



Rapid encoding of musical tones discovered in whole-brain connectivity

L. Bonetti^{a,b,c,d,*}, E. Brattico^{b,h}, F. Carlomagno^b, G. Donati^{d,b}, J. Cabral^{a,b,g}, N.T. Haumann^b, G. Deco^{e,f}, P. Vuust^b, M.L. Kringelbach^{a,b,c}

^a Centre for Eudaimonia and Human Flourishing, University of Oxford, United Kingdom

^b Center for Music in the Brain, Department of Clinical Medicine, Aarhus University & The Royal Academy of Music Aarhus/Aalborg, Denmark

^c Department of Psychiatry, University of Oxford, Oxford, United Kingdom

^d Department of Psychology, University of Bologna, Italy

^e Institució Catalana de la Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, Barcelona, 08010, Spain

^f Computational and Theoretical Neuroscience Group, Center for Brain and Cognition, Universitat Pompeu Fabra, Barcelona, Spain

^g Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, 4710-057 Braga, Portugal

^h Department of Education, Psychology, Communication, University of Bari Aldo Moro, Italy

ARTICLE INFO

Keywords:

Sound encoding
Brain dynamics
Memory
Magnetoencephalography (MEG)
Whole-brain functional connectivity

ABSTRACT

Information encoding has received a wide neuroscientific attention, but the underlying rapid spatiotemporal brain dynamics remain largely unknown. Here, we investigated the rapid brain mechanisms for encoding of sounds forming a complex temporal sequence. Specifically, we used magnetoencephalography (MEG) to record the brain activity of 68 participants while they listened to a highly structured musical prelude. Functional connectivity analyses performed using phase synchronisation and graph theoretical measures showed a large network of brain areas recruited during encoding of sounds, comprising primary and secondary auditory cortices, frontal operculum, insula, hippocampus and basal ganglia. Moreover, our results highlighted the rapid transition of brain activity from primary auditory cortex to higher order association areas including insula and superior temporal pole within a whole-brain network, occurring during the first 220 ms of the encoding process. Further, we discovered that individual differences along cognitive abilities and musicianship modulated the degree centrality of the brain areas implicated in the encoding process. Indeed, participants with higher musical expertise presented a stronger centrality of superior temporal gyrus and insula, while individuals with high working memory abilities showed a stronger centrality of frontal operculum. In conclusion, our study revealed the rapid unfolding of brain network dynamics responsible for the encoding of sounds and their relationship with individual differences, showing a complex picture which extends beyond the well-known involvement of auditory areas. Indeed, our results expanded our understanding of the general mechanisms underlying auditory pattern encoding in the human brain.

1. Introduction

Memory is doubtless one of the most crucial cognitive abilities of humans and animals, necessary to allow species to survive and act (Klein et al., 2010). Among its building elements, information encoding plays a fundamental role, allowing individuals to accomplish the crucial goal of learning from experience. In the last decades of neuroscience, much has been done on information encoding, working both with humans and animal models (Fazio et al., 2009; Herry and Johansen, 2014; Takeuchi et al., 2014). A large part of this research focused on encoding of visual and spatial stimuli (Hickey and Peelen, 2015; Stern et al., 1996), whereas a conspicuous number of studies explored auditory processes related to early and even pre-attentive elaboration of standard

and deviant sounds inserted in elementary auditory and more complex musical sequences (Conley et al., 1999; Lijffijt et al., 2009; Näätänen et al., 2007; Vuust et al., 2012).

The former approach led to several examples of visual encoding and recognition studies which employed faces as main stimuli, showing for instance the key role of fusiform gyrus (Dobs et al., 2019; Kanwisher et al., 1997; Schwarzlose et al., 2005).

Conversely, regarding auditory research a large number of studies deeply explored the early and pre-attentive processing of sounds and simple temporal sequences. For instance, this research highlighted several automatic event related potentials/fields (ERP/F) to standard and deviant sounds such as the well-known N100 and mismatch negativity (MMN). Indeed, it has been shown that N100 and MMN were modulated

* Corresponding author.

E-mail address: leonardo.bonetti@psych.ox.ac.uk (L. Bonetti).

<https://doi.org/10.1016/j.neuroimage.2021.118735>.

Received 23 June 2021; Received in revised form 30 September 2021; Accepted 14 November 2021

Available online 20 November 2021.

1053-8119/© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

by the stimulus characteristics, implying that an initial memory processing was happening already in the primary auditory cortex (Conley et al., 1999; Lijffijt et al., 2009; Näätänen et al., 2007; Vuust et al., 2012).

Other studies have used proper musical paradigms to better understand the brain mechanisms underlying processing of sounds and temporal sequences. Such choice was driven by the fact that music is the human art that mainly acquires meaning through the logical combination of objects (sounds) extended over time (Koelsch et al., 2004), and thus provides great opportunity to refine our understanding of how the brain encodes and recognises temporal information. Along this line, classic studies from Koelsch and colleagues revealed signs of musical syntax processing in the brain. Especially, they discovered the early right-anterior negativity (ERAN), a component of the ERP/F elicited by harmonically non-appropriate chords inserted in a sequence of coherent ones (Koelsch et al., 2004; Maess et al., 2001; Koelsch and Siebel, 2005). Similarly, additional studies using event-related potentials revealed that the brain selectively responded to violation of harmonic musical structure (Leino et al., 2007; Villarreal et al., 2011). On another note, it has been shown that music, as well as words, was able to prime the meaning of subsequent words, suggesting parallels between the brain processing of language and music (Koelsch et al., 2004).

Other studies showed the role of the neural entrainment to sounds and rhythms, proposing it as a key feature of both music processing and control mechanisms of neural sensory gain (Obleser and Kayser, 2019; Novembre and Iannetti, 2018). Such studies demonstrated that some neural processes presented a temporal alignment with exogenous stimuli arranged in regular sequences (e.g. regular series of beats or sounds). Furthermore, it has been shown that such mechanisms can present functional neuroplasticity being modulated by individual differences in musical expertise (Celma-Mirallas and Toro, 2019) and aging (Henry et al., 2017).

Although this research clarified several aspects of how the brain processes music and auditory information, it did not directly explain how the sounds were actually encoded. A more general approach on sound processing investigation reported the activation of auditory cortex brain regions such as Heschl's and superior temporal gyri in response to acoustic stimuli varying in temporal and spectral features (Richiardi et al., 2011; Warrier et al., 2009) as well as a contribution coming from higher-order brain areas (Husain et al., 2006; Langers and Melcher, 2011). Further studies focused on cognitively demanding tasks developed with musical materials, aiming to further understand how the brain processes music and how music can help us to better understand the brain. Along this line, several works focused on cognitive processes involved in music listening such as memory, attention, and evaluation. For instance, in a pioneer PET study by Zatorre et al. (1994) on melodic perception and memory, authors discovered that perceptual analysis of melodies recruited parts of the right superior temporal cortex, while pitch comparison of specific tones within the melodies involved mainly right prefrontal cortex. Similarly, Gaab et al. (2003) asked participants to compare different simple melodic sequences and detected activation of supramarginal and left inferior frontal gyri, superior temporal, superior parietal, posterior dorsolateral frontal and dorsolateral cerebellar regions. More recently, Kumar et al. (2016) showed that to hold in mind a single sound was crucial the activity and connectivity between three main brain areas: primary auditory cortex, hippocampus and inferior frontal gyrus. An additional fMRI study (Sikka et al., 2015) about recognition of familiar music proposed that successful task performance was associated to the activation of right superior temporal, superior frontal, bilateral inferior, bilateral precentral and left supramarginal gyri. Back to oscillations and neurophysiology, a recent work by Albouy et al. (2017) showed the brain activity underlying memory retention, highlighting that theta oscillations in the brain dorsal stream predicted participants' abilities to perform an auditory memory task where they held and manipulated sound information.

Even though these studies provided a detailed description of the brain areas mainly involved in a wide range of music processing tasks,

they did not reveal the fast-scale whole-brain connectivity patterns (either static or dynamic) underlying encoding of the single sounds forming a structured, musical prelude. Moreover, as explicitly stated in their review on processing and encoding of temporal sequences, Dehaene et al. (Dehaene et al., 2015) highlighted the urgency to unravel which is the specific contribution of cortical and subcortical brain networks to encoding of the single bits of information (e.g. sounds) forming temporal sequences such as a musical piece.

Thus, in our study we investigated the rapid whole-brain networks underlying the encoding of sounds forming a complex temporal sequence such as a full structured musical piece. On the one hand, understanding such topic is a key step to discover the brain mechanisms underlying encoding of music. On the other hand, our investigation may contribute to the understanding of the more general neural mechanisms underlying pattern encoding and attribution of meaning to sound information available in the external environment.

2. Materials and methods

2.1. Participants

Seventy volunteers participated in the study. However, the analyses were carried out on 68 of them since two have been excluded because of technical problems during part of the data acquisition. Thus, the final sample was composed by 68 participants (35 males and 33 females, age range: 18 – 42 years old, mean age: 24.88 ± 4.17 years). Since our experiment involved a well-known piano musical piece, we recruited 23 classical pianists (13 males and 10 females, age range: 18 – 34 years old, mean age: 24.83 ± 4.10 years), 21 non-pianist musicians (10 males and 11 females, age range: 22 – 39 years old, mean age: 24.29 ± 5.02 years) and 24 non-musicians (12 males and 12 females, age range: 21 – 35 years old; mean age: 25.46 ± 3.48 years).

All experimental procedures complied with the Declaration of Helsinki – Ethical Principles for Medical Research and were approved by the Ethics Committee of the Central Denmark Region (De Videnskabssetiske Komitéer for Region Midtjylland) (Klein et al., 2010-Vuust et al., 2012; Uddin, 2015; Leino et al., 2007).

2.2. Experimental design and stimuli

Participants' brain activity was recorded by using MEG. At first, we had a resting state session that has been used later as baseline for evaluating the brain functional connectivity during the task. Participants were required to sit down in the scanner for 10 minutes trying to relax but without closing their eyes. The room was dark, and participants were asked to fixate a cross on the screen and not to think about anything in particular. Then, to study the brain dynamics of sound encoding, we asked participants to actively listen to a MIDI version of the right-hand part of the entire prelude in C minor BWV 847 by Johann Sebastian Bach, as depicted in Figure 1a. To be noted, to have a musical piece with tones lasting approximately the same time and thus reduce the potential confounds introduced by rhythmic variety, we have slightly modified the last few bars of the Bach's prelude. In our version of the prelude, such bars were formed by tones with the same approximate duration. Importantly, the duration of each tone was on average 250ms, but it could vary in the short range: 220 - 280ms. This small variation was introduced to make the execution of the musical piece less artificial and more musical. Participants were required to try to memorize the prelude as much as possible. To facilitate their task, as well as to collect more data and increase the reliability of our findings, we played the musical piece four times. Afterwards, participants were engaged in a memory recognition task based on the Bach's prelude that they previously listened to. The recognition task has been thoroughly described in Bonetti et al. (Bonetti et al., 2020) and has been considered in this study as a validation task to ensure that participants actively listened to the Bach's prelude presented in our work and were able to distinguish

