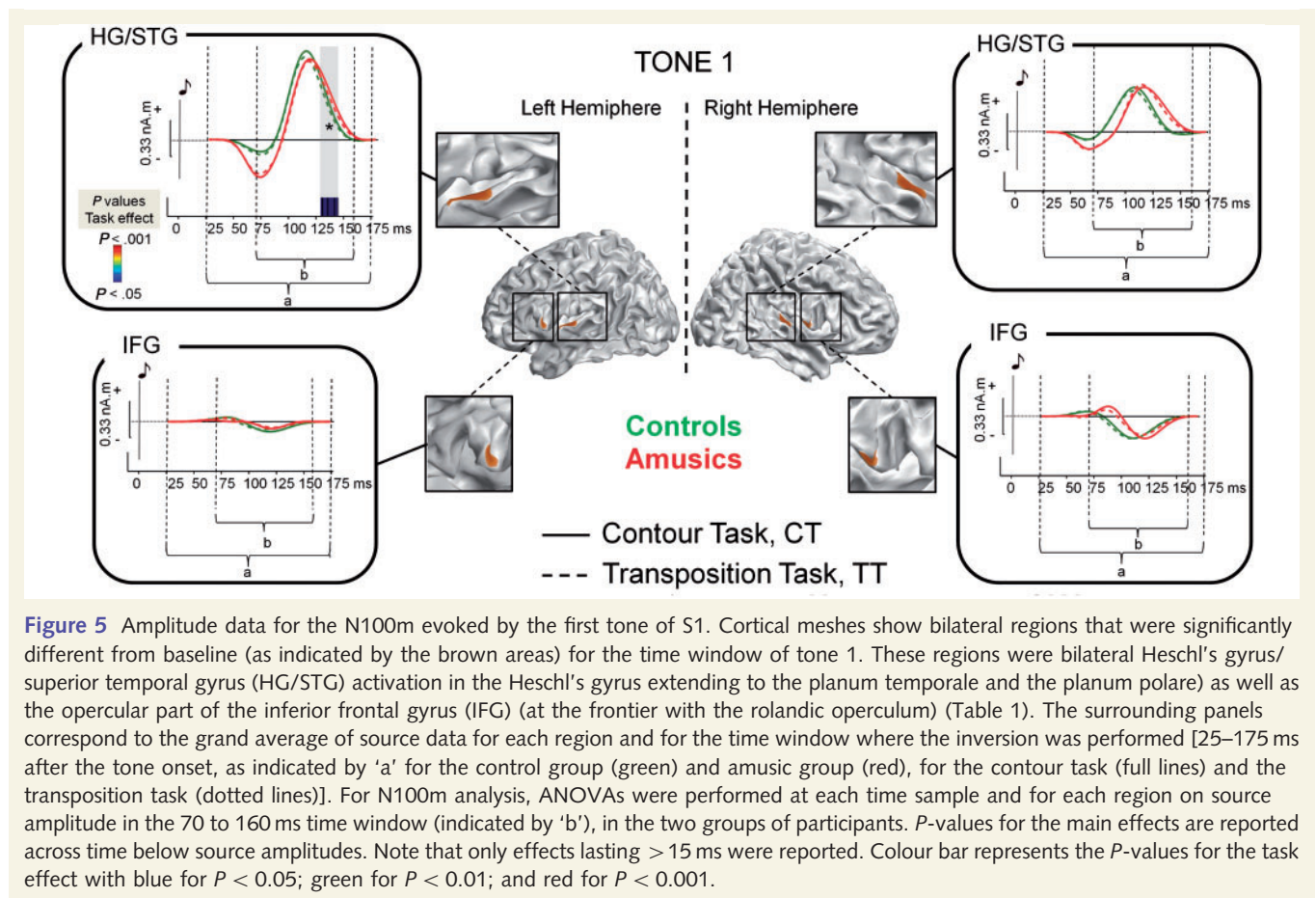
Yvert *et al.*, 2001). Typically, the evoked response for the first tone of a sequence is different and larger than that of subsequent tones (Fig. 4A and B), not only because of neural refractoriness, but also because the first tone can be considered as an 'infrequent/relatively unexpected' orienting stimulus because it is presented after a period of silence. In contrast, subsequent tones in a sequence (here, tones 2 to 6, presented without interstimulus interval) can be considered as 'frequent/expected' stimuli. The first tone should thus recruit more 'orienting' (non-specific) components (Alcaini *et al.*, 1994; Demarquay *et al.*, 2011) than tones 2 to 6. In comparison with the first tone, these subsequent tones should recruit more strongly the cortical areas implicating high-level processing or memory representation (relative to the recruitment of non-specific components). Source analyses thus explored the generators of P50m and N100m as a function of group and task, separately for the first tone and the subsequent tones of S1.

Source analyses of transient responses in S1 (P50m and N100m): regions of interest

For the P50m of tone 1, activity was significantly different from baseline in bilateral auditory regions (Heschl's gyrus, left hemisphere, $x = -48$, $y = -18$, $z = 4$; cluster surface = 53 mm^2; right
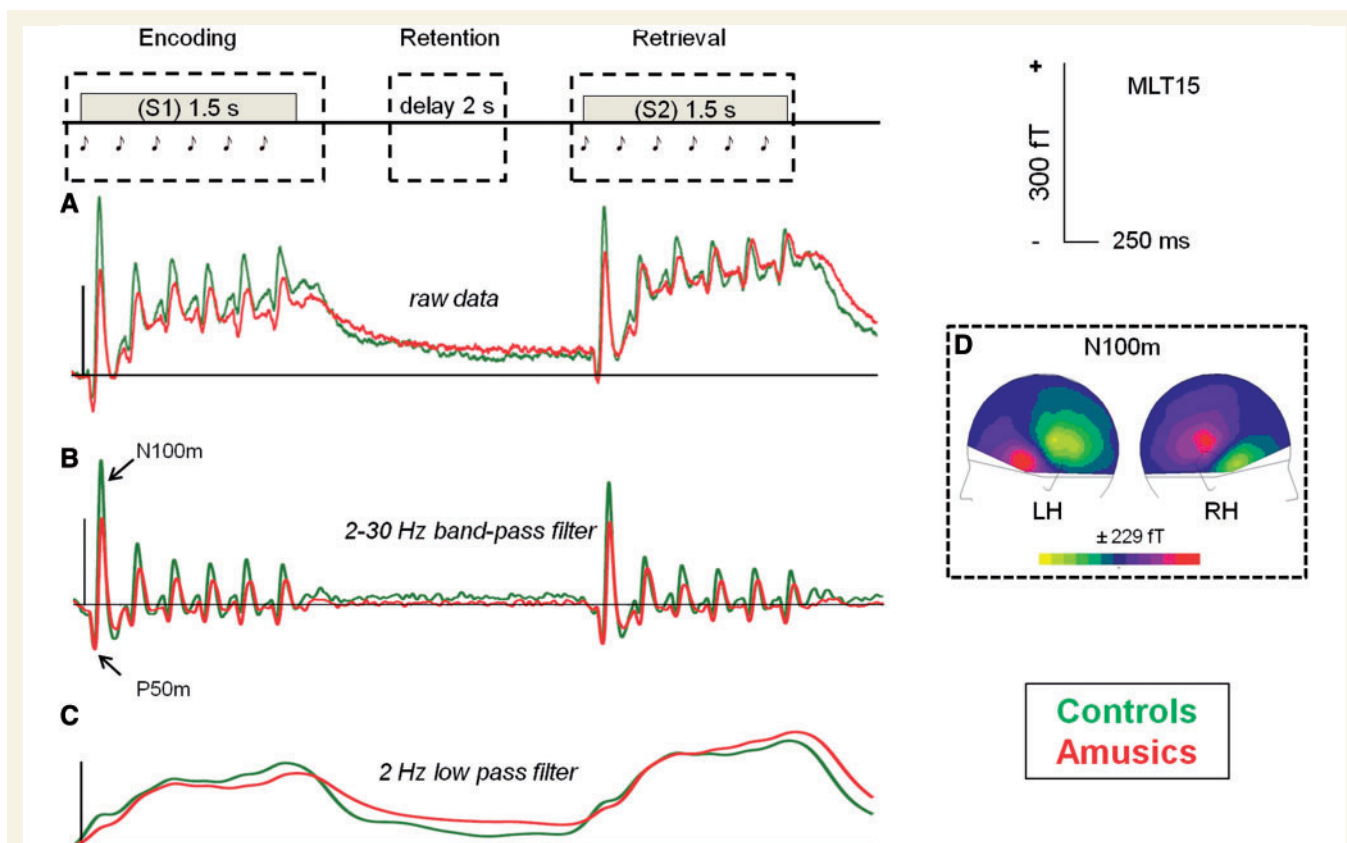
Figure 4 Grand average of a left temporal MEG sensor (MLT15, 'MEG left temporal') for a trial time window (-100 to 5500 ms), for the control group (green) and amusic group (red), collapsed across conditions (contour task, transposition task). (**A**) Event-related fields based on original MEG signals. (**B**) Event-related fields filtered with a 2–30 Hz band-pass filter to uncover P50m and N100m responses. (**C**) Event-related fields filtered with 2 Hz low-pass filter to uncover sustained evoked responses. (**D**) Sensor plot of the mean event-related fields in the 90–140 ms time-window for the first tone, averaged across groups and tasks. LH = left hemisphere; RH = right hemisphere.

Figure 5 Amplitude data for the N100m evoked by the first tone of S1. Cortical meshes show bilateral regions that were significantly different from baseline (as indicated by the brown areas) for the time window of tone 1. These regions were bilateral Heschl's gyrus/ superior temporal gyrus (HG/STG) activation in the Heschl's gyrus extending to the planum temporale and the planum polare) as well as the opercular part of the inferior frontal gyrus (IFG) (at the frontier with the rolandic operculum) (Table 1). The surrounding panels correspond to the grand average of source data for each region and for the time window where the inversion was performed [25–175 ms after the tone onset, as indicated by 'a' for the control group (green) and amusic group (red), for the contour task (full lines) and the transposition task (dotted lines)]. For N100m analysis, ANOVAs were performed at each time sample and for each region on source amplitude in the 70 to 160 ms time window (indicated by 'b'), in the two groups of participants. P-values for the main effects are reported across time below source amplitudes. Note that only effects lasting >15 ms were reported. Colour bar represents the P-values for the task effect with blue for $P < 0.05$; green for $P < 0.01$; and red for $P < 0.001$.

hemisphere, $x = 51$; $y = -11$; $z = 4$; cluster surface $= 25 \text{ mm}^2$). For the N100m, bilateral fronto-temporal regions with activity significantly different from baseline for tone 1 and for the average of tones 2 to 6 are presented in Table 1, together with the coordinates of emergent vertices for the highest peak amplitudes. Note that for the auditory regions, the vertex of maximum amplitude was located in Heschl's gyrus, but that the activation further extended to the planum temporale, planum polare and superior temporal gyrus. Note that, when we performed the same PPM analyses for each group separately, the same regions with the same peaks were emergent in control subjects and amusics, but with a smaller extent in the amusic group.

