scientific reports

Check for updates

OPEN

Operatic voices engage the default mode network in professional opera singers

Adél Bihari[1,2✉], Ádám Nárai[1,3], Boris Kleber[4], Judit Zsuga[5], Petra Hermann[1,6] & Zoltán Vidnyánszky[1,6✉]

Extensive research with musicians has shown that instrumental musical training can have a profound impact on how acoustic features are processed in the brain. However, less is known about the influence of singing training on neural activity during voice perception, particularly in response to salient acoustic features, such as the vocal vibrato in operatic singing. To address this gap, the present study employed functional magnetic resonance imaging (fMRI) to measure brain responses in trained opera singers and musically untrained controls listening to recordings of opera singers performing in two distinct styles: a full operatic voice with vibrato, and a straight voice without vibrato. Results indicated that for opera singers, perception of operatic voice led to differential fMRI activations in bilateral auditory cortical regions and the default mode network. In contrast, musically untrained controls exhibited differences only in bilateral auditory cortex. These results suggest that operatic singing training triggers experience-dependent neural changes in the brain that activate self-referential networks, possibly through embodiment of acoustic features associated with one's own singing style.

Keywords Operatic voice, Auditory processing, Musical expertise, Default mode network, fMRI

Becoming a professional musician demands extensive and persistent dedication to practice and learning over an extended period of time, often starting in early childhood[1,2]. This pursuit entails not only developing sensorimotor proficiency but also encompasses a profound transformation of sensory perceptual modalities required for performance monitoring and motor adjustments[3–5]. Simultaneously, it engages cognitive and affective processes, highlighting its multifaceted nature[6,7]. This comprehensive training has been associated with several anatomically plausible adaptations in brain regions involved in sensory processing, sensorimotor integration, emotion, and higher-level cognitive functions, causing a global impact on the structural and functional organization of the brain[8–11]. Robust changes in large-scale structural and functional networks across the brain, as recently demonstrated[12], substantiate and extend these observations, indicating that musical expertise shapes the connectivity across sensory, motor, and cognitive regions, potentially enhancing the information transfer among these neural systems even in task-free conditions[13]. These results highlight why musical training has emerged as an excellent framework for investigating the brain's ability to adapt and reorganize in response to experiences and environmental demands[14], with broader implications for understanding brain development, learning, and adaptation that extend beyond the realm of music itself[15].

While we continue to expand our understanding of how musical training may alter the brain, we remain relatively naïve about the effects of training an instrument that's so ubiquitous, we often take it for granted: our voice. Drawing inferences regarding the potential effect of training one's voice carries significant implications, since much of the research into training-related effects has focused on how long-term exposure to specific acoustic stimuli can alter auditory perception, considering that observable differences are most pronounced when the auditory input is both behaviourally relevant and actively trained[8,16–18]. This is consistent with the "action-perception" hypothesis, suggesting that perception is not a passive process but is actively shaped by our past

[1]Brain Imaging Centre, HUN-REN Research Centre for Natural Sciences, Budapest, Hungary. [2]Department of Cognitive Science, Faculty of Natural Sciences, Budapest University of Technology and Economics, Budapest, Hungary. [3]Doctoral School of Biology and Sportbiology, Institute of Biology, Faculty of Sciences, University of Pécs, Pécs, Hungary. [4]Center for Music in the Brain, Department of Clinical Medicine, Aarhus University and The Royal Academy of Music, Aarhus/Aalborg, Denmark. [5]Department of Psychiatry, Faculty of Medicine, University of Debrecen, Debrecen, Hungary. [6]These authors jointly supervised this work: Petra Hermann and Zoltán Vidnyánszky. ✉email: bihari.adel@ttk.hu; vidnyanszky.zoltan@ttk.hu