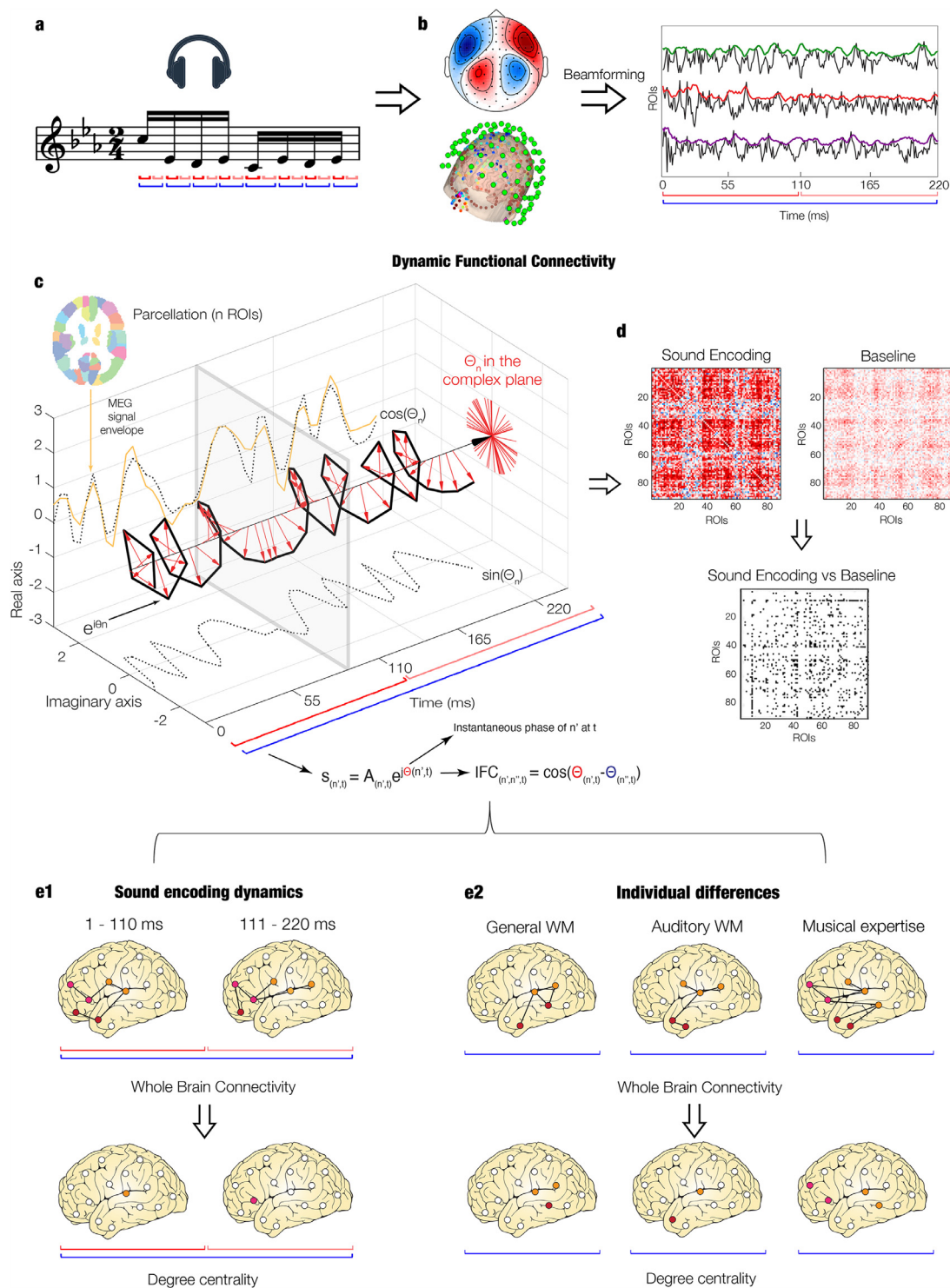


Fig. 1. Overview of the analysis pipeline

a – Participants were requested to attentively listen to a whole musical piece and to try to remember it as much as possible. As illustrated by the blue graphs, first we epoched and averaged together the brain signal underlying processing of each single tone. Then, first we focused our attention on the full time-window of sound encoding (1 – 220 ms) by computing analysis of brain activity and static functional connectivity (SFC). Second, we performed analyses on two main time-windows underlying sound encoding (as illustrated by the red graphs) by computing dynamic functional connectivity (DFC). **b** – MEG data during musical encoding has been collected, pre-processed and beamformed into source space within the 90 non-cerebellar brain regions of the AAL parcellation. Then, we calculated the envelope of the time-course of each brain region (top right). The red graphs show the two time-windows used in the DFC analysis (1 – 110 ms and 111 – 220 ms), while the blue refers to the full time-window used in the SFC. **c** – After computing analysis on brain activity and SFC, we investigated the DFC. To this aim, we computed the Hilbert transform of the envelope of each brain region and estimated the phase synchronization by calculating the cosine similarity between the instantaneous phases of each pair of brain regions. **d** – We obtained IFC matrices for both sound encoding task and resting state (used as baseline). Afterwards, we contrasted the task matrices versus the average of the baseline matrices to isolate the brain activity specifically associated to the sound encoding brain processes over time. **e1** – We used the whole-brain connectivity measures to derive the significant centrality of the brain regions within the whole-brain network. This was done for the two main time-windows of sound encoding (1 – 110 ms and 111 – 220 ms). **e2** – Similarly, contrasting participants who were grouped according to their level of WM and musical expertise, we computed the significant centrality of brain regions associated to those different skills.

it from different music when doing the recognition task a few minutes later.

The stimuli were designed by using Finale (MakeMusic, Boulder, CO) and then presented through Presentation software (Neurobehavioural Systems, Berkeley, CA). After the MEG recording, in the same or in another day, participants' brain structural images were collected by magnetic resonance imaging (MRI) exam. Furthermore, participants' general and auditory WM abilities and musical expertise were assessed. Specifically, their musical expertise was collected using the Gold-MSI questionnaire (Baker et al., 2018; Müllensiefen et al., 2014). With regards to general WM we adopted one of the most used psychological tests for assessing cognitive abilities, namely Wechsler Adult Intelligence Scale - IV (WAIS-IV) (Dumont et al., 2014; Wechsler, 1997), while for the auditory WM abilities we employed the Musical Earing Test (MET) (Wallentin et al., 2010), a newly developed tool that presents pair of complex melodies requiring participants to state whether they are the same or different.

2.3. Data acquisition

We acquired MEG and MRI data in two independent sessions. The MEG data were collected using an Elekta Neuromag TRIUX system (Elekta Neuromag, Helsinki, Finland) equipped with 306 channels. The scanner was located in a magnetically shielded room at the Aarhus University Hospital, Denmark. The data were collected at a sampling rate of 1000 Hz with an analogue filtering of 0.1–330 Hz. Before the exam, we adjusted the sound volume at 50 dB above the participants' minimum hearing threshold. Furthermore, by using a 3D digitizer (Polhemus Fastrak, Colchester, VT, USA) we recorded the participant's head shape and the continuous location of four headcoils, with respect to three anatomical landmarks (nasion, and left and right preauricular points). These data were then utilized to ensure a high-quality co-registration of the MEG data with the anatomical structure obtained during the MRI exam, at a later stage of the analysis pipeline.

The location of the headcoils was registered during the whole recording session using a continuous head position identification (cHPI) and therefore we tracked the exact head position within the MEG scanner at each moment. This allowed us to perform an accurate movement correction at a later stage of data analysis.

The MRI data consisted in structural T1. The acquisition parameters for the scan were: voxel size = $1.0 \times 1.0 \times 1.0$ mm (or 1.0 mm³); reconstructed matrix size 256×256 ; TE of 2.96 ms and TR of 5000 ms and a bandwidth of 240 Hz/Px.

2.4. Data pre-processing

We conducted Maxfilter (Taulu and Simola, 2006) noise reduction on the raw MEG sensor data (204 planar gradiometers and 102 magnetometers) for attenuating the interference that originated outside the scalp by applying signal space separation. Moreover, Maxfilter allowed us to correct for head movement and down-sample the data from 1000 Hz to 250 Hz.

The data were converted into Statistical Parametric Mapping (SPM) format and further processed in Matlab (MathWorks, Natick, Massachusetts, United States of America) by utilizing Oxford Centre for Human Brain Activity Software Library (OSL), a freely available toolbox that relies on a combination of Oxford Centre for fMRI of the Brain Software Library (FSL) (Woolrich et al., 2009), SPM (Penny et al., 2007) and Fieldtrip (Oostenveld et al., 2011), and in-house-built functions.

The data were high pass filtered (0.1 Hz threshold) to remove too low frequencies for being originated by the brain. We also used a notch filter (48–52 Hz) to control for interference of the electric current. The data were then down-sampled again to 150 Hz and few segments of the data, altered by large artifacts, were discarded after visual inspection. Then, to correct for eyeblinks and heartbeat artifacts, we calculated independent component analysis (ICA) to decompose the original signal

in independent components. Then, we individuated and discarded the components that picked up the heartbeat and eyeblink activities and we reconstructed the signal by using only the remaining components (Mantini et al., 2011). Finally, data were epoched according to the beginning of each of the 605 musical tones of the prelude (pre-stimulus time of 100 ms and post-stimulus time of 220 ms) and baseline corrected by removing the mean value of the pre-stimulus baseline from the entire trial. Therefore, our trials were represented by the segment of the signal starting with the onset of each musical tone and finishing with the offset of the same tone. This procedure was carried out also for the resting state. As conceivable, the resting state did not have any external stimulation, therefore we created trials in equal length and number at pseudorandom time-points of the recorded data.

2.5. Event related fields and power spectra analysis

Prior to performing connectivity analysis, as illustrated in Figure 1b, we tested the quality of our data by assessing the ERF and especially the N100, a well-known component of the ERF arising 100–150 ms after sound stimulation (Näätänen and Picton, 1987). To this purpose, we averaged together all trials obtained after epoching the data and combined planar gradiometers by mean root square (Bruno and Romani, 1989). Then, we calculated a t-test for each MEG gradiometer channel and time-point between the ERF to the sound and the averaged pre-stimulus brain activity. To correct for multiple comparisons, we adopted MCS (Kroese et al., 2011). Specifically, we reshaped the previously calculated statistics for obtaining, for each time-point, a two-D approximation of the MEG channels layout and we binarized it according to the *p*-values obtained from the previous t-tests (threshold $\alpha = 1.0e-12$). The resulting three-D matrix (*M3*) was therefore composed by 0s when the t-tests were not significant and 1s when they were. Then, we made 1000 permutations of the elements of the original binary matrix *M3*, identified the maximum cluster size of 1s for each permutation and built the distribution of the 1000 maximum cluster sizes. Finally, we considered significant the original clusters that had a size bigger than the 99.9% of the permuted data maximum cluster sizes.

To assess the contribution of each frequency, we estimated the power spectra of the brain signal by employing complex Morlet wavelet transform (from 1 to 40 Hz with 1-Hz intervals) (Daubechies, 1992). Then, we calculated a t-test for each time-point within the range: 0.050–0.200 seconds and the averaged power spectra of the 100ms pre-stimulus baseline. Emerging *p*-values were binarized according to threshold $\alpha = 1.0e-18$ and then submitted to a two-D MCS. Specifically, we calculated the clusters size of continuous significant values in frequency and time and then made 10000 permutations of the binarized *p*-values. For each permutation we detected the size of the maximum emerging cluster and built a reference distribution with one value for each permutation. Then, we considered significant the original clusters that had a size bigger than the 99.99% of the permuted data maximum cluster sizes.

Thresholds for binarizing the *p*-values matrices were very low since we were comparing the brain activity versus baseline and therefore, as conceivable, the results were highly significant and, to individuate the strongest contribution of MEG channels, time-points and frequencies, it was necessary to select very low thresholds.

2.6. Source reconstruction

As depicted in Figure 1b, the brain activity recorded on the scalp by the MEG sensors was reconstructed in source space. First, each individual T1-weighted MRI scan was co-registered to the standard Montreal Neurological Institute (MNI) template brain through an affine transformation and referenced to MEG sensors space by employing the Polhemus head shape data and the three fiducial points collected during MEG session. Second, the MRI co-registered image was used in the source reconstruction procedure that has been implemented separately for each par-

ticipant, as commonly done in the field (Hillebrand and Barnes, 2005; Huang et al., 1999; Brookes et al., 2007).