Source analyses of transient responses in S1 (P50m, N100m): amplitude

For P50m (tone 1), no significant effects or interaction were observed. Note that there were no differences between contour task and transposition task; and that we did not attempt to analyse the P50m of tones 2 to 6 because their amplitudes were rather small as a consequence of refractoriness (P50-gating), and because of the overlap with the end of the components evoked by the preceding tones.

For N100m (tone 1), whereas there was no significant effect of group, the main effect of task was significant: both amusics and

control subjects showed larger activation in the contour task in comparison to the transposition task in the left auditory cortex (129 to 145 ms).

The ANOVAs for N100m (average of tones 2 to 6) revealed group differences in source amplitude, with higher amplitudes for control participants than for amusic participants in the four regions of interest and the following time windows (Fig. 6): (i) right Heschl's gyrus/superior temporal gyrus (together with planum temporale and planum polare), 80–120 ms; (ii) right inferior frontal gyrus, 105–122 ms; (iii) left inferior frontal gyrus, 107–127 ms; and (iv) left Heschl's gyrus/superior temporal gyrus (together with planum temporale and planum polare), 95–126 ms. Moreover, the inverse difference (amusics > control subjects) was observed in the right Heschl's gyrus/superior temporal gyrus for the 135–153 ms time window. However, this inverse difference for the right Heschl's gyrus/superior temporal gyrus seems to be related to an increased latency of the N100m in the amusic brain (see below). No significant effect of task and no interaction between group and task were observed.

Source analyses of transient responses (P50m, N100m): latency

For P50m (tone 1), there were no significant effects or interaction. In analyses for N100m (tone 1 and average of tones 2 to 6) there
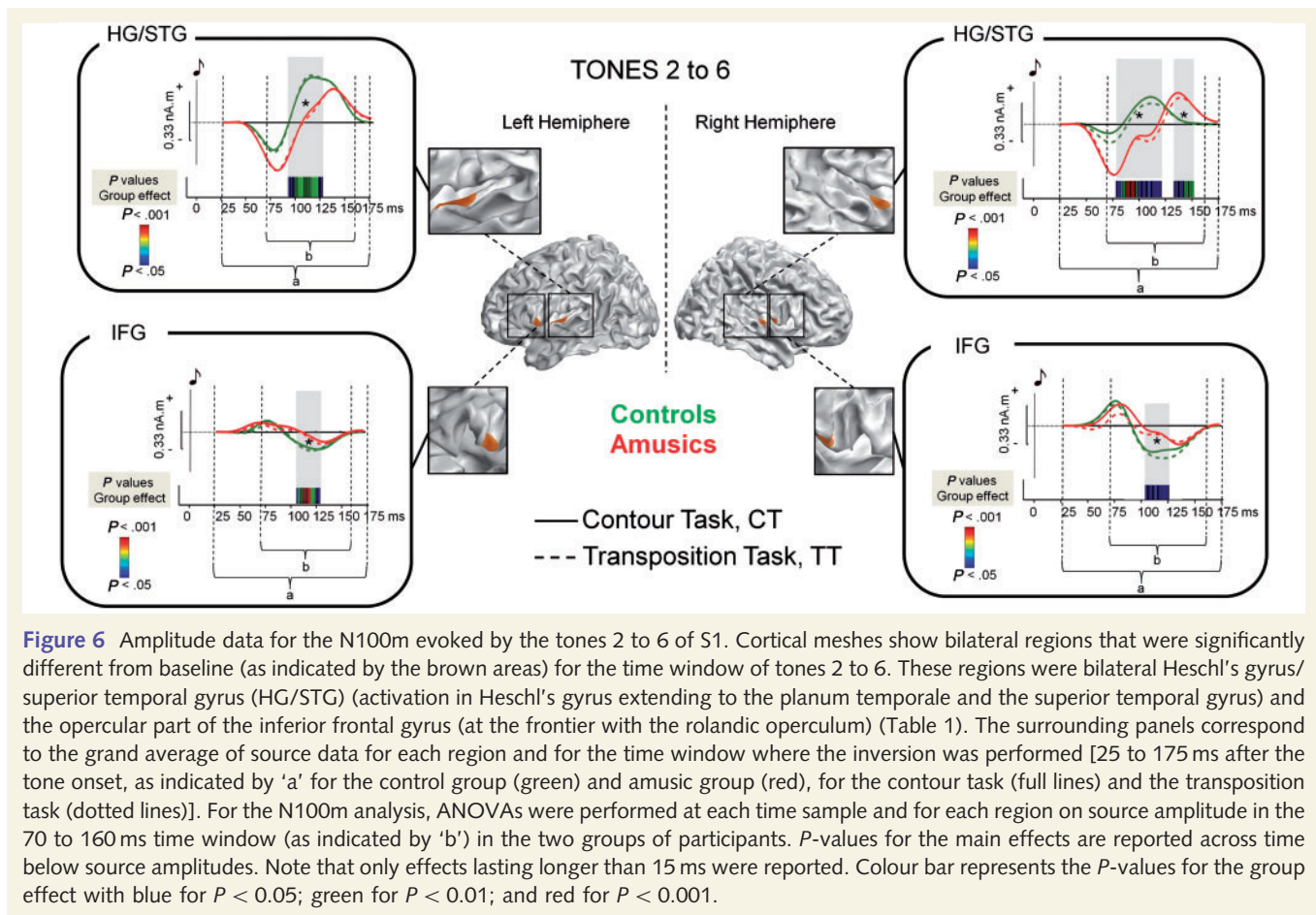
Figure 6 Amplitude data for the N100m evoked by the tones 2 to 6 of S1. Cortical meshes show bilateral regions that were significantly different from baseline (as indicated by the brown areas) for the time window of tones 2 to 6. These regions were bilateral Heschl's gyrus/superior temporal gyrus (HG/STG) (activation in Heschl's gyrus extending to the planum temporale and the superior temporal gyrus) and the opercular part of the inferior frontal gyrus (at the frontier with the rolandic operculum) (Table 1). The surrounding panels correspond to the grand average of source data for each region and for the time window where the inversion was performed [25 to 175 ms after the tone onset, as indicated by 'a' for the control group (green) and amusic group (red), for the contour task (full lines) and the transposition task (dotted lines)]. For the N100m analysis, ANOVAs were performed at each time sample and for each region on source amplitude in the 70 to 160 ms time window (as indicated by 'b') in the two groups of participants. P-values for the main effects are reported across time below source amplitudes. Note that only effects lasting longer than 15 ms were reported. Colour bar represents the P-values for the group effect with blue for $P < 0.05$; green for $P < 0.01$; and red for $P < 0.001$.

were significant main effects of: (i) group [$F(1,16) = 8.54$; $P < 0.01$], reflecting increased N100m latency of source activity for amusics in comparison to control subjects; (ii) tone rank [$F(1,16) = 19.73$; $P < 0.0001$], with increased latency for tones 2 to 6 in comparison with tone 1; and (iii) a marginally significant effect of region [$F(1,16) = 4.3$; $P = 0.054$], with increased latency in the inferior frontal gyrus in comparison with the Heschl's gyrus. The interaction between tone rank and task [$F(1,16) = 4.85$; $P = 0.042$] was significant. For tone 1, the latency did not differ between the two tasks ($P = 0.21$), but for tones 2 to 6, the latency was slightly longer for the contour task than for the transposition task, even though only marginally significantly ($P = 0.087$). In addition, the interaction between tone rank and region was significant [$F(1,16) = 9.33$; $P = 0.007$], and, most importantly, it was modulated by group, as indicated by the three-way interaction between tone rank, region and group [$F(1,16) = 5.17$; $P = 0.03$].

For tone 1 (Fig. 7), the increased latency in bilateral inferior frontal gyrus relative to bilateral Heschl's gyrus was more pronounced for amusics (reflecting a delayed frontal responses, $P < 0.003$) than for control subjects (where the difference was only marginally significant, $P = 0.06$). The latencies of the responses for tones 2 to 6 were similar to the latency of tone 1 (Fig. 7), except that, critically, for amusics, the Heschl's gyrus response for tones 2 to 6 was delayed by ~20 ms relative to control subjects ($P < 0.001$).

## Dynamic causal modelling of the transient responses in S1 for tones 2 to 6

Figure 8A shows the results of the comparisons between model families (exceedance probabilities obtained for each family inference). Figure 8B shows the network architecture of the winning model as well as the conditional estimates of the connection strengths associated with the connections that proved to be significantly modulated to explain the amusic response compared to the control response.

Posterior estimates obtained with the winning model enabled us to conclude that, compared with control subjects, amusic participants showed an abnormally increased lateral connectivity between the two A1, decreased intrinsic modulations in both auditory cortices, and decreased backward connectivity between the right inferior frontal gyrus and the right auditory cortex.