experiences and actions[19]. Further supporting this perspective, the Theory of Event Coding (TEC) and Prinz's model on a common coding approach postulate that sensory inputs and motor outputs are not separately processed in the brain but are encoded as "events" in a common representational medium, enhancing the efficiency of interaction with the environment[20,21]. The Predictive Coding of Music (PCM) model[22,23] further complements this view by framing music perception, action, emotion, and learning as recursive Bayesian processes, where the brain continuously minimizes prediction errors by comparing sensory input with top-down predictions. This model illustrates how musical training, such as that received by opera singers, refines these predictive mechanisms, leading to improved auditory and motor integration and corresponding cognitive skills through long-term experience-dependent plasticity. A compelling demonstration of the intimate relationship between sound production and perception is that cortical responses to musical stimuli are uniquely modulated by the specific timbre of one's own trained instrument[24-28]. Moreover, it has been shown that the primary laryngeal motor cortex involves separable motor and auditory representations of vocal pitch, with discrete neural populations handling specific pitch-related tasks during speech or singing[29]. When it comes to classical or operatic singing, vibrato is a particularly salient feature that provides a unique opportunity to explore how singers' net sensorimotor experience with this singing style may shape auditory perception and the ensuing brain responses.

In Western-style operatic singing, vibrato is an essential aspect of classical vocal training that must be mastered[30]. Characterized by a quasi-sinusoidal, low-frequency modulation of the fundamental frequency or pitch, vibrato affects all spectral timbral components[31]. Typically, it is quantified in terms of the rate and extent of this modulation, whereas the pitch we perceive corresponds to the mean frequency of this modulation[32]. It serves a crucial artistic function: it infuses the voice with richness and warmth, enhancing both resonance and emotional expression[33-35]. Given its importance in classical singing training, its psychoacoustic properties, and its artistic impact, vibrato presents an intriguing lens through which to investigate the neural mechanisms underlying experience-dependent action-perception processes.

Our present understanding of the brain function and structural differences associated with action-perception processes in singing training largely stems from studies that have utilized pitch matching and singing production paradigms to investigate sensorimotor integration. These studies collectively suggest an experience-dependent role of embodied motor and acoustic feedback control mechanisms, strongly engaging the auditory, somatosensory, and inferior parietal cortices, as well as frontal brain regions associated with performance monitoring and the insula[36-41]. With regards to the sensory modalities involved, the brain regions associated with embodied motor control not only support tactile information transmission relevant to singing but also have important links with corresponding auditory processes[10,35,42-44]. However, despite a steadily increasing number of studies investigating the neural mechanisms underlying human speech[45], understanding how singing training might influence perceptual processes remains an open question[46,47].

To address this gap, the current study aims to elucidate the specific neural mechanisms by which experience and training in operatic singing interact with the neural substrates underlying its perception, leveraging the presence or absence of vibrato as a defining element. Blood-oxygen-level-dependent (BOLD) responses were measured to ecologically valid recordings performed by trained opera singers, highlighting two contrasting vocal styles: a full operatic voice with vibrato, and a straighter voice without vibrato, which we refer to as the "natural" style. These responses were investigated in trained opera singers and subsequently compared to musically untrained controls. We hypothesize that listening to operatic versus natural singing styles will engage brain regions linked to auditory perception, sensorimotor integration, and higher-level cognitive functions more robustly in opera singers. By including musically untrained controls, we aim to determine whether comprehensive operatic singing training will result in heightened neural activity in these brain regions among singers relative to controls, following our hypothesis that specialized singing expertise has a profound impact on singing-related auditory processing in the brain.

Methods

Participants
In this study, 32 professionally trained classical opera singers and 35 controls without prior musical or vocal training were enrolled. None of the participants had a reported history of neurological or psychiatric disease. Due to excessive head motion and anatomical abnormalities, two controls were excluded from further analysis, resulting in a final sample size of N = 32 opera singers (age range: 20–52 years; age mean ± standard deviation (SD): 39.0 ± 7.3 years; 18 female, 5 left-handed) and N = 33 controls (age range: 20–49 years; age mean ± SD: 29.0 ± 8.2 years; 22 female, 2 left-handed). Professional singers' average age at the beginning of formal private singing training was 17.9 years (SD = 3.9; range = 13–29 years), and their average professional singing experience was 20.8 years (SD = 6.7, range = 6–31 years), as assessed by the translated form of the Montreal Music History Questionnaire (MMHQ)[48]. All participants provided written, informed consent before participation and were financially compensated for their time. The research protocol was designed and conducted in accordance with the Hungarian regulations and laws, and with the Declaration of Helsinki, and was approved by the National Institute of Pharmacy and Nutrition, Hungary (file number: OGYÉI/70,184/2017).