Our source reconstruction algorithm used an overlapping-spheres forward model and a beamformer approach as inverse model (Hillebrand and Barnes, 2005), with an eight-mm grid and both planar gradiometers and magnetometers. The spheres model represented the MNI-co-registered anatomy as a simplified geometric model, fitting a sphere separately for each MEG sensor (Kroese et al., 2011). The beamforming used a different set of weights sequentially applied to the source locations for individuating the contribution of each source to the activity recorded by the MEG sensors at each time-point (Hillebrand and Barnes, 2005; Hunt et al., 2012). To be noted, using an eight-mm grid we reconstructed the MEG signal in 3559 brain sources.

To assess the brain activity associated to the sound encoding task we submitted the beamformed reconstructed activity to first-level statistical analysis carried out by calculating a GLM for each time-point and at each reconstructed brain source (Hunt et al., 2012). Then, after calculating the absolute value of the reconstructed time-series to avoid sign ambiguity of the neural signal, we conducted group-level analysis, using one-sample t-tests with spatially smoothed variance obtained with a Gaussian kernel (full-width at half-maximum: 50 mm). Finally, to correct for multiple comparisons, a cluster-based permutation test (Hunt et al., 2012) with 5000 permutations has been calculated on group-level analysis results. Considering an α level = 0.05, we used a cluster forming threshold t -value = 1.7.

2.7. Static functional connectivity

To investigate functional connectivity, we used the source localized data obtained by the beamforming algorithm. Then, these data were constrained from the 3559 reconstructed brain sources into the 90 non-cerebellar regions of the AAL parcellation, a freely and widely used available template (Tzourio-Mazoyer et al., 2002), in line with previous MEG studies (Brookes et al., 2016; Cabral et al., 2014; Hindriks et al., 2015). This procedure was very important to reduce the dimensionality of the data and focus on a well-known parcellation of the brain which helped to directly compare our results with previous studies. To note, we performed this algorithm after applying a bandpass filter correspondent to 2 – 8 Hz to the data. This frequency band interval was the one characterized by the highest power, according to the previous time-frequency analysis described two paragraphs above. In particular, we chose 2 – 8 Hz instead of 2 – 5 Hz, as emerging from the power spectra statistics, to have a broader frequency range usually employed in studies on theta waves. Then, we proceeded with static functional connectivity analysis performed using Pearson's correlations. Since the length of our epoched data were quite short (36 time-samples with our sampling rate of 150 Hz, corresponding to 220 ms), to estimate more reliable SFC through Pearson's correlations, we concatenated and sub-averaged groups of trials. Specifically, our 605 trials were concatenated in seven chunks of data with 86 trials each (89 for one of the chunks) and then averaged together to obtain seven sub-averaged concatenated trials leading to a new matrix M . Such matrix M had dimensions: 90 brain regions x 252 time-points (seven concatenated sub-averaged trials of 36 time-points each: $36 \times 7 = 252$). Then, we performed source leakage correction (Colclough et al., 2015) by orthogonalization and calculated Pearson's correlations between the envelope (Cabral et al., 2014) of the time-series of each pair of brain areas. This procedure was carried out for both task and resting state (used as baseline) and resulted in two 90×90 matrices for each participant, one for the task and one for the baseline. Those two matrices were contrasted by applying Wilcoxon signed-rank test for each pair of brain areas. The resulting z-values matrix Z was submitted to a degree MCS for assessing which brain area was significantly central within the brain network, after contrasting task versus rest. In graph theory, the degree of each vertex v (here each brain area) of the graph G (here the matrix Z describing the whole-brain functional connectivity) is given by summing the connection strengths of v

with the other vertexes of G , returning a value of the centrality of each v in G (Rubinov and Sporns, 2010). In this MCS, we computed the degree of each vertex of Z , obtaining a 90×1 vector (s_i). Then, we made 10000 permutations of the elements in the upper triangle of Z and we computed a 90×1 vector $d_{v,p}$ containing the degree of each vertex v for each permutation p . Combining vectors $d_{v,p}$ we obtained the distribution of the degrees calculated for each permutation. We considered significant the degrees stored in s_i that were higher than the 99.9% of the degree distribution values calculated by permuting Z 10000 times. The threshold of 99.9% derived from simulations of the MCS function with matrices composed by uniformly distributed random values. Setting a 99.9% threshold yielded to a number of false positive nearly equal to zero, while a more common 95% threshold gave rise to few false positives.

2.8. Phase synchronization estimation

To unravel the brain dynamics of the sound encoding and thus refine our understanding of the dynamic changes of connectivity over time, we studied the phase synchronization over time between brain areas for theta band.

By applying Hilbert transform (Layer and Tomczyk, 2015) on the envelope of the reconstructed time-courses (matrix M described in the paragraph above) we obtained the analytic signal $S_{(n_i,t)}$ expressed by the following equation:

$$S_{(n_i,t)} = A_{(n_i,t)} e^{j\theta_{(n_i,t)}} \quad (1)$$

Where $A_{(n_i,t)}$ refers to the instantaneous amplitude and $\theta_{(n_i,t)}$ to the instantaneous phase of the signal for brain region n_i at time t . A graphical depiction of Hilbert transform is reported in Figure 1c. Then, since matrix M was made up by seven concatenated sub-averaged trials, after estimating the instantaneous phase, we discarded the time-samples corresponding to the first and last trials to prevent boundary artefacts introduced by instantaneous phase estimation and we averaged the remaining five, obtaining a new matrix $M2$ composed by the 90-brain region instantaneous phases x 36 time-samples. To estimate the phase synchronization between two brain areas n_i and n_m of the matrix $M2$ at time t , we calculated the cosine similarity expressed by equation (2):

$$IFC_{(n_i,n_m,t)} = \cos(\theta_{(n_i,t)} - \theta_{(n_m,t)}) \quad (2)$$

We carried out this procedure for each time-point and each pair of brain areas of the matrix $M2$, obtaining one 90×90 symmetric instantaneous functional connectivity (IFC) matrix for each time-point showing the phase synchronization of every pair of brain areas. This procedure was carried out for both task and resting state and is illustrated in Figure 1d.

2.9. Brain dynamics of sound encoding

To estimate the instantaneous connectivity matrix Z_t associated to the task, first we averaged over time the IFC matrix computed for the rest. Second, for each time-point t_i , we contrasted the task IFC matrix at time t versus the resting state IFC matrices by using Wilcoxon sing-rank test. Then, as described above, an MCS computed on Z_t assessed the significantly central brain regions within the brain network for each time-point. We refer to this measure as instantaneous brain degree (IBD_i). Moreover, the sum ($sIBD$) over time of IBD_i for the brain area n showed us its centrality within the whole-brain network during the whole sound encoding process, as expressed by equation (3):

$$sIBD_{(n)} = \sum_{t=1}^T IBD_{(n,t)} \quad (3)$$

Then, we focused our attention towards only two main, equally long time-windows (Figure 1e1): 1 - 110 ms and 111 - 220 ms.

Our interest in working on such fast time-windows related to the neural relevance of such timing. Indeed, as pointed out by

Dehaene et al. (2011), the first 200 – 250 ms of elaboration of stimuli seem to play a crucial role for making temporal information available in human awareness. Thus, we were particularly interested in exploring the brain connectivity dynamics of sound encoding in this framework. On top of this, we decided to break this short window down into two further time-windows slightly longer than 100 ms each. Since 100 ms is a key timing of ERF components (e.g. onset of N100, interval between N100 and P200), we thought was of particular interest to look into the underlying functional connectivity. On top of this, considering the noise inherent in the data and potentially added by our elaborated signal processing pipeline, we believed that looking into further and shorter time-windows may have increased the probability of detecting non-reliable findings.

Back to our analysis, we were interested in comparing the overall degree centrality of each of our 90 brain areas in the first versus second time-window. The degree centrality was expressed by the two 90 x one vectors named: $sIBD_{(n)1-110ms}$ and $sIBD_{(n)111-220ms}$. Then, for each brain region n we calculated the vector difference $vt do$, as expressed by equation (4):

$$vt do = sIBD_{(n)1-110ms} - sIBD_{(n)111-220ms} \quad (4)$$

Finally, to assess whether vector $vt do$ contained any significant value, we used a further MCS approach. Specifically, we made 10000 permutations of the matrices $IBD_{1-110ms}$ and $IBD_{111-220ms}$ and we summed the two matrices, independently, over the temporal dimension. This returned two 90 x one vectors, $vt o_p$ and $vt v_p$. Subtracting $vt v_p$ from $vt o_p$ we obtained the difference vector $vt d_p$ for the permutation p . Performing this procedure for every permutation, we obtained the distribution of the difference vectors. Then, we compared the original differences in vector $vt do$ with the distribution of difference vectors, independently for positive and negative differences, to get their associated p -values. Finally, we considered significant the brain areas whose p -values were lower than:

$$\alpha' = \alpha / (ROIs * tails) = 2.7e - 04 \quad (5)$$

where α corresponds to α level = 0.05, ROIs are the number of brain regions (90) and *tails* refers to the two tails of the normal distribution of difference vectors created by MCS (which corresponds to positive and negative differences in vectors $vt d_p$). In other words, since our hypothesis was that some brain regions were different in the two time-windows, but we did not hypothesize which ones, we looked at the results for each brain region and each direction of the difference (time-windows one > time-windows two and vice versa). Therefore, we had to correct for multiple comparisons by calculating the new strict threshold α' expressed by equation (5).

2.10. Brain dynamics of sound encoding and individual differences

In conclusion, as illustrated in Figure 1e2, we aimed to assess whether the neural networks underlying sound encoding differed across participants grouped according to general and auditory WM, and musical expertise. To highlight more clearly the differences, we selected for each WM test two groups formed by participants whose scores were at least one SD apart from each other. According to psychometric guidelines (Taylor and Heaton, 2001), this procedure is particularly relevant when aiming to clearly distinguish participants based on their WM (or general cognitive) abilities as measured by the test. Indeed, such approach was adopted since we wanted to include only participants that were clearly differentiated by the tests. With regards to the WAIS test groups the best scorers had a range of 110 – 130, while the worst of 76 – 93 (according to the standardization of the WAIS test one SD corresponds to 15). In relation to MET test, we observed that our best scorers had a range of 43 – 52 while worst scorers of 28 – 36. The mean across all participants was 40.24 with a standard deviation of 6.28. Thus, also in this case the two groups were differentiated by at least one SD. Finally, for musicianship, we considered the 24 non-musicians and the

musicians (both pianists and non-pianists) that received a formal musical education for at least 10 years, a threshold widely suggested by previous literature (e.g.) (Sloboda et al., 1996). Those participants were 24. This threshold was set in order to compare people with no musical expertise at all with individuals who engaged in a long-term professional education. In addition to those three measures of individual differences, we have computed a further analysis to control whether the previous familiarity of the Bach's prelude (i.e. participants who knew our Bach's prelude versus participants who had never heard it before the experiment) modulated the brain connectivity patterns during sound encoding. In this case, we had 34 participants who were already familiar with the Bach's prelude and 31 who had never heard it before the experiment. One participant did not answer this question and thus was excluded from the analysis. Finally, as described in the previous section, we focused on the number of times that each brain region was significantly central within the brain, contrasting those values across the two groups by an MCS analogous to the one described above. This procedure was carried out independently for the two WM and the musicianship analyses (since we had three independent tests, we divided the threshold α' described by equation (5) by four, obtaining a new threshold = 6.8e-05). In Figure SF1, to provide full information, brain regions centrality is depicted also for the remaining participants.