Sustained evoked responses during S1

Few studies have explored the role of the sustained evoked responses in the auditory domain, but some of these data sets have demonstrated that: (i) sustained evoked responses are involved in the cortical representation of behaviourally relevant sounds (Picton *et al.*, 1978; Bidet-Caulet *et al.*, 2007); and (ii) they are modulated by attentional processes (for a review, see Picton *et al.*, 1978). In line with Bidet-Caulet *et al.* (2007) who reported sustained evoked
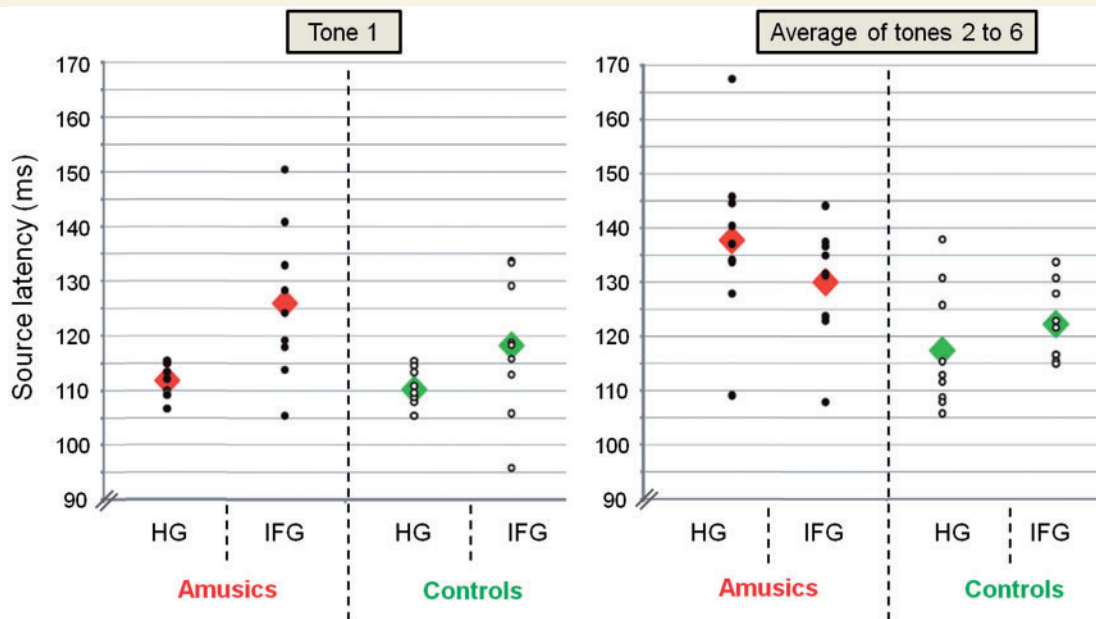
Figure 7 N100m latency (in ms) of amusic and control participants, calculated from the reconstructed source signal [amusics *n* = 9; control subjects *n* = 9 for bilateral Heschl's gyrus (HG) and bilateral inferior frontal gyrus (IFG); Green = control subjects; red = amusics] presented as function of tone rank (tone 1 and average of tones 2 to 6) and region (Heschl's gyrus, inferior frontal gyrus). Diamonds indicate average latency for each group and task (in ms); circles indicate participants' individual latency (in ms). Note that latencies were averaged across tasks (contour task, transposition task).
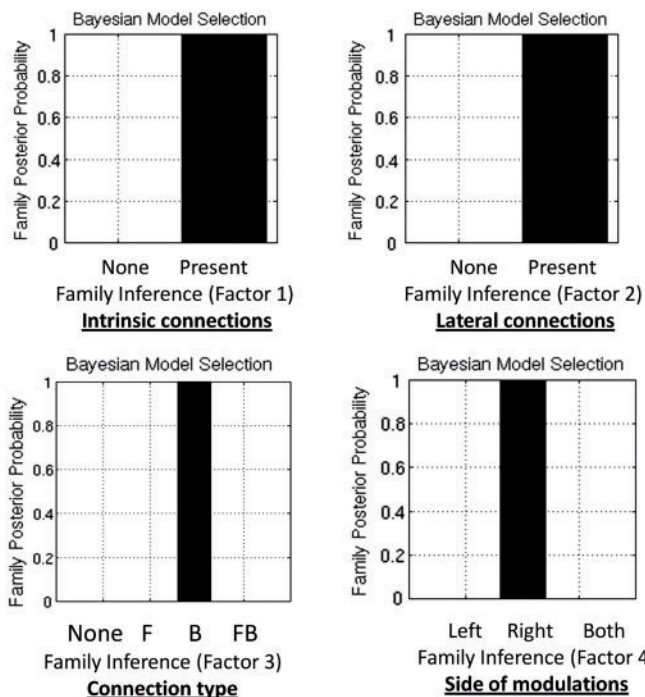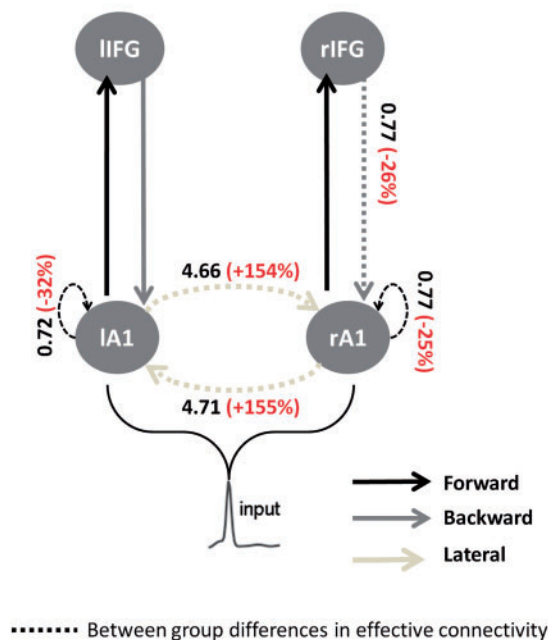

Figure 8 Dynamic causal modelling results. (**A**) Results of the four family-wise inferences. For each inference, the posterior probability of each model family is depicted. For each comparison, the family associated with a high posterior probability (*P* > 0.99) could be retained as the winning family. As each comparison yields a clear winning family, we could identify a winning model. (**B**) Winning model. Dashed arrows indicate modulated connections (i.e. connections that differ between groups) and solid arrows indicate fixed connections. Significant changes in effective coupling between control subjects and amusics are specified (in black: amount of coupling change between groups; in red: corresponding relative coupling with amusics coupling expressed in per cent of control coupling). B = Backward; F = Forward; FB = Forward Backward; lIFG = left inferior frontal gyrus; rIFG = right inferior frontal gyrus; lA1 = left primary auditory cortex; rA1 = right primary auditory cortex.
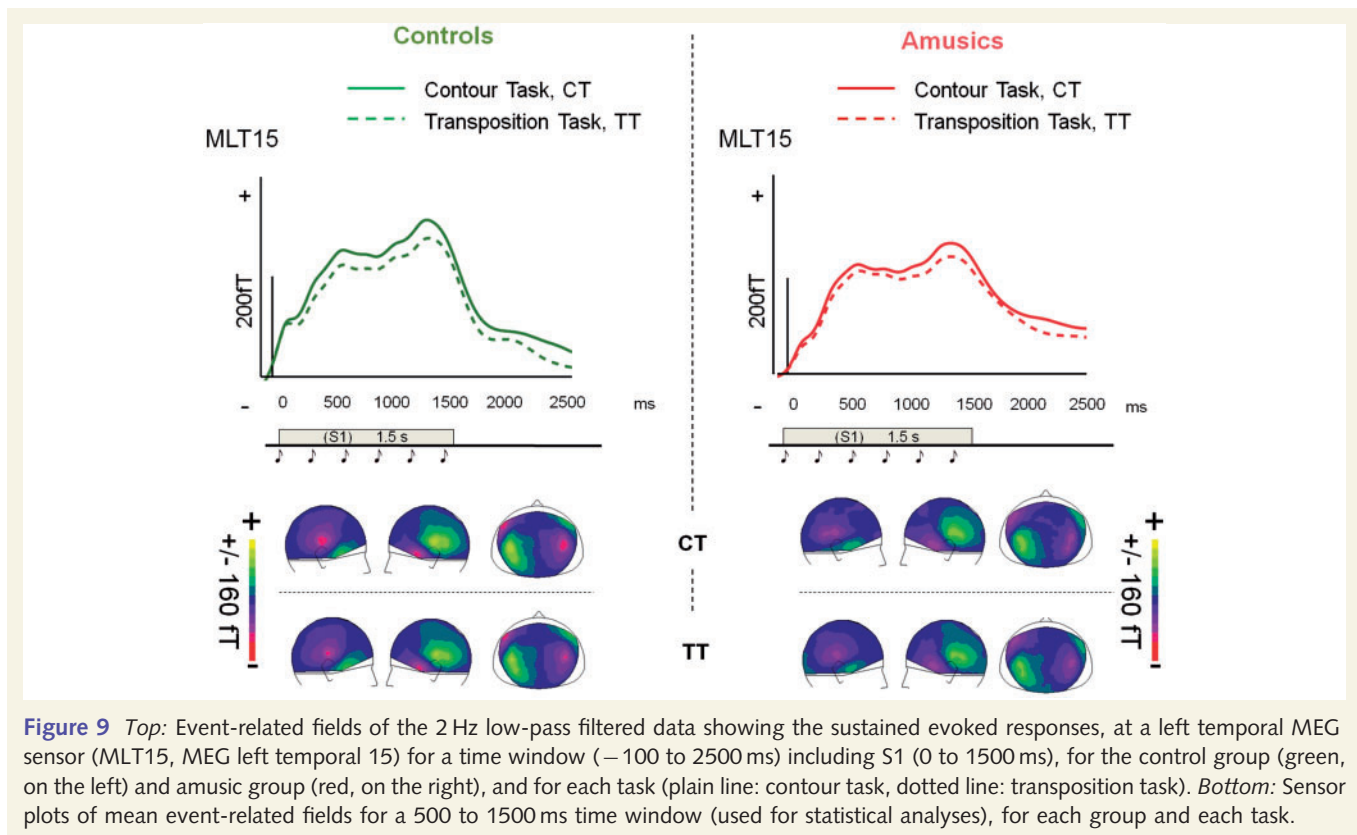
Figure 9 *Top:* Event-related fields of the 2 Hz low-pass filtered data showing the sustained evoked responses, at a left temporal MEG sensor (MLT15, MEG left temporal 15) for a time window (−100 to 2500 ms) including S1 (0 to 1500 ms), for the control group (green, on the left) and amusic group (red, on the right), and for each task (plain line: contour task, dotted line: transposition task). *Bottom:* Sensor plots of mean event-related fields for a 500 to 1500 ms time window (used for statistical analyses), for each group and each task.

responses in secondary auditory areas during long-lasting stimuli, the topographies of sustained evoked responses in the present data indicated bilateral sources in the auditory cortices for all participants. Sensor mean amplitude analyses revealed (i) a significant main effect of sensor site [$F(1,16) = 10.06$; $P = 0.005$], reflecting higher amplitude over temporal sensors in comparison to the frontal sensors; (ii) a significant main effect of hemisphere [$F(1,16) = 8.61$; $P = 0.009$], with higher amplitude in the left hemisphere in comparison with the right hemisphere; and (iii) a marginally significant main effect of task [$F(1,16) = 4.16$; $P = 0.058$], with increased mean amplitudes for the contour task in comparison with the transposition task. There was no significant main effect of group and no interaction implicating the group factor.

Retention of the melodic information: time–frequency analyses of the delay period

Gamma-power was analysed in the 30–40 Hz frequency band over the time window corresponding to the central part of the delay (2000–3000 ms). After source reconstruction, statistical comparisons were performed to compare groups for each task, and tasks for each group using two-sided *t*-tests (corrected for multiple comparisons using cluster-level statistics implemented in Fieldtrip). The results are depicted in Fig. 10.