Data acquisition initially focused on the opera group. After the analyses revealed significant effects related to stimuli style (natural vs. operatic), we proceeded to recruit control participants to determine if these effects were specific to opera singers. The analysis pipeline described takes the sequential nature of data acquisition into account.

Stimuli generation
Following the methodology described by Lévêque et al.[49,50], we initially generated a diverse pool of 120 pseudorandomized 5-tone isochronous melodies for use across multiple studies. These melodies were derived from all 7

modal scales and were designed with minimal tone repetition (maximum one repetition per melody). The melodies were transposed into three distinct vocal ranges to accommodate four different voice types: G2–C♯3 for men only (male-low), B3–E♯4 for both sexes (male-high and female-low), and C5–F♯5 for women only (female-high).

Melodies were recorded by two professionally trained and actively performing opera singers, one male (baritone) and one female (mezzo-soprano). Each singer performed the melodies using the vowel "o" in two distinct singing styles: "naturally" (using straight tones without vibrato) and "operatically" (using vibrato). The recordings were captured at a sampling rate of 48 kHz in a sound-proof recording studio. A metronome set at 80 beats per minute and the melody played on a piano delivered through headphones aided the singers in maintaining consistent tempo and pitch. Each recorded stimulus had an average ± SD duration of 4.03 ± 0.12 s.

Stimulus selection and preparation

For this specific experiment, we refined our stimulus selection to 48 melodies, each recorded with all stimulus type combinations (2 singing styles × 4 voice types), resulting in a total of 384 unique stimuli. During the selection process, we focused on minimizing length differences (maximum difference of 353 ms) for each melody across all voice types and singing styles. To ensure uniformity across conditions, melodies were trimmed at both ends to match the length of the shortest one. Onset and offset ramps were applied using a half-Hanning window, corresponding in length to the amount trimmed. Finally, each melody was normalized to ensure a consistent root mean square (RMS) value across all stimuli.

To optimize the use of our specialized participant pool, the curated set of 48 melodies was designed to allow for the investigation of multiple research questions for different studies. This approach aimed to enhance the efficiency of data collection to explore various dimensions of vocal perception, such as the impact of singing styles, vocal ranges, and the influence of the participants' and singers' sex on participants' auditory processing. For the current analysis, however, we focused specifically on the effect of singing style (operatic vs. natural) within the vocal range of B3–E♯4, reproducible by both male and female participants. This range was selected to avoid potential confounding effects related to different vocal capabilities and to maximize the number of trials available for robust statistical analysis. Melodies outside this range, while part of the original experimental design, were reserved for separate analyses addressing different research questions.

To align the vocal range of the singing stimuli with that of the participants, we assessed each participant's vocal capabilities. Male participants demonstrated an average vocal range of 28.5 semitones (range: 26–39), extending from a lower limit of C2–A2 to an upper limit of G4–C5. Female participants exhibited an average range of 35 semitones (range: 27–43), spanning from D3–A3 to B♭5–C7. Consequently, the selected tonal range for the singing stimuli (B3–E♯4) fell well within the vocal capabilities of all participants (A3–G4). This careful matching ensured that the fMRI data analysed reflected the neural responses to stimuli that participants could potentially produce themselves, enhancing the relevance and interpretability of our findings related to the experience-dependent neural activity during operatic voice perception.

Experimental procedure

Participants were positioned comfortably in the MRI scanner, with their heads secured to minimize movement. They were instructed to remain still and focus on the auditory stimuli throughout the session. The fMRI experiment was conducted over four functional runs using a block design, each approximately 12 min long. During stimuli presentation, participants listened to a series of 18.25 s long melody blocks, containing four melodies of the same singing style (either operatic or natural) and voice type (female-low, female-high, male-low, or male-high). The melodies were played consecutively for approximately 4000 ms each, with an interstimulus interval (ISI) of 750 ms between melodies. At the beginning of each run, there was a baseline period of 30 s. Following the baseline, the melody blocks alternated with rest periods of varying durations (8, 11, or 14 s) to allow for BOLD response decay and to reduce predictability of stimulus presentation. Each run included three blocks for each stimulus type combination (style × voice type), resulting in 24 melody blocks and 25 baseline blocks per run. During the experiment, each of the 48 melodies was presented once in a specific combination (e.g., melody no. 1 in the natural style female-low condition). Each melody appeared twice in a run, but in different combinations (e.g., melody no. 1 in the natural style female-low condition and in the operatic style male-high condition).