3. Results

3.1. Experimental design and data analysis overview

In our study, we aimed to investigate the fine-grained spatiotemporal dynamics of the brain during the encoding of sounds forming a full, structured musical piece. To this aim, we used MEG to record the brain functioning of 68 participants while they listened to a slightly edited musical instrumental digital interface (MIDI) version of the full prelude in C minor BWV 847 composed by Johann Sebastian Bach.

As described in the Methods and depicted in Figure 1a, participants were requested to attentively listen to the music, trying to memorize its structure and sounds as much as they could. The analysis pipeline employed in our study is depicted in Figure 1 (and described in detail in the Methods). Our results on brain functioning underlying sound encoding have been organised as follows: 1) sensor space and beamformed source localised activity, 2) static source localised connectivity, 3) dynamic source localised connectivity and 4) dynamic source localised connectivity in subsamples characterized by different levels of general (GWM) and auditory working memory (AWM) and musical expertise.

First, we detected the brain activity in MEG sensor space using univariate tests and Monte Carlo simulations (MCS). Then, we reconstructed the sources of the brain signal using a beamforming algorithm (Figure 1b). Second, we computed the static functional connectivity (SFC) by calculating Pearson's correlations between the envelope of each pair of brain areas. Third, we computed dynamic functional connectivity (DFC) using the instantaneous phase obtained from Hilbert transform for each time-point of the brain areas timeseries (Figure 1c). After contrasting the brain connectivity patterns for sound encoding versus resting state (Figure 1d), we computed the DFC for two short time-windows (1 – 110 ms, 111 – 220 ms), as depicted in Figure 1e1. Fourth, we analysed the DFC in subsamples of individuals characterized by different levels of GWM, AWM and musical expertise (Figure 1e2). DFC analysis consisted of detecting the whole-brain connectivity patterns and then the significant centrality of specific brain areas within the whole-brain network.

Notably, we employed a wide set of analytical techniques, ranging from the well-established investigation of brain activity at MEG sensor space and after source localisation to more recent approaches such as DFC. We believe that covering several different analytical techniques with different strengths and limitations is of great importance since allowed at the same time to obtain a collection of coherent results and explore new analytical solutions. Indeed, on the one hand our results

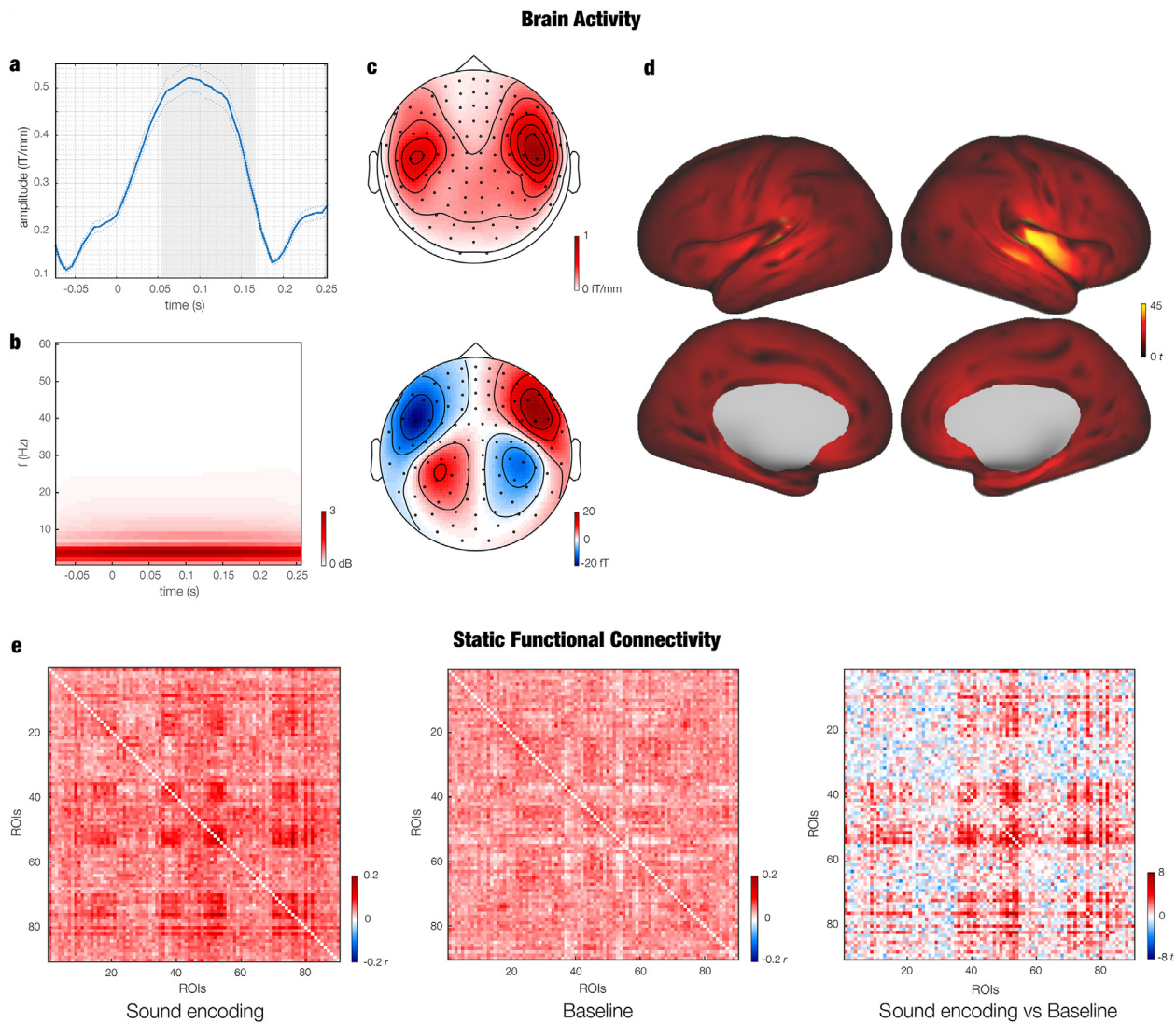


Fig. 2. *N100 component*

a – Waveform depicting the N100 component of the ERF. The plot shows the brain activity averaged over the significant gradiometer channels emerged by MCS. The grey area illustrates the significant time-window. Thinner lines depict standard errors. **b** – Power spectra depiction over time for all MEG channels. **c** – Gradiometers (top) and magnetometers (bottom) topoplots of the N100 component amplitude computed over the significant time-window emerged by MCS. Values correspond to neural signal in fT/mm for gradiometers and fT for magnetometers. **d** – Neural sources of the N100 component (t-values from group level analysis). **e** – SFC estimated by computing Pearson's correlations between the envelope of each pair of AAL brain regions. The left matrix refers to sound encoding task, while the middle one to resting state (used as baseline). Finally, the right matrix depicts the t-values emerged by contrasting task versus baseline.

confirmed and expanded the existent literature on the sound encoding in the brain by using very solid and established methodologies such as the investigation of brain activity and SFC. On top of this, we added a further layer of analyses which involved DFC. As shown in the following paragraphs, these last algorithms allowed us to complement our original analyses and obtain new exciting perspectives with a less common approach.

3.2. Event related fields and power spectra analysis

At first, as depicted in **Figure 2a** and **2c** to assess the quality of the data, we analysed the ERF associated to the processing of the tones. To this aim, after carrying out standard pre-processing steps (see Methods for details), we epoched the data in correspondence to each musical tone (trials), we averaged the trials and combined planar gradiometers by sum root square. Then, we calculated a t-test for each time-point in the time-range 0.050 – 0.200 seconds and for each gradiometer channel contrasting the task versus its own baseline. Finally, we performed cluster-based MCS to detect significant clusters (t-test threshold = 1.0e-

16, MCS threshold = 0.001). Results showed two significant clusters in the time-range 0.053 – 0.160 seconds. Specifically, we observed a larger cluster in the right hemisphere (cluster size: 81, $p < 0.001$) and a smaller one in the left (cluster size: 40, $p < 0.001$). These results are reported in detail in **Table ST1**.

Then, we reconstructed the neural sources of the signal by applying a beamforming approach and calculating a general linear model (GLM) for each of the 3559 reconstructed brain sources and each time-point within the significant time-range emerged by MCS on sensor data (0.053 – 0.160 seconds). Significant clusters were assessed by a cluster-based permutation test. **Figure 2d** shows that the activity was mainly localized within primary auditory cortex and insula. Complete results are reported in **Table ST2**. Finally, to restrict our subsequent functional connectivity analysis to a specific frequency band, we performed the power spectra associated to the task by complex Morlet wavelet transform (from 1 to 60 Hz with 1-Hz intervals) to define which frequencies were mainly involved in the sound processing. We then calculated t-tests for each frequency and time-point within the range 0.050 – 0.200 seconds and the averaged power spectra of the baseline. Binarized val-

ues (threshold = 1.0×10^{-18}) were submitted to a two-dimensional (two-D) MCS (threshold = 0.001). As reported in **Figure 2b**, the analysis returned a significant cluster (size: 69, $p < 0.001$) for frequency range: 3 – 5 Hz within the time-range: 0.053 – 0.200 seconds.

3.3. Static functional connectivity

We constrained the reconstructed sources of the MEG signal from the 3559 brain sources obtained from the beamforming algorithm to the 90 non-cerebellar parcels of automated anatomical labelling (AAL) parcellation. Since the length of our epoched data were quite short (36 time-samples with our sampling rate of 150 Hz, corresponding to 220ms), to estimate more reliable SFC through Pearson's correlations, we concatenated and sub-averaged groups of trials. This procedure returned a final time-series matrix M , made up by seven concatenated sub-averaged trials, with dimensions: 90 brain regions $\times$ 252 time-points. Then, we performed source leakage correction by orthogonalization (Colclough et al., 2015) and calculated Pearson's correlations between the envelope of the time-series of each pair of brain areas. This procedure was carried out for both task and resting state (used as baseline) and resulted in two 90×90 matrices for each participant, one for the task and one for the rest. Those two matrices were contrasted by applying Wilcoxon signed-rank test for each pair of brain areas. The resulting z -values matrix Z was submitted to a degree MCS (MCS threshold = 0.001) for assessing which brain area was significantly central within the brain network, after contrasting task versus rest.

As depicted in **Figure 2e** and **Figure SF2**, the analysis returned a significant centrality of the following brain regions: left precentral ($p = 3.5 \times 10^{-5}$), Rolandic operculum ($p < 1.0 \times 10^{-7}$), caudate ($p < 1.0 \times 10^{-7}$), putamen ($p = 4.4 \times 10^{-6}$), thalamus ($p < 1.0 \times 10^{-7}$), Heschl's gyrus ($p = 3.0 \times 10^{-5}$), temporal superior ($p < 1.0 \times 10^{-7}$), right temporal pole middle ($p = 6.6 \times 10^{-6}$), temporal pole superior ($p < 1.0 \times 10^{-7}$), temporal superior ($p = 1.8 \times 10^{-5}$), Heschl's gyrus ($p < 1.0 \times 10^{-7}$), thalamus ($p < 1.0 \times 10^{-7}$), pallidum ($p < 1.0 \times 10^{-7}$), putamen ($p < 1.0 \times 10^{-7}$), amygdala ($p < 1.0 \times 10^{-7}$), hippocampus ($p < 1.0 \times 10^{-7}$), insula ($p < 1.0 \times 10^{-7}$), frontal medial orbital cortex ($p = 6.5 \times 10^{-4}$), subgenual ($p < 1.0 \times 10^{-7}$), Rolandic operculum ($p < 1.0 \times 10^{-7}$), frontal inferior operculum ($p = 8.2 \times 10^{-5}$).