Group effect

For the contour task, control subjects showed increased gamma synchronization (relative to baseline) in the right dorso-lateral prefrontal cortex [Brodmann area (BA) 9/46; $x = 45$, $y = 31$, $z = 25$;

$K_E = 6795$ mm^3] in comparison with the amusic participants. For the transposition task, amusics showed increased gamma synchronization in the left dorso-lateral prefrontal cortex (BA9/46; $x = −37$, $y = 26$, $z = 29$; $K_E = 7870$ mm^3) and the left temporo-parietal junction (BA39; $x = 46$, $y = −69$, $z = 28$; $K_E = 3430$ mm^3) in comparison with the control participants.

Task effect

For control participants, increased gamma synchronization was observed for the contour task in comparison with the transposition task in the right inferior frontal gyrus (opercular part: $x = 56$, $y = 14$, $z = 20$; $K_E = 4220$ mm^3) and in the left inferior frontal gyrus (opercular part: $x = −59$, $y = 6$, $z = 9$; $K_E = 5590$ mm^3). For the amusic participants, no significant modulations of gamma oscillations with the tasks were observed.

Source analyses of the transient responses of the changed tone in S2

To investigate if amusics' altered encoding and retention of melodies in memory could be further associated to altered retrieval of melodic information, we analysed the brain responses evoked by the changed tone in S2 of different trials (for correct responses). The difference wave (different trial − same trial) observed at the sensor level revealed that for control subjects, the processing of the changed tone was associated to two evoked responses. The first one was elicited ~150 ms after the onset of the changed tone; and the second one peaked at 500 ms after the tone onset (Fig. 11A). This biphasic response was not observed in amusics (Fig. 11B). To compare the
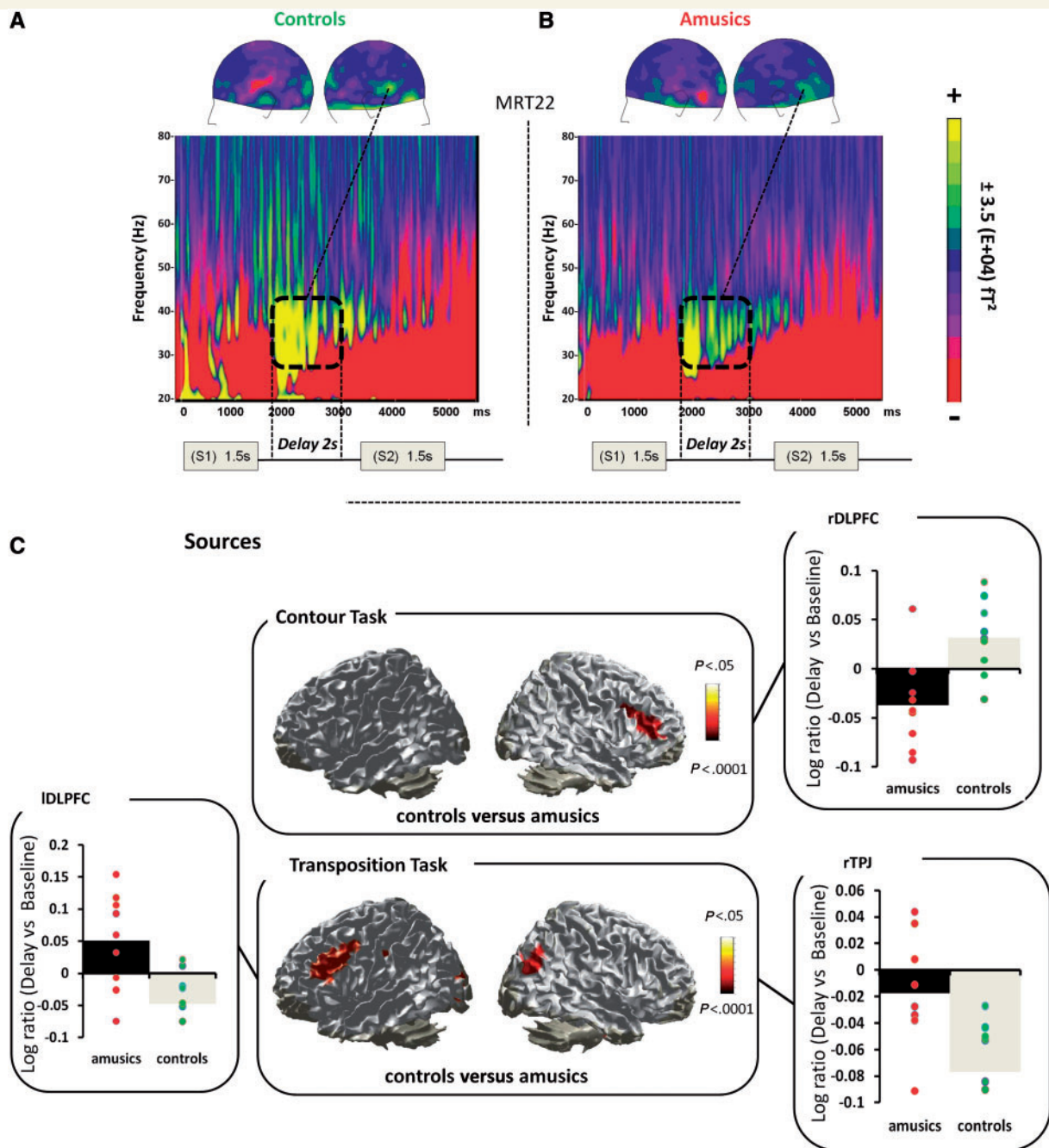
Figure 10 (**A** and **B**) Time–frequency plot of a right temporal MEG sensor (MRT22, 'MEG right temporal') for a trial time window (−100 to 5500 ms) including S1 (0 to 1500 ms), delay (1500 to 3500 ms) and S2 (3500 to 5000 ms) collapsed across conditions (contour task, transposition task), for the control group (**A**) and the amusic group (**B**). The time–frequency power values are plotted after subtraction of the mean power values of the baseline for each frequency. (**C**) Cortical meshes showing the statistics of two sided *t*-tests (corrected for multiple comparisons using cluster-level statistics) for the group comparison for each task (*upper* panel: contour task; *lower* panel: transposition task). *P*-values for the group effect are colour-coded with yellow for *P* < 0.05, red for *P* < 0.01 and black for *P* < 0.001. The surrounding panels correspond to the grand average of source data for each group and for each region. Black bars = amusics; grey bars = control subjects; circles = individual data; rDLPFC = right dorsolateral prefrontal cortex; lDLPFC = left dorsolateral prefrontal cortex; rTPJ = right temporo-parietal junction.

two participant groups for the processing of the changed tone in S2 and to investigate whether the same network was recruited by the two groups, we performed source modelling of the difference waveform (i.e. between different trials and same trials for correct

responses). The bilateral fronto-temporal regions where activity was significantly different from baseline are presented in Table 2. Note that for the auditory regions, the vertex of maximum amplitude was located more laterally in the superior temporal plane than
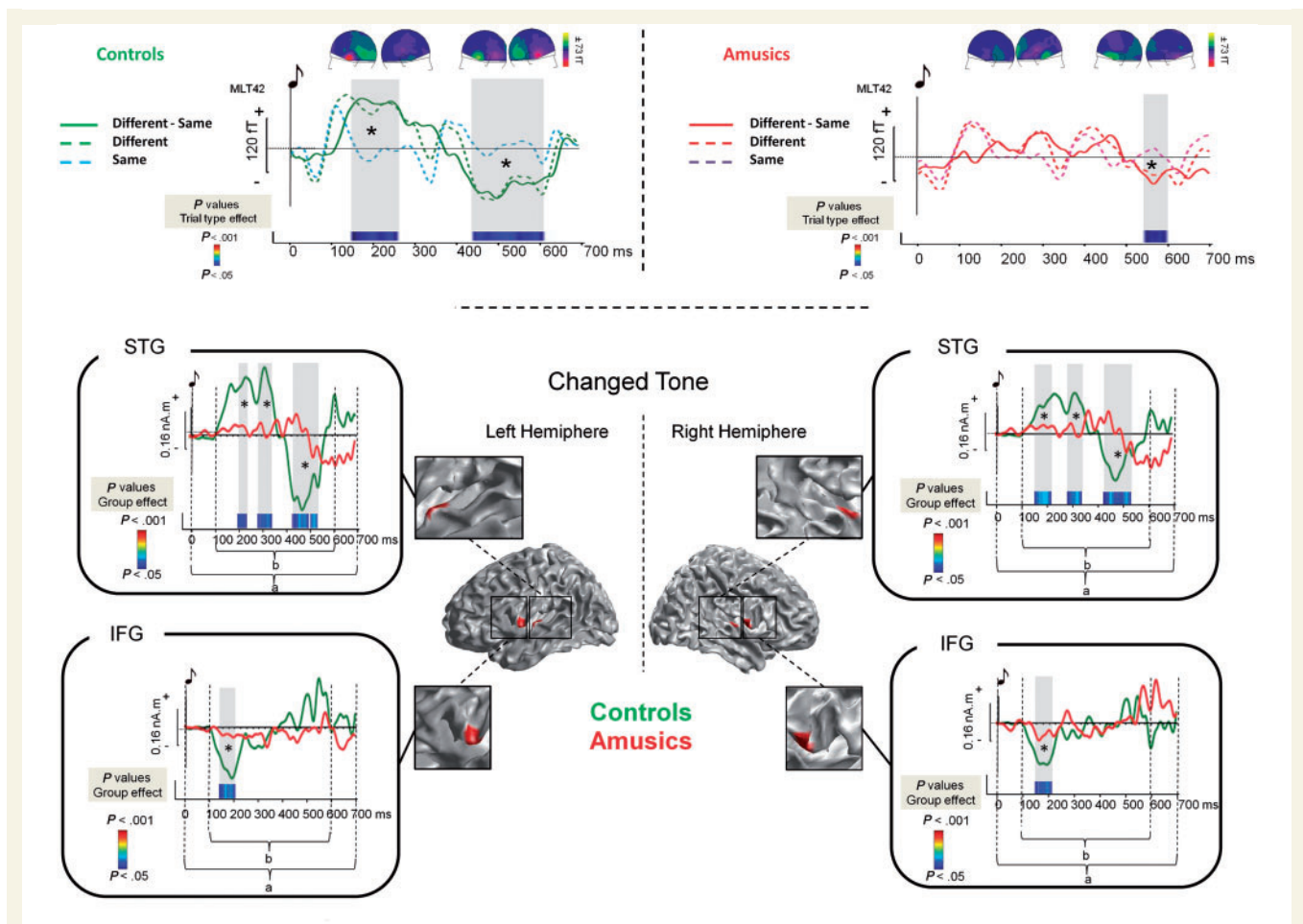
Figure 11 (**A** and **B**) Grand average of a left temporal MEG sensor (MLT42) for a 0 to 700 ms time window after the onset of the changed tone in S2 for the contour task for each group and each type of trial. (**A**) For control subjects, green dotted line = different trials, correct responses; blue dotted line = same trials, correct responses; green plain line = difference wave (different trials − same trials for correct responses). (**B**) For amusics, red dotted line = different trials, correct responses; purple dotted line = same trials, correct responses; red plain line = difference wave (different trials − same trials for correct responses). Two-sample *t*-tests were performed at each time sample on sensor amplitudes in the 0 to 700 time window in the two groups of participants. *P*-values are reported across time in the lower panel with blue for $P < 0.05$, green for $P < 0.01$ and red for $P < 0.001$. Note that only effects lasting >15 ms were reported. (**C**) Source reconstruction of the brain responses specifically evoked by the changed tone in S2. Cortical meshes show bilateral regions that were significantly different from baseline (as indicated by the brown areas). These regions were the bilateral superior temporal gyrus (STG) as well as the opercular part of the inferior frontal gyrus (IFG) (Table 2). The surrounding panels correspond to the grand average of source data for each region and for the time window where the inversion was performed [0 to 700 ms after the changed tone onset, as indicated by 'a' for the control group (green) and the amusic group (red)]. Two sample *t*-tests were performed at each time sample and for each region on source amplitude in the 100 to 600 ms time window (as indicated by 'b') in the two groups of participants. *P*-values are reported across time below the source amplitudes with blue for $P < 0.05$, green for $P < 0.01$ and red for $P < 0.001$. Note that only effects lasting longer than 15 ms were reported.