Auditory stimuli were delivered binaurally using MRI-compatible earphones (CONFON HP-IE-01, MR Confon GmbH, Magdeburg, Germany). Visual fixation was maintained by a dark-grey dot displayed on a mid-grey background on an MRI-compatible LCD screen (32″ NNL LCD Monitor, NordicNeuroLab, Bergen, Norway; 1920 × 1080 pixels; 60 Hz). The screen was positioned at a distance of 142 cm from the participant's eyes, and viewed through a mirror attached to the MRI head coil. The presentation of stimuli and the timing of the experimental procedure were controlled using custom-written scripts in MATLAB (R2015a, The MathWorks Inc., Natick, MA, USA) utilizing the Psychophysics Toolbox[51,52].

To ensure participants' attention to the auditory stimuli, they were engaged in a one-back memory task during the melody blocks. Participants had to identify and remember if the current melody was identical to the immediately preceding one and report the total number of one-back repetitions at the end of each run using a response grip (NordicNeuroLab, Bergen, Norway).

Three additional runs featuring natural, operatic, and pure tone (sine wave) melodies were conducted in a second fMRI session following a similar experimental setup. However, data from these additional runs were reserved for separate analyses not included in the current study.

Data acquisition

Data were acquired on a Siemens Magnetom Prisma 3 T MRI scanner (Siemens Healthcare GmbH, Erlangen, Germany) at the Brain Imaging Centre, Research Centre for Natural Sciences. A standard Siemens 32-channel

head coil was applied. For functional measurements, a blipped-controlled aliasing in parallel imaging (CAIPI) simultaneous multi-slice gradient-echo-EPI sequence[53] was used with sixfold slice acceleration. Full brain coverage was obtained with an isotropic 2 mm spatial resolution (208×208 mm field-of-view (FOV); anterior-to-posterior phase encoding direction; 104×104 in-plane matrix size; 54 slices; 25% slice gap) and a repetition time (TR) of 710 ms, without in-plane parallel imaging. A partial Fourier factor of 7/8 was used to achieve an echo time (TE) of 30 ms. Flip angle (FA) was 59°. Image reconstruction was performed using the Slice-GRAPPA algorithm[53] with LeakBlock kernel[54]. T1-weighted 3D MPRAGE anatomical imaging was performed using twofold in-plane GRAPPA acceleration with isotropic 1 mm spatial resolution (TR/TE/FA = 2300 ms/3 ms/9°; FOV = 256×256 mm).

Stimulus data processing

To assess the specific characteristics of vibrato in the singing stimuli (B3–E♯4) that fell well within the vocal capabilities of all participants, fundamental frequencies (f_0) of melodies matching participants' vocal range were extracted using the autocorrelation method in Praat (version 6.1.38). All f_0 trajectories were segmented at tone boundaries by thresholding the low-pass filtered ($f_c = 3$ Hz) trajectories at the fundamental frequency of two consecutive tones. To ensure clarity and remove transients, 50 ms were trimmed from the beginning and end of each segment. Vibrato parameters, such as rate and extent, were determined using a discrete Fourier transform-based method[55] with refined spectral peak detection to accommodate tones lacking distinct vibrato. Parameters were quantified by identifying the maximal spectral peak within the 4–8 Hz modulation range using Matlab's findpeaks() function. In cases where no spectral peak was detected, the vibrato rate was set to NaN, and the vibrato extent was set to zero.

Processing of neuroimaging data

The functional images were first spatially realigned using SPM12's two pass procedure. All realigned functional images were then co-registered with the T1-weighted anatomical image, which also facilitated surface reconstruction using the recon-all pipeline provided by FreeSurfer 7.1.1[56]. Once co-registered, functional images were resampled to fsaverage space and underwent iterative diffusion smoothing (40 iterations) as part of the CONN preprocessing pipeline to increase signal-to-noise ratio and to compensate for intersubject variability in functional activation.