3.4. Dynamic functional connectivity

To unravel the dynamics of the functional connectivity during sound encoding, we calculated the instantaneous phase of the signal envelope of the matrix M described in the previous paragraph by applying Hilbert transform. Then, since matrix M was made up by seven concatenated sub-averaged trials, after estimating the instantaneous phase, we discarded the time-samples corresponding to the first and last trials to prevent boundary artefacts introduced by instantaneous phase estimation and we averaged the remaining five, obtaining a new matrix $M2$ composed by the 90-brain region instantaneous phases $\times$ 36 time-samples. Afterwards, to estimate the phase synchronisation between each pair of brain areas we computed the cosine of the difference (cosine similarity) of those instantaneous phases (**Figure 1c**). This procedure resulted in a 90 brain regions $\times$ 90 brain regions DFC matrix for each time-point and each condition (task and rest). Then, analogous to the SFC analysis, to estimate the instantaneous connectivity specifically associated to the task, we contrasted the sound encoding versus rest DFC matrices using Wilcoxon sing-rank test. Then, as described above, a degree MCS assessed the significantly central brain regions within the brain network underlying the sound encoding task. Finally, in this case we wanted to investigate the different degree centrality of the brain areas in the two subsequent time-windows that we defined in the study (1 – 110 ms and 111 – 220 ms). We have performed this by contrasting the degree centrality of the 90 brain areas for time-window one versus time-window two and correcting for multiple comparisons by using MCS. As illustrated in **Figure 3** and **Figure SF3**, the results highlighted that right Rolandic operculum ($p < 1.0 \times 10^{-7}$) and Heschl's gyrus ($p = 2.2 \times 10^{-6}$) were

more central within the first 110ms, while right insula ($p = 2.2 \times 10^{-6}$) and superior temporal pole ($p = 2.2 \times 10^{-6}$) were more central in the second time-window.

3.5. Dynamic functional connectivity and individual differences

In conclusion, we aimed to assess whether the neural networks activated during sound encoding differed across participants grouped according to WM abilities (general, assessed by WAIS-IV, and auditory, assessed by MET) and musical expertise. To highlight more clearly the differences, we selected for each WM test two groups formed by participants whose scores were at least one SD apart from each other. According to psychometric guidelines (Taylor and Heaton, 2001), this procedure is particularly relevant when aiming to clearly distinguish participants based on their WM abilities as measured by the test. Indeed, such approach was adopted since we wanted to include only participants that were clearly differentiated by the tests. With regards to the WAIS test groups the best scorers had a range of 110 – 130, while the worst of 76 – 93 (according to the standardization of the WAIS test one SD corresponds to 15). In relation to MET test, we observed that our best scorers had a range of 43 – 52 while worst scorers of 28 – 36. The mean across all participants was 40.24 with a standard deviation of 6.28. Thus, also in this case the two groups were differentiated by at least one SD. Finally, for musicianship, we considered the 24 non-musicians and the musicians (both pianists and non-pianists) that received a formal musical education for at least 10 years. Those participants were 24. This threshold was set to compare people with no musical expertise at all with individuals who engaged in a long-term professional education. In addition to those three measures of individual differences, we have computed a further analysis to control whether the previous familiarity of the Bach's prelude (i.e. participants who knew our Bach's prelude versus participants who had never heard it before the experiment) affected the brain connectivity patterns during sound encoding. Thus, for each of the three variables (AWM, GWM and musical expertise), we calculated independent contrasts between the groups and corrected for multiple comparisons by using MCS.

Figure 4 shows that higher GWM was associated to higher centrality of right Rolandic operculum ($p = 1.2 \times 10^{-5}$) and lower GWM to left ($p = 6.4 \times 10^{-5}$) and right putamen ($p = 6.4 \times 10^{-5}$). With regards to AWM skills, the best participants reported higher centrality of right ($p < 1.0 \times 10^{-7}$) and left insula ($p = 6.6 \times 10^{-6}$), left frontal middle orbital cortex ($p = 1.1 \times 10^{-6}$), right temporal middle gyrus ($p = 6.6 \times 10^{-6}$) while worst ones had a higher centrality of right occipital inferior ($p < 1.0 \times 10^{-7}$), occipital superior ($p = 6.4 \times 10^{-5}$) and frontal medial orbital cortex ($p = 6.4 \times 10^{-5}$). Finally, musicians exhibited higher centrality of right insula ($p < 1.0 \times 10^{-7}$), subgenual cortex ($p = 5.7 \times 10^{-5}$), left supplementary motor area ($p = 1.1 \times 10^{-6}$), while non-musicians of right caudate ($p < 1.0 \times 10^{-7}$) and occipital inferior ($p < 1.0 \times 10^{-7}$), as illustrated in **Figure 4**. Finally, the last analysis on familiarity with the Bach's prelude showed that participants who had never heard the prelude before the experiment presented a significantly stronger centrality of the right medial orbito-frontal cortex ($p < 1.0 \times 10^{-7}$) when encoding the sounds. Additional information about WM, analysis methods and sound encoding brain networks related to all participants are provided in the Methods section and in Supplementary Materials (**Figure SF1**).

Depiction of the connectivity underlying sound encoding between the significantly different brain regions emerged contrasting participants' groups and the rest of the brain. For each of the five columns, the connectivity is depicted within brain templates (top) and in schemaballs (middle). For each pair of brain templates, the left one is a posterior refiguration from the left hemisphere, while the right one a posterior representation. Finally, we depicted within brain templates the significant differences of brain regions' centrality obtained by contrasting different groups of participants (bottom of each column). The top left column refers to all participants to have an immediate visual comparison, the top middle column represents participants grouped for general WM abil-

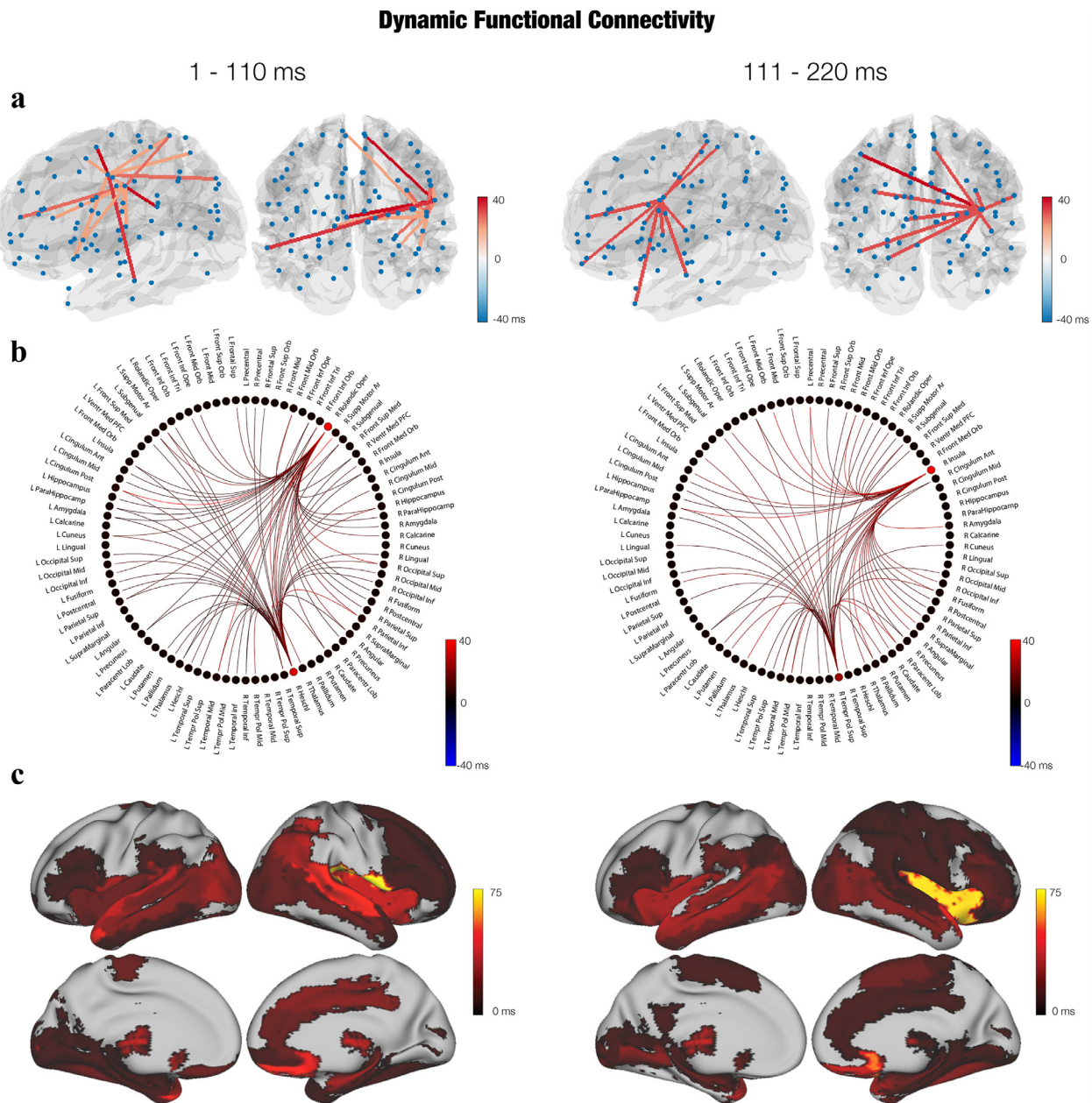


Fig. 3. Dynamic functional connectivity of sound encoding

a – Connectivity underlying sound encoding depicted within a brain template showing the strongest connections between the significantly central brain regions and the rest of the brain during two main rapid time-windows (1 – 110 ms on the left and 111 – 220 ms on the right). For each pair of brain templates, the left one is a refiguration from the left hemisphere, while the right one a posterior depiction. **b** – Schemaball showing the strongest connections between the significantly central brain regions and the rest of the brain during the two time-windows. **c** – Significantly central brain regions within the whole brain network during sound encoding depicted in the two time-windows. In all depictions, the colorbar values refer to the temporal extent (in ms) of the brain regions' significance.

ities, the top right relates to auditory WM abilities. Then, the bottom left column shows the contrast between musicians and non-musicians and the bottom right column shows the difference between participants familiar versus non-familiar with the Bach's prelude. The colorbars depict high WM scorers, musicians and participants familiar with the Bach's prelude with red shades, while low WM scorers, non-musicians and participants non-familiar with the Bach's prelude with blue shades. Values show the temporal extent (in ms) of the significant differences.

4. Discussion

In this study, we investigated the rapid spatiotemporal brain mechanisms for encoding of sounds forming a full, structured musical piece, as

compared to resting state. Notably, investigating the brain functioning within the first 220 ms after sounds onset, we detected the brain mechanisms that may be partly responsible for making temporal information available to human awareness (Dehaene et al., 2011).