Table 2 Frontal and temporal generators of the change-specific response within S2

Lobe	Region	Hemisphere	x	y	z	mm²	Number of vertices
Frontal	Inferior frontal gyrus, pars opercularis	Right	55	4	6	49	10
		Left	−54	3	6	65	12
Temporal	Superior temporal gyrus/planum temporale	Right	55	−12	5	182	25
		Left	−52	−12	5	117	16

We report regions where activity was significantly different from baseline ($P < 0.05$, Bonferroni corrected across participants), for at least one-third of the time window of interest, for at least three participants, as assessed from the obtained posterior probability maps (see 'Materials and methods' section for details). Coordinates correspond to the vertex with maximal amplitude within each region (coordinates are in MNI space).

observed for the transient responses in S1. For the frontal generators, the coordinates were close to the coordinates for S1, but the activations were more extended. In addition, note that when we performed the same posterior probability map analyses for each group separately, the four clusters were emergent in control subjects (but only right and left frontal generators were emergent in amusics).

Two sample *t*-tests revealed group differences in source amplitude, with higher amplitudes for control participants than for amusic participants in the four regions of interest and in the following time windows: (i) right superior temporal gyrus, 150–210 ms, 290–340 ms, 425–520 ms; (ii) right inferior frontal gyrus, 160–210 ms; (iii) left inferior frontal gyrus, 150–205 ms; and (iv) left superior temporal gyrus, 190–230 ms, 280–340 ms, 425–530 ms (Fig. 11C).

Discussion

The present study investigated the cerebral correlates of pitch perception and memory in congenital amusics and matched control subjects. Using MEG, we investigated the amusic brain during encoding, retention, and retrieval of melodic information. The major finding is that pitch deficits in congenital amusia can be traced down to early brain responses. During the encoding of a melody (S1), auditory N100m components were observed after each tone for both groups of participants, but were strongly reduced and delayed in amusics. Source reconstruction analyses provided evidence of an altered recruitment of frontal and temporal N100m generators in the amusic brain during the encoding of the S1 melodies. Dynamic causal modelling of the N100m revealed an abnormally increased effective connectivity between the right and left auditory cortices in amusics, reduced intrinsic connections within the bilateral auditory cortices, as well as reduced backward connectivity between the right inferior frontal gyrus and the right auditory cortex in comparison with control subjects. During the retention of the melodic information, gamma oscillations revealed an altered recruitment of the right dorsolateral prefrontal cortex in amusics for the more difficult memory task (contour task). Finally, these altered responses observed in both encoding and retention of the melodic information of S1 were associated to an altered retrieval of the melodic information: amusics showed reduced brain responses elicited by the changed tone of the second melody (S2) in a bilateral fronto-temporal network. As predicted on the basis of previous findings (Hyde *et al.*, 2006, 2007; Loui *et al.*, 2009), we observed brain morphological anomalies in terms of white matter concentration and grey matter concentration in the right inferior frontal gyrus as well as in the right superior temporal gyrus of amusic participants in comparison to control participants. The convergence between functional and structural brain differences provided evidence for abnormalities in a fronto-temporal pathway (including functional anomaly in the auditory cortex) associated with the pitch encoding and short-term memory dysfunctions in congenital amusia.

Short-term memory deficit in congenital amusia

Control participants' performance was high for both the contour task and transposition task. Amusic participants' performance was unimpaired for the transposition task, but was strongly impaired for the contour task (all amusic participants exhibited a deficit in comparison with control participants). Pitch discrimination deficits can be excluded as the sole origin of the impaired performance as several amusics' pitch thresholds were comparable to control subjects' pitch thresholds (see also Tillmann *et al.*, 2009). Additionally, we observed impaired short-term memory in amusics in comparison to control subjects even for changed tones of large interval sizes in the contour task (i.e. the large changes were superior to all amusics' pitch discrimination threshold), thus allowing us to argue that the deficit in short-term memory performance for the contour task was not merely a consequence of increased pitch discrimination threshold in amusics. The impaired performance of amusics in the contour task is in agreement with previously reported deficits in the short-term memory of tones (Gosselin *et al.*, 2009; Tillmann *et al.*, 2009; Williamson and Stewart, 2010; Williamson *et al.*, 2010) and suggests that amusic individuals experience difficulties in maintaining the memory trace of melodic contour information.

Structural abnormalities in the amusic brain

VBM analyses revealed decreased white matter concentration and increased grey matter concentration in the right inferior frontal gyrus (BA 47) of amusics (relative to control subjects). These findings are in agreement with Hyde *et al.* (2006, 2007), who suggested that cortical abnormalities in the amusic brain might occur consequently to an anomaly in cortical development. Amusics' white matter abnormalities are hypothesized to reflect anomalous connectivity between auditory and frontal cortical areas (Hyde *et al.*, 2006, 2007, 2011; Loui *et al.*, 2009). This hypothesis was also supported by the positive correlation between white matter concentration in right inferior frontal gyrus and right superior temporal gyrus in amusic participants of our present study. Moreover, our present data revealed less grey matter concentration in the right superior temporal gyrus for amusics in comparison with control subjects; this result contrasts with cortical thickness data of Hyde *et al.* (2007). This opposite result pattern can be related to the methodological differences between cortical thickness and VBM. In other domains, several studies have demonstrated opposite result patterns in a given brain region when comparing cortical thickness or VBM measures. For example, Park *et al.* (2009) found a significant reduction of grey matter concentration in the primary and associative visual cortices in a group of blind participants using VBM. However, using cortical thickness and surface area measures in the same group of participants, they found a thicker cortex in the same areas. The authors attributed these volumetric atrophies to the decreased cortical surface area despite increased cortical thickness. They proposed that the two measures are complementary and might differently reflect morphological alteration during the developmental period (see Jiang *et al.*, 2009, for converging

evidence of opposite results when comparing data obtained with the two methods). Combining these different approaches of cortical anatomy in amusia remains to be done within the same participants, but nevertheless, the present data agreed with Hyde *et al.* (2007) in revealing structural abnormalities in the right superior temporal gyrus of the amusic brain for grey matter concentrations.

Encoding of melodies in congenital amusia

Source reconstruction of the N100m allowed us to disentangle activity in bilateral auditory regions (in Heschl's gyrus/superior temporal gyrus/planum temporale/planum polare) and in the pars opercularis of the inferior frontal gyrus (BA 44). This observation is in line with numerous studies showing the major role of these areas in music perception and memory (Zatorre *et al.*, 1994; Griffiths *et al.*, 1999; Schulze *et al.*, 2009; see also Griffiths, 2001; Peretz and Zatorre, 2005 for reviews). For the N100m component, the observed frontal generators are in line with previous EEG data (Näätänen and Picton, 1987; Alcaini *et al.*, 1994; Giard *et al.*, 1994; Pantev *et al.*, 1995; see Trainor and Unrau, 2012 for review), suggesting that separate frontal and temporal neural systems mediating different processes could be activated during the N1-time range. Further evidence comes from intracranial recordings (Edwards *et al.*, 2005), which indicated that during the N1-time range, the most strongly activated brain regions were not only the superior surface of the temporal lobe, but also some areas of the frontal lobe.

Note that in comparison with tone 1, tones 2 to 6 recruited a more lateral part of Heschl's gyrus (and planum temporale). This finding is in agreement with Patterson *et al.* (2002) showing that the lateral part of the Heschl's gyrus is recruited by melodies in comparison to fixed-pitch sequences, suggesting that this part of the auditory cortex plays a role in melodic processing. It is well known that bilateral auditory regions (including superior temporal gyrus, Heschl's gyrus, planum polare and planum temporale) play a critical role in the encoding of acoustic features of individual tones (Griffiths, 1999, 2001) and in melody perception (Zatorre *et al.*, 1994; Liégeois-Chauvel *et al.*, 1998) and are more strongly recruited during high-load conditions, which require active rehearsal (Zatorre *et al.*, 1994; Griffiths *et al.*, 1999), such as retaining a pitch in memory while subsequent tones are presented. In addition to the participation of auditory cortices, Griffiths (1999, 2001; Griffiths *et al.*, 2000) has suggested that higher-order auditory patterns and information are analysed by distributed networks including temporal and frontal lobes, which are necessary for online maintenance and encoding of tonal patterns. For frontal regions, the inferior frontal gyrus has been further suggested to play a role in integrating sequential auditory events and in the encoding of tonality (Zatorre *et al.*, 1994; Griffiths *et al.*, 1999; Gaab *et al.*, 2003; Peretz and Zatorre, 2005; Schulze *et al.*, 2009, 2011).

For the first tone, source analyses revealed near-normal cortical distribution and activity, both for the P50m (e.g. Yvert *et al.*, 2001) and the N100m components in the amusic brain. As also reported by Peretz *et al.* (2005, 2009) and Hyde *et al.* (2011), the functioning of amusics' auditory cortices was unimpaired for a relatively simple acoustic task (the first tone of the sequence was presented after a period of silence, and should thus largely recruit non-specific components). However, latency analyses of the N100m revealed slightly delayed inferior frontal gyrus responses in amusics in comparison to control subjects, thus providing new functional evidence of an altered fronto-temporal network, even for the encoding of the first tone of a sequence.