Statistical modelling of neuroimaging data

For the first-level analysis of fMRI data, a standard voxel-wise general linear model (GLM) was implemented using SPM12. This model included regressors for operatic and natural voice stimuli convolved with the canonical hemodynamic response function to model BOLD responses. Only stimuli of the female-low and male-high voice types (B3–E♯4 vocal range) were included as primary regressors for the purpose of this study, ensuring that the analyses accurately reflected the participants' physiological capabilities to produce these sounds. This selective approach allows for a focused comparison of operatic versus natural singing styles within the participants' own vocal range, enhancing the relevance and interpretability of our findings. Additionally, operatic and natural voice stimuli from the female-high (C5–F♯5) and male-low (G2–C♯3) ranges were also included as independent regressors to account for their overall effects, although these data were not the focus of the current analysis and are not discussed further.

Temporal high-pass filtering was applied with a cutoff frequency of 1/128 Hz to remove low-frequency drifts. Movement-related variance was controlled by incorporating the estimated movement parameters from the realignment step, their temporal derivatives, and squares[57] into the model as nuisance covariates. Additionally, these movement-related parameters underwent band-stop filtering using a Butterworth IIR filter with a stopband of 0.2–0.5 Hz and a filter order of 12 to mitigate respiration-related effects[58,59]. Outlier scans, defined by framewise displacement exceeding 0.9 mm or global BOLD signal changes surpassing 5 SDs in the Artifact Detection Tools[60] implemented in the CONN toolbox were also accounted for using nuisance covariates. Temporal autocorrelations were modelled using SPM12's FAST model.

The first-level GLM produced contrast images based on parameter estimates (β values) of operatic and natural stimuli regressors, which were then used as input for second-level whole-brain random-effects analyses (one-sample t-tests). These primary analyses were conducted using a cluster-based permutation testing framework[61] implemented in the CONN toolbox, applying a two-sided FDR-corrected P value threshold of $P < 0.05$ for cluster-level significance to correct for multiple comparisons. Each significant cluster's mass (sum of the t-squared statistics across all voxels) and FDR-corrected P values (P_{FDR}) are reported.

Preprocessing and statistical analyses of the imaging data were performed using the SPM12 (v7771) toolbox (Wellcome Trust Centre for Neuroimaging, London, UK), the CONN (v.20.b) toolbox[62], as well as custom-made scripts running on MATLAB R2017a (The MathWorks Inc., Natick, MA, USA).

Region of interest (ROI) analysis

In our primary analysis above, we investigated the whole brain effect of stimuli style within the opera group by comparing BOLD responses to operatic versus natural voice stimuli. Significant differences (see Results for details) were observed bilaterally in early auditory regions, including the primary auditory cortex (A1), medial belt (MBelt), posterior belt (PBelt), and lateral belt (LBelt) complexes, as well as in the A4 region of the auditory association cortex, as identified by the Human Connectome Project Multi-Modal Parcellation atlas of human cortical areas (HCP-MMP1.0)[63]. Additionally, significant effects related to stimuli style were detected bilaterally in core areas of the default mode network (DMN), specifically the precuneus/posterior cingulate cortex (pCun/PCC) and the medial prefrontal cortex (mPFC), according to the 200-region Kong2022 parcellation atlas[64].

To determine if the effects observed in the voxel-wise analysis were unique to the opera group, we conducted an ROI analysis across both groups using atlas-based ROIs. Percent signal change (PSC) of voxels was calculated for each participant within each ROI. Separate linear mixed effects (LME) models were then applied to these PSC averages for the auditory and the core DMN ROIs. The models incorporated fixed-effects for group, stimuli style, hemisphere, ROI, along with their interactions, and random effect for participant to account for their inter-individual variability. Since stimuli style, hemisphere, and ROI are within-subject factors, and their effect may vary significantly between individuals, we also included random slopes for these. Given the significant age difference between groups (mean age was 39 in the opera and 29 in the control group), age could be a confounding factor when investigating the group effects. To mitigate this, we controlled for age by including it and its interaction with group as additional fixed effects.

Statistical analyses were performed in R (4.1.3)[65] using the afex (1.2.1)[66] and emmeans (1.8.5)[67] packages. Simple-effects were corrected for multiple comparisons using the single-step correction of the multcomp (1.4.23) package[68].