We detected significant activation and centrality, primarily in the right hemisphere, of several brain regions linked to memory, attentional and auditory processes such as primary auditory cortex, frontal operculum, basal ganglia, insula and hippocampus. Additional analysis employing phase synchronization and therefore dynamical changes over time of the connectivity patterns highlighted stronger centrality of auditory cortex regions such as right Heschl's and superior temporal gyri as well as frontal operculum within the first 110 ms of the processing of each sound. Conversely, the second time-window that we defined

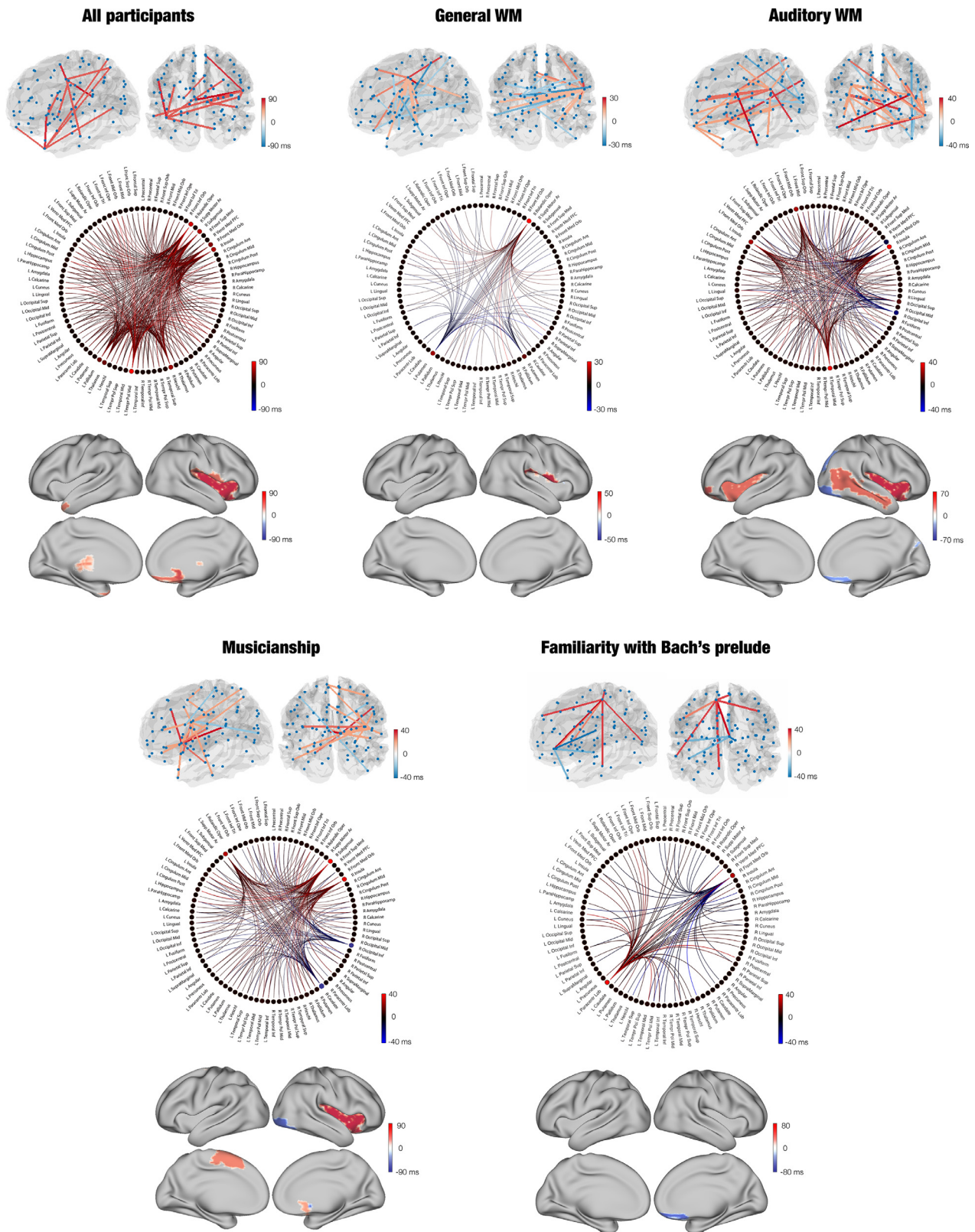


Fig. 4. Brain dynamics and general and auditory WM, musical expertise, and familiarity with the Bach's prelude.

(110 – 220 ms) showed a higher centrality of right insula and superior temporal pole. In conclusion, we presented results connecting individual differences among WM skills and musical expertise, and the brain network underlying sound encoding. Specifically, stronger auditory WM

skills and musical expertise were linked to higher centrality of subgenual cortex, insula and supplementary motor area, while higher general WM abilities were connected to stronger centrality of right frontal operculum.

4.1. Brain activity, static functional connectivity and sound encoding

Coherent with our results, primary auditory cortex has been widely shown responsible for processing of sound stimulations by a plethora of well-known studies (Winer and Schreiner, 2011; Zatorre, 2003). Additionally, we detected the strongest activity within the right hemisphere, another result largely described and replicated by previous studies (Tervaniemi and Hugdahl, 2003). Although our focus was on brain connectivity, replicating the activation of the auditory cortex in response to sound stimulation was a first necessary sanity check to be made. Moving toward connectivity analysis, we observed a rather interesting picture. While the brain activity was largely reconstructed on the primary auditory cortex and insula, the brain static connectivity analysis revealed a large set of higher-order brain areas that were highly correlated between each other and central within the whole-brain network. Notably, this result survived the correction for source leakage (Colclough et al., 2015), an issue with MEG source reconstruction that could have altered these findings. Conversely, this evidence strongly suggests that to process and encode sounds the sole activation of primary auditory cortex may need to be complemented by a further brain mechanism, which is the synchronization of a subnetwork of brain areas. Indeed, such areas, even if not strongly active, presented very similar time series which may be interpreted as the neural marker of a dense communication happening within this subnetwork of areas (Yoo et al., 2018). Remarkably, looking at the brain regions forming this subnetwork, we found that most of them were areas not directly implicated in auditory processing. For instance, we detected brain regions such as hippocampus, an area repeatedly connected to memory encoding (Frank et al., 2000; Preston and Eichenbaum, 2013; Vago et al., 2014) and frontal operculum, a brain region that has been linked to linguistic production and processing (Behroozmand et al., 2015; Indefrey et al., 2001). Considering the main functions of these areas highlighted by previous literature, we argue that their involvement in our task may relate to the encoding of the information carried by the sounds. Another key central area that we detected was insula, whose involvement could be related to the salience of the stimuli to be encoded, coherent with studies that showed the role of insula in processing stimulation salience (Cauda et al., 2011; Uddin, 2015). Basal ganglia also played a role in the sound encoding brain network. These subcortical regions have been shown important for several different tasks and are likely most known for their involvement in motor activities and associative learning (Ashby et al., 2010; Packard and Knowlton, 2002). Since in this study participants were actively attempting to memorize the musical piece, basal ganglia centrality within the brain may be interpreted as a sign of the learning process occurring while listening to the piece. In conclusion, our results suggested the relevance to conduct both activity and connectivity analyses to obtain a more complete picture of the sound encoding processes.

On another relevant note, it is important to underline that in our task participants actively listened to the musical piece and tried to memorize it as much as possible. This is not a typical approach used in 'free/naturalistic' listening to music (Taruffi et al., 2017; Alluri et al., 2012; Haumann et al., 2021) and has been employed in our study since we wanted to characterize the specific functional connectivity patterns of active sound encoding. Previous studies on naturalistic listening to music highlighted large networks of brain areas involved in passive music processing (Toivainen et al., 2014; Toivainen et al., 2020). Interestingly, those studies reported the insurgence of brain networks which shared similarities with the default-mode network (Taruffi et al., 2017), probably because passive listening to music led participants towards mind-wandering states. Other analogous studies reported activity in motor areas of the brain (Alluri et al., 2012) and in regions mainly related to the limbic system (Brattico et al., 2011) when listening to music. These findings may relate to the relevance that movement has when processing music (especially for participants with musical expertise) as well as for the emotional content that a naturalistic listening to music can easily evoke. Interestingly, the network of brain areas that we de-

tected was different, since it did not relate much with default mode or motor networks, but comprised areas mostly related to auditory, memory, attentional and evaluative processes. Further, this network evolved even more clearly when contrasting it with the functional connectivity that we observed in resting state. Thus, our findings suggest that the synchronization between these areas reflect the brain functional connectivity underlying the active encoding of the sounds. Future studies should replicate our results and further expand the knowledge on sound encoding by directly comparing active encoding of sound information not only versus resting state (as done in the current study) but also versus passive listening to music.

4.2. Dynamic functional connectivity of sound encoding

A further relevant result that we achieved comes from the study of functional connectivity rapid dynamics that allowed us to identify two main time-windows of sound encoding brain processes. This procedure returned two similar networks of brain regions that were however differentiated by a diverse centrality of few areas. Primary auditory cortex was central mainly in the early time window whereas its connectivity was less marked in the later time interval which was more distant from the sound onset. This result relates to previous literature highlighting its role for the first sensorial processing of upcoming auditory stimuli (Elhilali et al., 2004). Interestingly, our results confirmed the relevance of auditory cortex not only in terms of activation, but also with regards to synchronization with the other brain areas in the first rapid processing of the sounds. More surprisingly, the same result regarded frontal operculum which could be as well important for the brain network associated to the first processes of sound encoding. This result was less expected since frontal operculum has usually been associated to later processes in language studies (Rolston et al., 2015). Interestingly, our evidence may suggest that when encoding sounds the frontal operculum could be needed already in the initial processing of the sounds. On the contrary, right insula was predominantly central within the second time-window. In accordance with what we described above, insula may relate to salience of the stimuli and in this case could have played an important role for achieving a more fine-grained encoding and categorization of the upcoming sounds. Once again, this interpretation would be coherent with previous studies highlighting insula's role in salience appraisal of stimuli (Cauda et al., 2011; Uddin, 2015). In conclusion, by investigating the brain connectivity dynamics within the first 220 ms after sounds onset, we described the neural mechanisms in terms of functional connectivity that are presumably responsible for making temporal information available to human awareness, as suggested by previous investigations that explored the dynamics of human conscious processing (Dehaene et al., 2011). Finally, although we employed solid signal processing methods, we wish to highlight that dynamic connectivity measures are intriguing, but possibly more prone to artifacts than more established techniques such as static functional connectivity analysis that we have previously discussed. Thus, while our results on static functional connectivity provides us with a solid and exciting new perspective on the brain mechanisms underlying sound encoding, the dynamic functional connectivity outcomes should be managed with a balanced mix of interest and caution.

4.3. Sound encoding, working memory abilities and musical expertise

An additional achievement of this study is represented by the modulation of sound encoding brain networks based on participants' WM skills and musical expertise. First, general WM was associated to centrality of frontal operculum. This result can be seen considering the several studies that highlighted the fundamental role of frontal and prefrontal cortices for WM tasks (Christophel et al., 2017; Constantinidis and Klingberg, 2016; Fuster, 2015). Indeed, in our study higher versus lower WM skills participants had a stronger centrality of the most frontal brain region that formed the significant sound encoding brain

network. Remarkably, when considering auditory WM, we did not observe any difference related to frontal operculum. However, we detected a significant centrality of bilateral insula and right temporal middle cortex. Considering that these regions have been shown important for complex processing of auditory stimuli (Kumar et al., 2016; Remedios et al., 2009), we may speculate that their involvement in an auditory memory task may offer an additional help to encode sounds. While this may happen for best auditory WM scorers, worse participants may rely only on more primitive auditory cortices. Although these results provide new interesting perspectives, we highlight that the current study focused on the fast-scale whole-brain connectivity underlying encoding of single sounds and not on WM tasks. Thus, our results suggest that general cognitive abilities (such as WM) can produce a small, but significant modulation of the brain encoding of single sounds, as suggested by other studies linking WM skills and brain processing of sounds (Bonetti et al., 2018). However, they do not tell us much on the brain processes underlying WM tasks such as information storage during goal-directed behaviour (Sreenivasan et al., 2014) involving sound information. Future studies are therefore called for to investigate such relevant topic and integrate the existing knowledge.