For tones 2 to 6, the N100m differed between amusics and control subjects in amplitude and latency in Heschl's gyrus/superior temporal gyrus and inferior frontal gyrus, thus placing the deficit of the amusic brain early in the auditory processing stream. Amplitude data revealed that amusic participants recruited less strongly a bilateral fronto-temporal network than did control subjects. This network has been previously shown to allow auditory information to be processed, encoded, maintained on-line and related to previous elements of a sequence (Zatorre *et al.*, 1994, 2002; Griffiths, 2001; Janata *et al.*, 2002*a, b*; Gaab *et al.*, 2003; Peretz and Zatorre, 2005). Observing an abnormal recruitment of frontal (inferior frontal gyrus) and temporal (Heschl's gyrus/superior temporal gyrus) regions during the encoding of tones 2 to 6 in amusics can be taken as functional correlates of the deficit in perception and memory of tone sequences in congenital amusia. It might be argued that decreased attention could explain the decreased N100m amplitude of amusics in the contour task. However, two aspects of the data allow us to reject this argument: (i) if amusics were paying less attention because, for example, they found the contour task too difficult, the amplitude of their N100m should have been attenuated in comparison with the transposition task. However, no interaction between Group and Task was observed for the N100m amplitude, thus suggesting that participants' attentional or task-related strategies did not have a major influence on the amplitude differences in the present N100m data; (ii) the sustained evoked responses data revealed that task-related (the difficult contour task versus the easier transposition task) attentional modulation is preserved in congenital amusia. Indeed, this task effect on sustained evoked responses is in line with the hypothesis that the contour task should involve more attentional processes than the transposition task, as these responses are known to be modulated by attention (Picton *et al.*, 1978; Bidet-Caulet *et al.*, 2007). As there were no significant main effect of group and no significant interactions involving the group factor, the data pattern suggests that amusic participants maintained a sustained attention level, even for the difficult contour task.

For tones 2 to 6, increased latency of the N100m was observed in frontal and temporal regions for the amusic group (relative to the control group). It can be interpreted as reflecting delayed encoding of S1 in the amusic brain as well as impaired high-level processing and stimulus representations (e.g. delayed responses associated with difficulties to maintain tone sequences in memory). The most interesting point of these data is the increased Heschl's gyrus latency for tones 2 to 6 in the amusic group. Together with the amplitude data, this finding provides the first evidence of functional abnormalities in the auditory cortex in congenital amusia.

To further assess the differences between the two participant groups during the encoding of the melodies, we used dynamic causal modelling to investigate the effective connectivity between the frontal and temporal sources of the N100m. Most approaches of connectivity in the MEG/EEG literature use functional measures, such as phase-synchronization, temporal correlations or coherence, to establish statistical dependencies between activities in two different cortical areas. Although functional connectivity can be used to establish statistical dependency, it does not provide information about the causal architecture of the interactions. In dynamic causal modelling, this influence is parameterized in a causal model, which can then be estimated using Bayesian inference. Using the same underlying cortical architecture, but differing between participant groups in the modulations of specific connection types, we observed decreased intrinsic connectivity for amusics in both auditory cortices relative to control subjects, as well as increased connectivity in amusics between the two auditory cortices, as previously observed with functional MRI (Hyde et al., 2011). These abnormalities observed in both intrinsic and lateral connections in the auditory cortex provide additional evidence of functional anomaly of the auditory cortex in amusia (together with the observed delayed N100m responses). More precisely, the observed hyper-connectivity between the two auditory cortices in the amusic brain might be a marker of the primary deficit, as also observed in other developmental disorders (see Wolf et al., 2010 for dyslexia; Hyde et al., 2011 for converging evidence in amusia) or rather reveal compensatory mechanisms of the amusic brain. The latter would suggest that amusics might compensate for an impoverished processing in the right auditory cortex by recruiting the contralateral auditory cortex. These functional abnormalities in both auditory cortices were associated to decreased right frontal-to-temporal connectivity in amusics. This decreased right fronto-temporal connectivity is in agreement with the functional MRI data from Hyde et al. (2011), showing decreased connectivity between right inferior frontal gyrus and right superior temporal gyrus in amusics in comparison with control subjects during passive listening. Note that our present data, which are based on the effective connectivity approach (as opposed to functional connectivity used in Hyde et al., 2011), allow us to more precisely ascribe abnormal brain responses in amusics to reduced backward connections, but not forward connections.

Maintenance of melodic information in memory

To investigate neural correlates of maintenance of tone information in memory, we have analysed oscillatory synchronization in the gamma frequency during the retention period. We observed enhanced gamma-power in the right dorsolateral prefrontal cortex in control subjects in comparison with amusics for the difficult contour task. This finding is in agreement with previous data showing that high-level representations of task-relevant information are reflected in the gamma-frequency (Jensen et al., 2007), and that increased gamma-frequency power in fronto-temporal areas is related with short-term memory processing (Kaiser et al., 2003). This hypothesis is also supported by the task

comparison for control subjects, showing that the contour task recruited more strongly a bilateral frontal network involving the inferior frontal gyrus. Moreover, this right-lateralized group effect is in agreement with neuroimaging data showing that the dorsolateral prefrontal cortex participates in short-term memory and working memory (Glahn et al., 2002; Jerde et al., 2011). Note, however, that these modulations with tasks were not observed in amusics, thus providing another evidence of altered processing of melodic information in the amusic brain. In addition, for the easier transposition task, increased gamma synchronization was observed in the left inferior frontal gyrus and in the right temporo-parietal junction in amusics in comparison with control subjects, suggesting that the relative involvement of each hemisphere might differ between the two groups. The interaction between frontal and temporo-parietal regions is known to be implicated in working memory for musical material (Zatorre et al., 1994; Jerde et al., 2011), and the present finding thus suggests that, in order to perform as well as control subjects for the easier transposition task, amusics had to recruit more strongly this network. Following this hypothesis, we suggest that for the difficult contour task, this compensatory strategy was not sufficient to overcome the deficit in pitch memory.

Retrieval of melodic information

Source reconstruction of the event-related fields elicited by the changed tone (difference wave between different trials and same trials) allowed us to observe activity in bilateral superior temporal gyrus and in the pars opercularis of the inferior frontal gyrus (BA 44) in the control group. Observing that the same network is recruited during encoding and retrieval of melodic information is in line with research showing the role of these areas in music perception and memory (Zatorre et al., 1994, 2002; Griffiths, 2001; Janata et al., 2002a, b; Gaab et al., 2003; Peretz and Zatorre, 2005). The present data revealed an abnormal recruitment of frontal (inferior frontal gyrus) and temporal (superior temporal gyrus) regions during the retrieval of melodic information in the amusic brain; data that further reflect the functional correlates of the deficit in short-term memory of melodic information in congenital amusia.

Convergence between functional and anatomical data

MEG source locations revealing group differences were congruent with bilateral fronto-temporal networks well-known to be involved in short-term memory for pitch sequences (Zatorre et al., 1994; Griffiths et al., 1999; Schulze et al., 2009), and attentive listening to musical sequences (Janata et al., 2002b) (Table 3 and Fig. 12). In the right superior temporal gyrus, anatomical group differences were spatially congruent with the present MEG data, and activations peaks previously reported in positron emission tomography (PET)/functional MRI studies investigating the active processing of pitch sequences (i.e. short-term memory, attentive listening). However, for frontal regions, anatomical and functional group differences were observed in different parts of the inferior frontal gyrus. VBM analyses (present data and data of Hyde et al., 2006) revealed abnormal grey and white matter concentrations

Table 3 MNI coordinates (mm) of fronto-temporal activations reported in previous functional MRI/PET and DTI studies

Reference	Lobe	Region	Contrasts	Hemisphere	x	y	z
Zatorre *et al.*, 1994	Frontal	Insula/inferior frontal gyrus pars opercularis	Tone judgement–passive melodies	Right	38	20	5
				Left	−31	22	8
	Temporal	Superior temporal gyrus	Passive melodies–scanner noise	Right	62	−25	3
Griffths *et al.*, 1999	Frontal	Inferior frontal gyrus, pars opercularis	Pitch memory task–resting	Right	30	26	2
				Right	40	18	8
				Left	−48	10	4
	Temporal	Heschl's gyrus		Right	66	−30	6
Janata *et al.*, 2002a	Frontal	Inferior frontal gyrus, pars opercularis	Listen–resting state	Right	45	15	5
				Right	−38	19	5
	Temporal	Superior temporal gyrus	Listen–resting state	Right	60	19	5
			Attend–resting state	Left	−60	−15	5
				Left	−52	−19	5
Hyde *et al.*, 2011	Frontal	Inferior frontal gyrus, pars opercularis	Amusics-control subjects, pitch changed – fixed pitch	Right	34	32	2
	Temporal	Heschl's gyrus/planum temporale	Pitch co-variation in control subjects	Right	56	−12	4
		Planum temporale	Pitch co-variation in amusics	Right	60	−18	8
		Heschl's gyrus/planum temporale		Left	−44	−30	8
Loui *et al.*, 2009	Frontal	Inferior frontal gyrus, pars opercularis	Seed regions for tractography	Right and left			
	Temporal	Posterior superior temporal gyrus		Right and left			

The first three studies were run with typical individuals, whereas the last two studies compared amusics and control groups.

in the amusic brain in the anterior part of the right inferior frontal gyrus (BA 45/47; Fig. 12), whereas the MEG data revealed functional abnormalities in more posterior parts of bilateral inferior frontal gyrus (BA 44) during encoding and retrieval of melodic information (S1 and S2), and in a more dorsal part (BA9/46) during the retention. The anterior part of the right inferior frontal gyrus has also been shown to exhibit abnormal blood oxygen level-dependant deactivation in the amusic brain during passive listening to tone sequences (Hyde *et al.*, 2011). The posterior part of the inferior frontal gyrus was used by Loui *et al.* (2009) as a seed region for DTI tractography revealing an anomalous connectivity in the right arcuate fasciculus in congenital amusia (Fig. 12 and Table 3).

For music processing, these different subregions of right inferior frontal gyrus have been previously reported in typical individuals' brains (e.g. Maess *et al.*, 2001; Koelsch *et al.*, 2002 for BA 44; Levitin and Menon, 2003 for BA 45/47; Tillmann *et al.*, 2003, 2006). Hyde *et al.* (2006, 2011) discussed their VBM and functional MRI data of amusic participants in link with two subregions involved in music perception and structure processing. They suggested that BA 44 and BA 45/47 are part of the same network of right inferior frontal cortex and are involved in pitch sequence perception and integration. Similarly, Hagoort (2005) discusses the role of the left inferior frontal gyrus in language processing by regrouping BA 44, 45 and 47, notably their role in enabling the

integration of structural information. While for language processing, subregions have also been specified for their respective roles in syntactic processing and phonological memory (BA 44) and semantic processing (BA 45/47) (Friederici, 2002, 2012), research in music processing still needs to investigate more specifically the respective roles of these subregions in right inferior frontal gyrus for pitch perception and memory in both normal-functioning and amusic brains.