Results
Vibrato characteristics of stimuli
Analysis of melodies within participants' reproducible vocal range revealed distinct characteristics: the vibrato extent was 16.5 ± 7.0 cents (mean ± SD) for natural tones and 126.5 ± 16.1 cents for operatic tones. Similarly, vibrato rate was measured at 5.81 ± 0.57 Hz for natural and 5.76 ± 0.13 Hz for operatic tones. Melodies outside this range, specifically from the male-low and female-high melody sets, were utilized for addressing different research questions not covered in this report.

Whole brain analysis results
The whole brain analysis of the stimuli style contrast in opera singers revealed significant activations in several brain regions involved in sound processing. Specifically, significantly higher BOLD responses were found bilaterally for operatic compared to natural voice stimuli in the auditory cortex (left cluster mass = 74,484.6, $P_{FDR} = 0.0042$; right cluster mass = 50,768.28, $P_{FDR} = 0.0042$), including the early auditory regions, such as the A1, MBelt, PBelt, and LBelt, as well as the auditory association area A4. Furthermore, significant differences in BOLD responses were found bilaterally, between operatic and natural voice stimuli in the core regions of the DMN, namely in the pCun/PCC (left cluster mass = 17,346.09, $P_{FDR} = 0.013$; right cluster mass = 9178.01, $P_{FDR} = 0.027$) and the mPFC (left cluster mass = 16,539.71, $P_{FDR} = 0.013$; right cluster mass = 8336.37, $P_{FDR} = 0.027$). Finally, we identified a weak, yet statistically significant cluster (cluster mass = 5610.90, $P_{FDR} = 0.042$) exclusively in the left hemisphere, partially encompassing the left inferior parietal lobule (IPL) (see Fig. 1).

Conversely, the whole brain analysis by stimuli style in the control group (non-singers) revealed significant activations confined to the bilateral auditory cortex. Notably, BOLD responses to operatic versus natural voice stimuli were significantly higher, manifesting bilaterally in the auditory cortex with nearly identical regional extents to those observed in opera singers (left cluster mass = 82,857.32, $P_{FDR} = 0.00063$; right cluster mass = 84,947.41, $P_{FDR} = 0.00063$).

ROI analysis results
To determine if the singing style effects identified in the opera group were specific to opera singers, we conducted an atlas-based ROI analysis comparing BOLD responses between opera singers and controls across auditory cortical and core DMN regions.

Auditory cortical regions
The ROI analysis targeting auditory cortical regions—specifically A1, MBelt, PBelt, LBelt, and A4—demonstrated a significant main effect of stimuli style (F(1,63) = 77.47, $P < 0.0001$), indicating higher BOLD responses to operatic versus natural voice stimuli across all ROIs and hemispheres for both opera singers and control participants. Notably, the interaction between stimuli style and group was not significant (F(1,63) = 0.40, $P = 0.53$) (see Fig. 2), highlighting a universal, experience-independent effect of stimuli style in the auditory cortex.

Core DMN regions
In the core DMN regions, specifically the pCun/PCC and the mPFC, our analysis revealed both a significant main effect of stimuli style (F(1,63) = 8.26, $P = 0.0055$) and a significant interaction effect between group and stimuli style (F(1,63) = 4.09, $P = 0.047$). This interaction was predominantly driven by a pronounced difference in the opera group, where a significant preference for operatic over natural voice stimuli was observed (t = 3.44, $P = 0.0021$); such a distinction was not evident in the control group (t = 0.61, $P = 0.79$). Specifically, opera singers showed weaker negative BOLD responses to operatic compared to natural voice stimuli, a phenomenon not seen in the controls (see Fig. 3). Additionally, no significant differences were found between groups when analysing responses to operatic or natural stimuli independently ($|t| \leq 1.64$, $P \geq 0.16$).