In relation to musical expertise, the higher centrality of left supplementary motor area and right insula for musicians compared to non-musicians is worthy to be mentioned. Since motor learning is a key feature of musical practice, we claim that musicians may also recruit motor areas when encoding sounds. This result could be seen in light of several studies that showed the role of motor brain areas in musicians during music listening (Alluri et al., 2017; Burunat et al., 2015; Bangert et al., 2006). Similarly, also insula has been shown more active when musicians compared to non-musicians listened to an early rehearsed and familiar musical piece (Mutschler et al., 2007). Finally, the analysis on the familiarity with the Bach's prelude showed that participants who had never heard the prelude before the experiment presented a significantly stronger centrality of the right medial orbito-frontal cortex during the encoding of sounds. Although this last result may lead to different interpretations, we might argue that to encode the sounds participants with no prior knowledge of the musical piece needed a higher synchronization of a frontal area such as the medial orbito-frontal cortex which has been shown implicated in several cognitive tasks by previous research (Gourley et al., 2016; Elliott et al., 2000) and thus may play a privileged role also in the encoding of sounds.

5. Conclusions

In conclusion, this study revealed the rapid spatiotemporal dynamics of brain activity and connectivity underlying encoding of sounds forming a full, structured musical piece, as compared to resting state, in a large sample of almost seventy participants. Remarkably, the integration between activity and connectivity provided us with a complete picture of the brain networks involved in this complex cognitive process, networks that the sole brain activity missed to reveal. Indeed, while the brain activity was largely reconstructed on the primary auditory cortex and insula, the large-scale whole-brain connectivity analysis revealed a wide set of higher-order brain areas that were highly synchronized between each other and central within the whole-brain network. Remarkably, during the encoding process such areas were not strongly active but presented very similar time series which may indicate the dense communication happening within this subnetwork of areas. Moreover, these brain regions are not classically implicated in auditory processing, suggesting that encoding of sounds requires the integration of several different brain structures that may collaborate to extract and store the information carried by the sounds. Further, our investigation within the first 220 ms after sounds onset allowed us to detect the brain mechanisms that may be partly responsible for making temporal information available to human awareness (Dehaene et al., 2011). On top of this, our study highlighted the potential role of DFC and phase synchronization analyses to unravel the rapid transition of the brain connectivity pat-

terns from primary auditory cortex to higher order association areas. Indeed, our results suggested that primary auditory cortex centrality would be implicated in the first processing of the present sound, while the integration between insula and superior temporal pole with the rest of the brain may play a crucial role for a more fine-grained elaboration of the sounds to be encoded. Finally, the DFC approach allowed us to complement our results with additional analyses that revealed a subtle but relevant modulation of the brain networks underlying sound encoding in participants characterized by different levels of WM abilities and musical expertise.

Taken together, these results advanced our knowledge of the brain connectivity functional networks and rapid dynamics occurring when sound information is encoded. Further, our findings may even provide a further understanding of the general mechanisms underlying encoding of the single objects (i.e. sounds) of a sequence (i.e. musical piece) in the human brain. As a natural development of our work, future research is called for to broaden our results by looking into the brain connectivity emerging during encoding not only of single sounds, but of longer sequences.

Credit author statement

L. Bonetti: Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization. **E. Brattico:** Conceptualization, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition. **F. Carlomagno:** Methodology, Software, Formal analysis, Visualization. **G. Donati:** Data curation, Investigation, Writing – original draft, Visualization. **J. Cabral:** Methodology, Software, Formal analysis. **N.T. Haumann:** Methodology, Software, Formal analysis, Data curation. **G. Deco:** Investigation, Supervision, Software. **P. Vuust:** Validation, Resources, Writing – review & editing, Supervision, Funding acquisition, Project administration. **M.L. Kringelbach:** Conceptualization, Methodology, Validation, Resources, Writing – review & editing, Supervision, Funding acquisition, Project administration.

Author contributions

LB, EB, MLK and PV conceived the hypotheses and designed the study. LB, FC, GDO JC, NTH, MLK performed pre-processing and statistical analysis. GDE, EB, MLK, GDO and PV provided essential help to interpret and frame the results within the neuroscientific literature. LB wrote the first draft of the manuscript and, together with FC and MLK, prepared the figures. All the authors contributed to and approved the final version of the manuscript.

Declaration of Competing Interests

The authors declare no competing interests.

Data availability

The codes are available on GitHub (<https://github.com/leonardob92/LBPD-1.0.git>). The anonymized neuroimaging data from the experiment will be made available upon reasonable request. Regarding the data, we will be able to share it when it is completely anonymized and cannot lead in any way to the original participants identity, according to Danish regulations. Otherwise, a data sharing agreement must be made. Please, contact the author Leonardo Bonetti for further information (leonardo.bonetti@clin.au.dk).

Acknowledgements

We thank Riccardo Proietti, Giulio Carraturo, Mick Holt, Holger Friis for their assistance in the neuroscientific experiment. We also thank the

psychologist Tina Birgitte Wisbech Carstensen for her help with the administration of psychological tests and questionnaires.

The Center for Music in the Brain (MIB) is funded by the [Danish National Research Foundation](#) (project number DNRF117).

LB is supported by [Carlsberg Foundation](#) (project number CF20-0239), Center for Music in the Brain and Linacre College of the University of Oxford.

MLK is supported by the [ERC Consolidator Grant: CAREGIVING](#) (n. 615539), Center for Music in the Brain, and Center for Eudaimonia and Human Flourishing funded by the Pettit and Carlsberg Foundations.

GD is supported by the Spanish Research Project PSI2016-75688-P (AEI/FEDER, EU), by the European Union's Horizon 2020 Research and Innovation Programme under grant agreements n. 720270 (HBP SGA1) and n. 785907 (HBP SGA2), and by the Catalan AGAUR Programme 2017 SGR 1545.

JC is supported by Portuguese Foundation for Science and Technology CEECIND/03325/2017, Portugal.

Additionally, we thank the Italian section of *Mensa: The International High IQ Society* for the economic support provided to the author Francesco Carlomagno and the University of Bologna for the economic support provided to the author Giulia Donati and the student assistants Riccardo Proietti and Giulio Carraturo.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.inffus.2021.12.002](https://doi.org/10.1016/j.inffus.2021.12.002).

References

- Albouy, P., Weiss, A., Baillet, S., Zatorre, R.J., 2017. Selective Entrainment of Theta Oscillations in the Dorsal Stream Causally Enhances Auditory Working Memory Performance. *Neuron* doi:[10.1016/j.neuron.2017.03.015](https://doi.org/10.1016/j.neuron.2017.03.015).
- Alluri, V., Toivianen, P., Jääskeläinen, I.P., Glerean, E., Sams, M., Brattico, E., 2012. Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *Neuroimage* 59 (4), 3677–3689.
- Alluri, V., et al., 2017. Connectivity patterns during music listening: evidence for action-based processing in musicians. *Hum. Brain Mapp.* 38, 2955–2970.
- Ashby, F.G., Turner, B.O., Horvitz, J.C., 2010. Cortical and basal ganglia contributions to habit learning and automaticity. *Trends Cogn. Sci.* 14, 208–215.
- Baker, D.J., Ventura, J., Calamia, M., Shanahan, D., Elliott, E.M., 2018. Examining musical sophistication: a replication and theoretical commentary on the Goldsmiths Musical Sophistication Index. *Musicae Scientiae* 24, 411–429.
- Bangert, M., et al., 2006. Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *Neuroimage* 30, 917–926.
- Behroozmand, R., et al., 2015. Sensory-motor networks involved in speech production and motor control: an fMRI study. *Neuroimage* 109, 418–428.
- Bonetti, L., Haumann, N.T., Brattico, E., Kliuchko, M., Vuust, P., Särkämö, T., Näätänen, R., 2018. Auditory sensory memory and working memory skills: association between frontal MMN and performance scores. *Brain research* 1700, 86–98.
- Bonetti, L., Brattico, E., Carlomagno, F., Cabral, J., Stevner, A., Deco, G., Whybrow, P.C., Pearce, M., Pantazis, Y., Vuust, P., Kringelbach, M.L., 2020. Spatiotemporal whole-brain dynamics of auditory patterns recognition. *bioRxiv* 2020.06.23.165191; doi:[10.1101/2020.06.23.165191](https://doi.org/10.1101/2020.06.23.165191).
- Brattico, E., Alluri, V., Bogert, B., Jacobsen, T., Vartiainen, N., Nieminen, S.K., Tervaniemi, M., 2011. A functional MRI study of happy and sad emotions in music with and without lyrics. *Frontiers in psychology* 2, 308 Chicago.
- Brookes, M.J., et al., 2007. Beamformer reconstruction of correlated sources using a modified source model. *Neuroimage* 34, 1454–1465.
- Brookes, M.J., et al., 2016. A multi-layer network approach to MEG connectivity analysis. *Neuroimage* 132, 425–438.
- Bruno, A.C., Romani, G.L., 1989. Neuromagnetic localization performed by using planar gradiometer configurations. *J. Appl. Phys.* 65, 2098.
- Burunat, I., et al., 2015. Action in perception: prominent visuo-motor functional symmetry in musicians during music listening. *PLoS One* 10, e0138238.
- Cabral, J., et al., 2014. Exploring mechanisms of spontaneous functional connectivity in MEG: How delayed network interactions lead to structured amplitude envelopes of band-pass filtered oscillations. *Neuroimage* 90, 423–435.
- Cauda, F., et al., 2011. Functional connectivity of the insula in the resting brain. *Neuroimage* 55, 8–23.
- Celma-Miralles, A., Toro, J.M., 2019. Ternary meter from spatial sounds: Differences in neural entrainment between musicians and non-musicians. *Brain and cognition* 136, 103594.
- Christophel, T.B., Klink, P.C., Spitzer, B., Roelfsema, P.R., Haynes, J.D., 2017. The Distributed Nature of Working Memory. *Trends Cogn. Sci.* 21, 111–124.
- Colclough, G.L., Brookes, M.J., Smith, S.M., Woolrich, M.W., 2015. A symmetric multi-variate leakage correction for MEG connectomes. *Neuroimage* 117, 439–448.
- Conley, E.M., Michalewski, H.J., Starr, A., 1999. The N100 auditory cortical evoked potential indexes scanning of auditory short-term memory. *Clin. Neurophysiol.* 110, 2086–2093.
- Constantinidis, C., Klingberg, T., 2016. The neuroscience of working memory capacity and training. *Nat. Rev. Neurosci.* 17, 438–449.
- Daubechies, I., 1992. *Ten Lectures on Wavelets*. SIAM Press, Philadelphia PA.
- Dehaene, S., Changeux, J.P., Naccache, L., 2011. The global neuronal workspace model of conscious access: From neuronal architectures to clinical applications. In: Dehaene, S., Christen Y. (Eds.), *Characterizing Consciousness: From Cognition to the Clinic?*. Research and Perspectives in Neurosciences. Springer, Berlin, Heidelberg.
- Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., Pallier, C., 2015. The Neural Representation of Sequences: From Transition Probabilities to Algebraic Patterns and Linguistic Trees. *Neuron* 88, 2–19.
- Dobs, K., Isik, L., Pantazis, D., Kanwisher, N., 2019. How face perception unfolds over time. *Nat. Commun.* 10, 1258.
- Dumont, R., Willis, J.O., Veizel, K., Zibulsky, J., 2014. *Wechsler Adult Intelligence Scale—Fourth Edition*. In *Encyclopedia of Special Education*. Wiley.
- Elhilali, M., Fritz, J.B., Klein, D.J., Simon, J.Z., Shamma, S.A., 2004. Dynamics of precise spike timing in primary auditory cortex. *J. Neurosci.* 24, 1159–1172.
- Elliott, R., Dolan, R.J., Frith, C.D., 2000. Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies. *Cerebral cortex* 10 (3), 308–317.
- Fazio, P., et al., 2009. Encoding of human action in Broca's area. *Brain* 132, 1980–1988.
- Frank, L.M., Brown, E.N., Wilson, M., 2000. Trajectory encoding in the hippocampus and entorhinal cortex. *Neuron* 27, 169–178.
- Fuster, J.M., 2015. *The Prefrontal Cortex*. Academic Press/Elsevier.
- Gaab, N., Gaser, C., Zaehle, T., Jancke, L., Schlaug, G., 2003. Functional anatomy of pitch memory - An fMRI study with sparse temporal sampling. *NeuroImage* doi:[10.1016/S1053-8119\(03\)00224-6](https://doi.org/10.1016/S1053-8119(03)00224-6).
- Gourley, S.L., Zimmermann, K.S., Allen, A.G., Taylor, J.R., 2016. The medial orbitofrontal cortex regulates sensitivity to outcome value. *Journal of Neuroscience* 36 (16), 4600–4613.
- Haumann, N.T., Lumaca, M., Kliuchko, M., Santacruz, J.L., Vuust, P., Brattico, E., 2021. Extracting human cortical responses to sound onsets and acoustic feature changes in real music, and their relation to event rate. *Brain Research* 1756.
- Henry, M.J., Herrmann, B., Künke, D., Obleser, J., 2017. Aging affects the balance of neural entrainment and top-down neural modulation in the listening brain. *Nature communications* 8 (1), 1–11.
- Herry, C., Johansen, J.P., 2014. Encoding of fear learning and memory in distributed neuronal circuits. *Nat. Neurosci.* 17, 1644–1654.
- Hickey, C., Peelen, M.V., 2015. Neural mechanisms of incentive salience in naturalistic human vision. *Neuron* 85, 512–518.
- Hillebrand, A., Barnes, G.R., 2005. Beamformer Analysis of MEG Data. *Int. Rev. Neurobiol.* 68, 149–171.
- Hindriks, R., et al., 2015. Role of white-matter pathways in coordinating alpha oscillations in resting visual cortex. *Neuroimage* 106, 328–339.
- Huang, M.X., Mosher, J.C., Leahy, R.M., 1999. A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG. *Phys. Med. Biol.* 44, 423–440.
- Hunt, L.T., et al., 2012. Mechanisms underlying cortical activity during value-guided choice. *Nat. Neurosci.* 15, 470–476.
- Husain, F.T., McKinney, C.M., Horwitz, B., 2006. Frontal cortex functional connectivity changes during sound categorization. *Neuroreport* 17, 617–621.
- Indefrey, P., et al., 2001. A neural correlate of syntactic encoding during speech production. *Proc. Natl. Acad. Sci. USA* 98, 5933–5936.
- Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17, 4302–4311.
- Klein, S.B., Robertson, T.E., Delton, A.W., 2010. Facing the future: memory as an evolved system for planning future acts. *Mem. Cognit.* 38, 13–22.
- Koelsch, S., Siebel, W.A., 2005. Towards a neural basis of music perception. *Trends in cognitive sciences* 9 (12), 578–584.
- Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., Friederici, A.D., 2004. Music, language and meaning: brain signatures of semantic processing. *Nature neuroscience* 7 (3), 302–307 Chicago.
- Kroese, D.P., Taimre, T., Botev, Z.I., 2011. *Handbook of Monte Carlo Methods*. Wiley.
- Kumar, S., Joseph, S., Gander, P.E., Barascud, N., Halpern, A.R., Griffiths, T.D., 2016. A brain system for auditory working memory. *Journal of Neuroscience* doi:[10.1523/JNEUROSCI.4341-14.2016](https://doi.org/10.1523/JNEUROSCI.4341-14.2016).
- Kumar, S., Joseph, S., Gander, P.E., Barascud, N., Halpern, A.R., Griffiths, T.D., 2016. A brain system for auditory working memory. *Journal of Neuroscience* doi:[10.1523/JNEUROSCI.4341-14.2016](https://doi.org/10.1523/JNEUROSCI.4341-14.2016).
- Langers, D.R.M., Melcher, J.R., 2011. Hearing without listening: functional connectivity reveals the engagement of multiple nonauditory networks during basic sound processing. *Brain Connect* 1, 233–244.
- Layer, E., Tomczyk, K., 2015. *Hilbert transform. Studies in Systems, Decision and Control*. Springer, Cham.
- Leino, S., Brattico, E., Tervaniemi, M., Vuust, P., 2007. Representation of harmony rules in the human brain: Further evidence from event-related potentials. *Brain research* 1142, 169–177.
- Lijffijt, M., et al., 2009. P50, N100, and P200 sensory gating: relationships with behavioral inhibition, attention, and working memory. *Psychophysiology* 46, 1059–1068.
- Müllensiefen, D., Gingras, B., Musil, J., Stewart, L., 2014. Measuring the facets of musicality: The Goldsmiths Musical Sophistication Index (Gold-MSI). *Pers. Individ. Dif.* 60, S35.
- Maess, B., Koelsch, S., Gunter, T.C., Friederici, A.D., 2001. Musical syntax is processed in Broca's area: an MEG study. *Nature neuroscience* 4 (5), 540–545.