Conclusion

In congenital amusics, N100m components were abnormal and strongly delayed in bilateral inferior frontal gyrus and Heschl's gyrus/superior temporal gyrus during the encoding of melodies, frontal gamma synchronization was decreased during the retention of melodic information, and fronto-temporal responses were altered during the retrieval of this information. These functional anomalies were related to abnormal grey matter and white matter concentrations in the same brain regions and to a deficit in memory processing observed at the behavioural level. This data set is in agreement with current hypotheses about the role of frontal and temporal structures (including auditory cortices) and of the fronto-temporal pathway in music processing, as well as its impairment in this 'musical handicap'. The present study
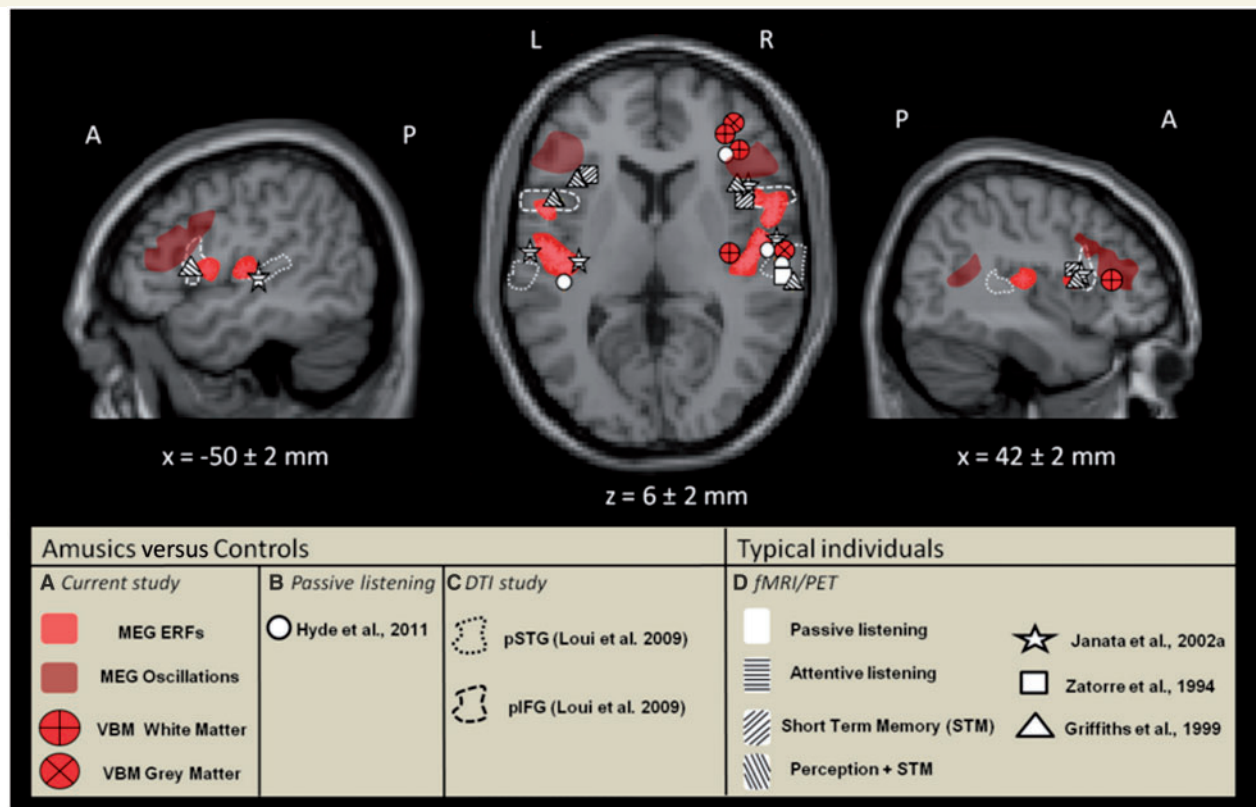
Figure 12 Comparison of MEG and VBM group differences observed in the present study (**A** in red and purple) with previous functional MRI (**B**) and DTI (**C**) studies comparing control subjects and amusics and with activation peaks observed in previous studies using PET or functional MRI in typical individuals (**D**) (Table 3). (**B**) Coordinates of the functional group difference in the right inferior frontal gyrus and of pitch co-variation data for both amusics and control subjects in the auditory cortex (Hyde *et al.*, 2011, circles; Table 3). (**C**) Approximate location of the posterior inferior frontal gyrus (pIFG) and posterior superior temporal gyrus (pSTG) seed regions used in the tractography study by Loui *et al.* (2009) (bold dotted lines and plain dotted line, respectively). Note that these regions were manually selected by a neurologist who was blind concerning the protocol and previously reported activations. (**D**) Coordinates of activations obtained with pitch material during auditory short-term memory tasks (Zatorre *et al.*, 1994, squares; Griffths *et al.*, 1999, triangles), attentive listening (Janata *et al.*, 2002, stars), and passive listening (Zatorre *et al.*, 1994, squares). Right-oriented bars = short-term memory; left-oriented bars = pitch perception and memory; horizontal bars = attentive listening; white background = passive listening. Activations are displayed on the single subject T$_1$ image provided by SPM8. The figure depicts activation with coordinates that were 2 mm up or down the central slice. R = right; L = left; A = anterior; P = posterior); PET = positron emission tomography.

improves our understanding not only of congenital amusia itself, but also, more generally, of typical brain functioning related to auditory perception and memory, as well as music processing. In particular, control participants' data reveal that the same bilateral fronto-temporal areas were recruited during the different stages of encoding, maintenance and retrieval of auditory information, suggesting the same networks are involved in these different aspects of melody processing and memory representation. Research investigating the neural correlates of music processing is an expanding research domain that also addresses the question of shared neural correlates with language processing (Patel, 2003, 2008). This comparison can also be extended to the understanding of developmental language impairments in parallel to the present musical impairments (as previously suggested by Hyde *et al.*, 2006). Beyond providing new data for the understanding of the deficits in the amusic brain, our data raises issues about the cerebral correlates of developmental disorders in general, notably by

suggesting that cerebral correlates of related deficits can be observed already in early brain responses. This later point represents a major contribution to the comprehension of neural correlates underlying congenital amusia as previous data have suggested that the functional neural anomaly mainly lies outside the auditory cortex (Peretz *et al.*, 2005, 2009; Moreau *et al.*, 2009; Hyde *et al.*, 2011).

Acknowledgements

Funding

References

Aguera PE, Jerbi K, Caclin A, Bertrand O. ELAN: a software package for analysis and visualization of MEG, EEG, and LFP signals. Comput Intell Neurosci 2011; 2011: 158970.

Alcaini M, Giard MH, Thevenet M, Pernier J. Two separate frontal components in the N1 wave of the human auditory evoked response. Psychophysiology 1994; 31: 611–15.

Ashburner J, Friston KJ. Unified segmentation. Neuroimage 2005; 26: 839–51.

Ayotte J, Peretz I, Hyde KL. Congenital amusia: a group study of adults afflicted with a music-specific disorder. Brain 2002; 125: 238–51.

Bidet-Caulet A, Fischer C, Besle J, Aguera P, Giard MH, Bertrand O. Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. J Neurosci 2007; 27: 9252–61.

Bouet R, Jung J, Delpuech C, Ryvlin P, Isnard J, Guenot M, et al. Towards source volume estimation of interictal spikes in focal epilepsy using magnetoencephalography. Neuroimage 2012; 59: 3955–66.

Caclin A, McAdams S, Smith BK, Giard MH. Interactive processing of timbre dimensions: an exploration with event-related potentials. J Cogn Neurosci 2008; 20: 49–64.

David O, Friston KJ. A neural mass model for MEG/EEG: coupling and neuronal dynamics. Neuroimage 2003; 20: 1743–55.

David O, Kiebel SJ, Harrison LM, Mattout J, Kilner JM, Friston KJ. Dynamic causal modeling of evoked responses in EEG and MEG. Neuroimage 2006; 30: 1255–72.

Demarquay G, Caclin A, Brudon F, Fischer C, Morlet D. Exacerbated attention orienting to auditory stimulation in migraine patients. Clin Neurophysiol 2011; 122: 1755–63.

Edwards E, Soltani M, Deouell LY, Berger MS, Knight RT. High gamma activity in response to deviant auditory stimuli recorded directly from human cortex. J Neurophysiol 2005; 94: 4269–80.

Foxton J, Dean J, Gee R, Peretz I, Griffiths T. Characterization of deficits in pitch perception underlying 'tone deafness'. Brain 2004; 127: 801–10.

Friederici AD. Towards a neural basis of auditory sentence processing. Trends Cogn Sci 2002; 6: 78–84.

Friederici AD. The cortical language circuit: from auditory perception to sentence comprehension. Trends Cogn Sci 2012; 16: 262–8.

Friston K, Harrison L, Daunizeau J, Kiebel S, Phillips C, Trujillo-Barreto N, et al. Multiple sparse priors for the M/EEG inverse problem. Neuroimage 2008; 39: 1104–20.

Friston KJ, Penny W. Posterior probability maps and SPMs. Neuroimage 2003; 19: 1240–9.

Gaab N, Gaser C, Zaehle T, Jancke L, Schlaug G. Functional anatomy of pitch memory–an fMRI study with sparse temporal sampling. Neuroimage 2003; 19: 1417–26.

Garrido MI, Kilner JM, Kiebel SJ, Stephan KE, Friston KJ. Dynamic causal modelling of evoked potentials: a reproducibility study. Neuroimage 2007; 36: 571–80.

Giard MH, Perrin F, Echallier JF, Thévenet M, Froment JC, Pernier J. Dissociation of temporal and frontal components in the human auditory N1 wave: a scalp current density and dipole model analysis. Electroencephalogr Clin Neurophysiol 1994; 92: 238–52.

Glahn DC, Kim J, Cohen MS, Poutanen VP, Therman S, Bava S, et al. Maintenance and manipulation in spatial working memory: dissociations in the prefrontal cortex. Neuroimage 2002; 17: 201–13.

Gosselin N, Jolicoeur P, Peretz I. Impaired memory for pitch in congenital amusia. Ann NY Acad Sci 2009; 1169: 270–2.