Discussion
In this study, we investigated how extensive musical training—in particular, Western operatic singing—can shape brain responses to auditory perception of different singing styles. Focusing on vibrato, one of the acoustically most salient features of operatic singing, we compared BOLD responses of trained opera singers and musically untrained controls when listening to operatic (i.e., with vibrato) and natural (i.e., without vibrato) singing styles. Whole-brain results demonstrated that opera singers exhibited increased BOLD responses in bilateral auditory
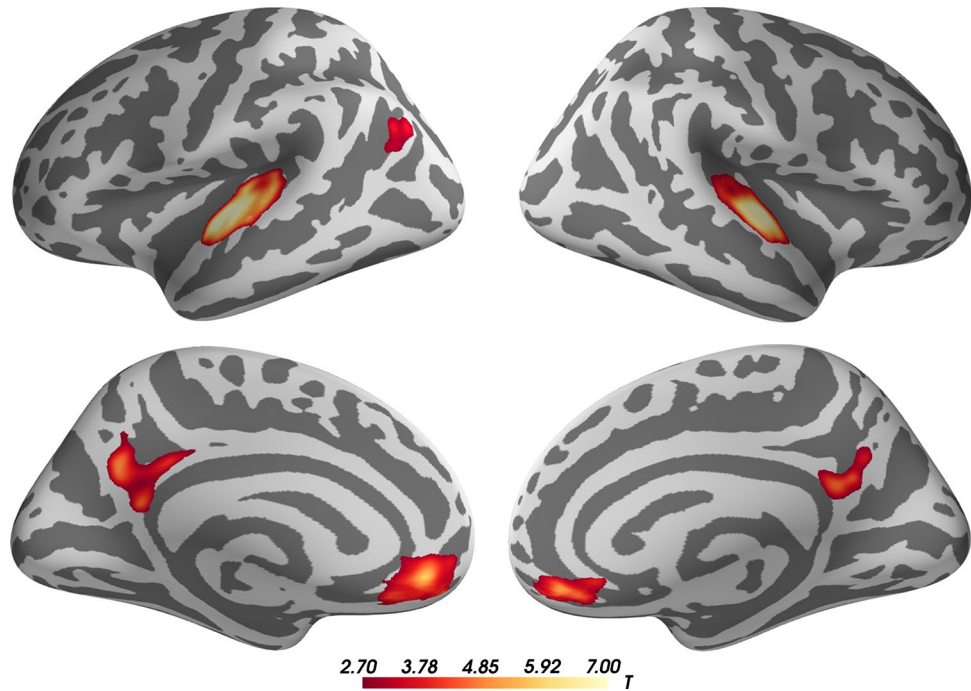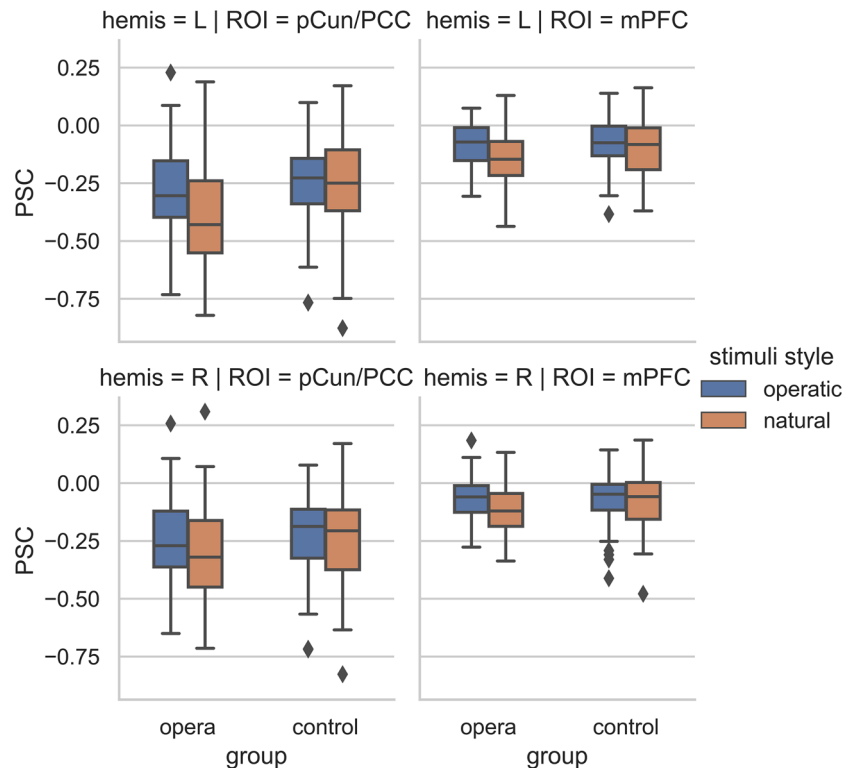
Fig. 1. Stimuli style effect on BOLD responses in opera singers. Heatmaps on inflated fsaverage cortical brain models show second-level statistical t-maps of the operatic vs. natural stimuli style contrast thresholded at FDR corrected $P < 0.05$. Significant stimuli style effects were revealed bilaterally in auditory cortical regions (A1, MBelt, PBelt, LBelt, and A4) shown in