- Mantini, D., et al., 2011. A signal-processing pipeline for magnetoencephalography resting-state networks. *Brain Connect* 1, 49–59.
- Mutschler, I., et al., 2007. A rapid sound-action association effect in human insular cortex. *PLoS One* 28, e259.
- Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24, 375–425.
- Näätänen, R., Paavilainen, P., Rinne, T., Alho, K., 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590.
- Novembre, G., Iannetti, G.D., 2018. Tagging the musical beat: Neural entrainment or event-related potentials? *Proceedings of the National Academy of Sciences* 115 (47), E11002–E11003.
- Obleser, J., Kayser, C., 2019. Neural entrainment and attentional selection in the listening brain. *Trends in cognitive sciences* 23 (11), 913–926 Chicago.
- Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., FieldTrip, 2011. Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.*, 156869 2011.
- Packard, M.G., Knowlton, B.J., 2002. Learning and Memory Functions of the Basal Ganglia. *Annu. Rev. Neurosci.* 25, 563–593.
- Penny, W., Friston, K., Ashburner, J., Kiebel, S., Nichols, T., 2007. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press/Elsevier.
- Preston, A.R., Eichenbaum, H., 2013. Interplay of hippocampus and prefrontal cortex in memory. *Curr. Biol.* 23, 764–773.
- Remedios, R., Logothetis, N.K., Kayser, C., 2009. An auditory region in the primate insular cortex responding preferentially to vocal communication sounds. *J. Neurosci.* 29, 1034–1045.
- Richiardi, J., Eryilmaz, H., Schwartz, S., Vuilleumier, P., Van De Ville, D., 2011. Decoding brain states from fMRI connectivity graphs. *Neuroimage* 56, 616–626.
- Rolston, J.D., Englot, D.J., Benet, A., Li, J., Cha, S., Berger, M.S., 2015. Frontal operculum gliomas: language outcome following resection. *Journal of neurosurgery* 122 (4), 725–734 Chicago.
- Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069.
- Schwarzlose, R.F., Baker, C.I., Kanwisher, N., 2005. Separate face and body selectivity on the fusiform gyrus. *J. Neurosci.* 25, 11055–11059.
- Sikka, R., Cuddy, L.L., Johnsrude, I.S., Vanstone, A.D., 2015. An fMRI comparison of neural activity associated with recognition of familiar melodies in younger and older adults. *Frontiers in Neuroscience* doi:10.3389/fnins.2015.00356.
- Sloboda, J.A., Davidson, J.W., Howe, M.J., Moore, D.G., 1996. The role of practice in the development of performing musicians. *British journal of psychology* 87 (2), 287–309.
- Sreenivasan, K.K., Vytlačil, J., D'Esposito, M., 2014. Distributed and dynamic storage of working memory stimulus information in extrastriate cortex. *Journal of cognitive neuroscience* 26 (5), 1141–1153.
- Stern, C.E., et al., 1996. The hippocampal formation participates in novel picture encoding: evidence from functional magnetic resonance imaging. *Proc. Natl. Acad. Sci. USA* 93, 8660–8665.
- Takeuchi, T., Duzkiewicz, A.J., Morris, R.G.M., 2014. The synaptic plasticity and memory hypothesis: encoding, storage and persistence. *Philos. Trans. R. Soc. B. Biol. Sci.* 369, 20130288.
- Taruffi, L., Pehrs, C., Skouras, S., Koelsch, S., 2017. Effects of sad and happy music on mind-wandering and the default mode network. *Scientific reports* 7 (1), 1–10.
- Taulu, S., Simola, J., 2006. Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. *Phys. Med. Biol.* 51, 1759–1768.
- Taylor, M.J., Heaton, R.K., 2001. Sensitivity and specificity of WAIS-III/WMS-III demographically corrected factor scores in neuropsychological assessment. *Journal of the International Neuropsychological Society* 7 (7), 867–874.
- Tervaniemi, M., Hugdahl, K., 2003. Lateralization of auditory-cortex functions. *Brain Res. Rev.* 43, 231–246.
- Toivainen, P., Alluri, V., Brattico, E., Wallentin, M., Vuust, P., 2014. Capturing the musical brain with Lasso: Dynamic decoding of musical features from fMRI data. *Neuroimage* 88, 170–180.
- Toivainen, P., Burunat, I., Brattico, E., Vuust, P., Alluri, V., 2020. The chronnectome of musical beat. *NeuroImage* 216.
- Tzourio-Mazoyer, N., et al., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289.
- Uddin, L.Q., 2015. Salience processing and insular cortical function and dysfunction. *Nat. Rev. Neurosci.* 16, 55–61.
- Vago, D.R., Wallenstein, G.V., Morris, L.S., 2014. Hippocampus. In: Aminoff, M.J., Daroff, R.B. (Eds.), *Encyclopedia of the Neurological Sciences*. Academic Press/Elsevier, Waltham, MA.
- Villarreal, E.A.G., Brattico, E., Leino, S., Østergaard, L., Vuust, P., 2011. Distinct neural responses to chord violations: a multiple source analysis study. *Brain Research* 1389, 103–114; Koelsch, S., Jentschke, S., 2010. Differences in electric brain responses to melodies and chords. *Journal of cognitive neuroscience* 22 (10), 2251–2262.
- Vuust, P., Brattico, E., Seppänen, M., Näätänen, R., Tervaniemi, M., 2012. The sound of music: differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm. *Neuropsychologia* 50 (7), 1432–1443.
- Wallentin, M., Nielsen, A.H., Friis-Olivarius, M., Vuust, C., Vuust, P., 2010. The Musical Ear Test, a new reliable test for measuring musical competence. *Learn. Individ. Differ.* 20, 188–196.
- Warrier, C., et al., 2009. Relating structure to function: Heschl's gyrus and acoustic processing. *J. Neurosci.* 29, 61–69.
- Wechsler, D., 1997. *Wechsler adult intelligence scale, Third Edition (WAIS-III)* San Antonio.
- Winer, J.A., Schreiner, C.E., 2011. *The Auditory Cortex*. Springer, New York.
- Woolrich, M.W., et al., 2009. Bayesian analysis of neuroimaging data in FSL. *Neuroimage* 45, S173–S186.
- Yoo, K., Rosenberg, M.D., Hsu, W.T., Zhang, S., Li, C.S.R., Scheinost, D., Chun, M.M., 2018. Connectome-based predictive modeling of attention: Comparing different functional connectivity features and prediction methods across datasets. *Neuroimage* 167, 11–22.
- Zatorre, R.J., Evans, A.C., Meyer, E., 1994. Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience* doi:10.1523/jneurosci.14-04-01908.1994.
- Zatorre, R.J., 2003. Sound analysis in auditory cortex. *Trends Neurosci* 26, 229–230.