Griffiths TD. Human complex sound analysis. Clin Sci (Lond) 1999; 96: 231–4.

Griffiths TD. The neural processing of complex sounds. Ann N Y Acad Sci 2001; 930: 133–42.

Griffiths TD, Johnsrude I, Dean JL, Green GG. A common neural substrate for the analysis of pitch and duration pattern in segmented sound? Neuroreport 1999; 10: 3825–30.

Griffiths TD, Penhune V, Peretz I, Dean JL, Patterson RD, Green GG. Frontal processing and auditory perception. Neuroreport 2000; 11: 919–22.

Gross J, Kujala J, Hamalainen M, Timmermann L, Schnitzler A, Salmelin R. Dynamic imaging of coherent sources: studying neural interactions in the human brain. Proc Natl Acad Sci USA 2001; 98: 694–9.

Guthrie D, Buchwald JS. Significance testing of difference potentials. Psychophysiology 1991; 28: 240–4.

Hagoort P. On Broca, brain, and binding: a new framework. Trends Cogn Sci 2005; 9: 416–23.

Hyde KL, Lerch JP, Zatorre RJ, Griffiths TD, Evans AC, Peretz I. Cortical thickness in congenital amusia: when less is better than more. J Neurosci 2007; 27: 13028–32.

Hyde KL, Peretz I. Brains that are out of tune but in time. Psychol Sci 2004; 15: 356–60.

Hyde KL, Zatorre R, Griffiths T, Lerch J, Peretz I. Morphometry of the amusic brain: a two-site study. Brain 2006; 129: 2562–70.

Hyde KL, Zatorre RJ, Peretz I. Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. Cereb Cortex 2011; 21: 292–9.

Janata P, Birk JL, Van Horn JD, Leman M, Tillmann B, Bharucha JJ. The cortical topography of tonal structures underlying Western music. Science 2002a; 298: 2167–70.

Janata P, Tillmann B, Bharucha JJ. Listening to polyphonic music recruits domain-general attention and working memory circuits. Cogn Affect Behav Neurosci 2002b; 2: 121–40.

Jensen O, Kaiser J, Lachaux JP. Human gamma-frequency oscillations associated with attention and memory. Trends Neurosci 2007; 30: 317–24.

Jerde TA, Childs SK, Handy ST, Nagode JC, Pardo JV. Dissociable systems of working memory for rhythm and melody. Neuroimage 2011; 57: 1572–9.

Jiang J, Zhu W, Shi F, Liu Y, Li J, Qin W, et al. Thick visual cortex in the early blind. J Neurosci 2009; 29: 2205–11.

Johnsrude I, Penhune VB, Zatorre RJ. Functional specificity in the right human auditory cortex for perceiving pitch direction. Brain 2000; 123: 155–63.

Kaiser J, Lutzenberger W. Human gamma-band activity: a window to cognitive processing. Neuroreport 2005; 16: 207–11.

Kaiser J, Ripper B, Birbaumer N, Lutzenberger W. Dynamics of gamma-band activity in human magnetoencephalogram during auditory pattern working memory. Neuroimage 2003; 20: 816–27.

Kiebel SJ, Garrido MI, Moran R, Chen CC, Friston KJ. Dynamic causal modeling for EEG and MEG. Hum Brain Mapp 2009; 30: 1866–76.

Koelsch S, Gunter TC, v Cramon DY, Zysset S, Lohmann G, Friederici AD. Bach speaks: a cortical "language-network" serves the processing of music. Neuroimage 2002; 17: 956–66.

Levitin DJ, Menon V. Musical structure is processed in "language" areas of the brain: a possible role for Brodmann Area 47 in temporal coherence. Neuroimage 2003; 20: 2142–52.

Liegeois-Chauvel C, Peretz I, Babai M, Laguitton V, Chauvel P. Contribution of different cortical areas in the temporal lobes to music processing. Brain 1998; 121: 1853–67.

Litvak V, Mattout J, Kiebel S, Phillips C, Henson R, Kilner J, et al. EEG and MEG data analysis in SPM8. Comput Intell Neurosci 2011; 2011: 852961.

Loui P, Alsop D, Schlaug G. Tone deafness: a new disconnection syndrome? J Neurosci 2009; 29: 10215–20.

Maess B, Koelsch S, Gunter TC, Friederici AD. Musical syntax is processed in Broca's area: an MEG study. Nat Neurosci 2001; 4: 540–5.

Mainy N, Kahane P, Minotti L, Hoffmann D, Bertrand O, Lachaux JP. Neural correlates of consolidation in working memory. Hum Brain Mapp 2007; 28: 183–93.

Mandell J, Schulze K, Schlaug G. Congenital amusia: an auditory-motor feedback disorder? Restor Neurol Neurosci 2007; 25: 323–34.

Mattout J, Henson RN, Friston KJ. Canonical source reconstruction for MEG. Comput Intell Neurosci 2007; 2007: 67613.

Mattout J, Phillips C, Penny WD, Rugg MD, Friston KJ. MEG source localization under multiple constraints: an extended Bayesian framework. Neuroimage 2006; 30: 753–67.

Moreau P, Jolicoeur P, Peretz I. Automatic brain responses to pitch changes in congenital amusia. Ann NY Acad Sci 2009; 1169: 191–4.

Näätänen R, Picton T. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiol 1987; 24: 375–425.

Pantev C, Bertrand O, Eulitz C, Verkindt C, Hampson S, Schuierer G, et al. Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. Electroencephalogr Clin Neurophysiol 1995; 94: 26–40.

Park HJ, Lee JD, Kim EY, Park B, Oh MK, Lee S, et al. Morphological alterations in the congenital blind based on the analysis of cortical thickness and surface area. Neuroimage 2009; 47: 98–106.

Patel AD. Language, music, syntax and the brain. Nat Neurosci 2003; 6: 674–81.

Patel AD. Music, language and the brain. Oxford: Oxford University Press; 2008.

Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD. The processing of temporal pitch and melody information in auditory cortex. Neuron 2002; 36: 767–76.

Penny WD, Stephan KE, Daunizeau J, Rosa MJ, Friston KJ, Schofield TM, et al. Comparing families of dynamic causal models. PLoS Comput Biol 2010; 6: e1000709.

Peretz I. Processing of local and global musical information by unilateral brain-damaged patients. Brain 1990; 113: 1185–205.

Peretz I, Ayotte J, Zatorre R, Mehler J, Ahad P, Penhune V, et al. Congenital amusia: a disorder of fine-grained pitch discrimination. Neuron 2002; 33: 185–91.

Peretz I, Brattico E, Jarvenpaa M, Tervaniemi M. The amusic brain: in tune, out of key, and unaware. Brain 2009; 132: 1277–86.

Peretz I, Brattico E, Tervaniemi M. Abnormal electrical brain responses to pitch in congenital amusia. Ann Neurol 2005; 58: 478–82.

Peretz I, Champod S, Hyde K. Varieties of musical disorders: the Montreal Battery of Evaluation of Amusia. Ann NY Acad Sci 2003; 999: 58–75.

Peretz I, Zatorre RJ. Brain organization for music processing. Annu Rev Psychol 2005; 56: 89–114.

Picton TW, Woods DL, Proulx GB. Human auditory sustained potentials: part I. The nature of the response. Electroencephalogr Clin Neurophysiol 1978; 45: 186–97.

Ramirez RR, Kopell BH, Butson CR, Hiner BC, Baillet S. Spectral signal space projection algorithm for frequency domain MEG and EEG denoising, whitening, and source imaging. Neuroimage 2011; 56: 78–92.

Schulze K, Gaab N, Schlaug G. Perceiving pitch absolutely: comparing absolute and relative pitch possessors in a pitch memory task. BMC Neurosci 2009; 10: 106.

Schulze K, Mueller K, Koelsch S. Neural correlates of strategy use during auditory working memory in musicians and non-musicians. Eur J Neurosci 2011; 33: 189–96.

Stewart L. Congenital amusia. Curr Biol 2006; 16: R904–6.

Stewart L. Fractionating the musical mind: insights from congenital amusia. Curr Opin Neurobiol 2008; 18: 127–30.

Tallon-Baudry C, Bertrand O. Oscillatory gamma activity in humans and its role in object representation. Trends Cogn Sci 1999; 3: 151–62.

Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J. Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. J Neurosci 1996; 16: 4240–9.

Tallon-Baudry C, Kreiter A, Bertrand O. Sustained and transient oscillatory responses in the gamma and beta bands in a visual short-term memory task in humans. Vis Neurosci 1999; 16: 449–59.

Tesche CD, Uusitalo MA, Ilmoniemi RJ, Huotilainen M, Kajola M, Salonen O. Signal-space projections of MEG data characterize both distributed and well-localized neuronal sources. Electroencephalogr Clin Neurophysiol 1995; 95: 189–200.

Tillmann B, Schulze K, Foxton JM. Congenital amusia: a short-term memory deficit for non-verbal, but not verbal sounds. Brain Cogn 2009; 71: 259–64.

Tillmann B, Janata P, Bharucha JJ. Activation of the inferior frontal cortex in musical priming. Ann N Y Acad Sci 2003; 999: 209–11.

Tillmann B, Koelsch S, Escoffier N, Bigand E, Lalitte P, Friederici AD, von Cramon DY. Cognitive priming in sung and instrumental music: activation of inferior frontal cortex. Neuroimage 2006; 31: 1771–82.

Trainor LJ, Unrau AJ. Development of pitch and music perception. In: Werner L, Fay RR, Popper AN, editors. Springer handbook of auditory research: human auditory development. New York: Springer; 2012. p. 223–54.

Williamson VJ, McDonald C, Deutsch D, Griffiths TD, Stewart L. Faster decline of pitch memory over time in congenital amusia. Adv Cogn Psychol 2010; 6: 15–22.

Williamson VJ, Stewart L. Memory for pitch in congenital amusia: beyond a fine-grained pitch discrimination problem. Memory 2010; 18: 657–69.

Wolf RC, Sambataro F, Lohr C, Steinbrink C, Martin C, Vasic N. Functional brain network abnormalities during verbal working memory performance in adolescents and young adults with dyslexia. Neuropsychologia 2010; 48: 309–18.

Yvert B, Crouzeix A, Bertrand O, Seither-Preisler A, Pantev C. Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. Cereb Cortex 2001; 11: 411–23.

Zatorre RJ, Belin P, Penhune VB. Structure and function of auditory cortex: music and speech. Trends Cogn Sci 2002; 6: 37–46.

Zatorre RJ, Evans AC, Meyer E. Neural mechanisms underlying melodic perception and memory for pitch. J Neurosci 1994; 14: 1908–19.