lateral view (top), in core regions of the default mode network (pCun/PCC and mPFC) shown in medial view (bottom) and only in the left hemisphere, partially overlapping with the left IPL shown in lateral view (top left).
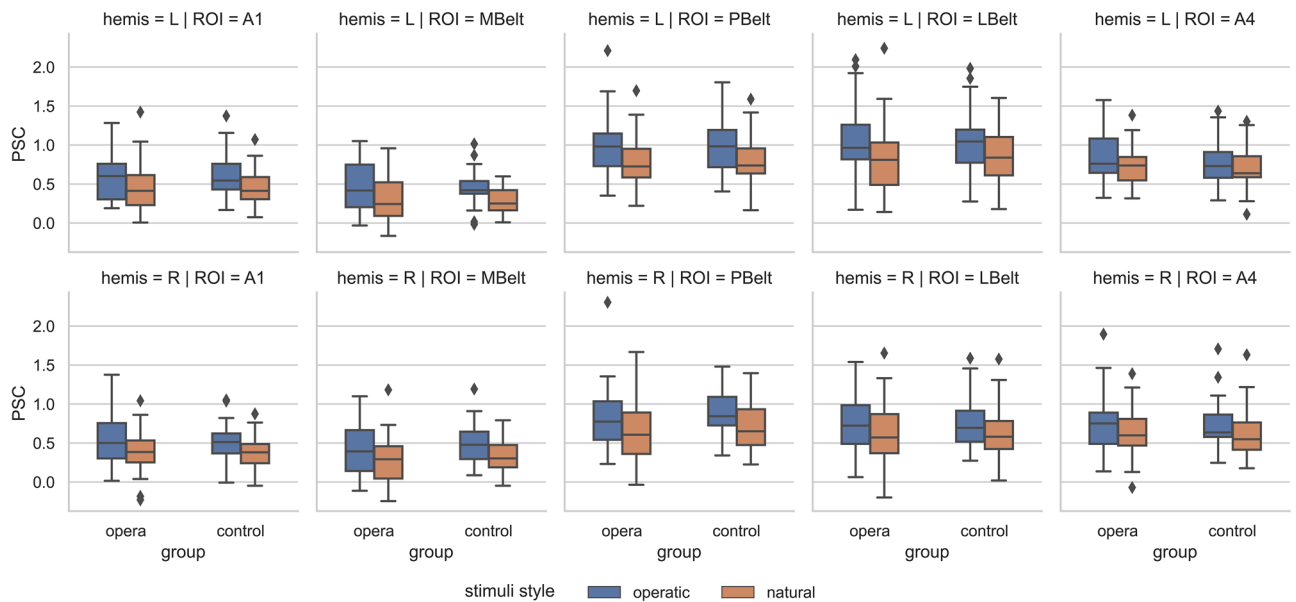
Fig. 2. Stimuli style effect on BOLD responses in auditory cortical regions. Boxplots show the effect of stimuli style (operatic vs. natural) and group (opera vs. control) on fMRI percent signal change (PSC) in both left (L) and right (R) hemispheric (hemis) auditory regions (A1, MBelt, PBelt, LBelt and A4 ROIs).

cortical regions and weaker negative responses in core regions of the DMN—including the pCun/PCC and the mPFC—as well as the left IPL when listening to operatic compared to natural singing voices. In contrast, among non-singers, differential BOLD responses were restricted to the bilateral auditory cortex.

Building upon these findings, an ROI analysis highlighted significant group differences in BOLD responses to operatic and natural singing voices. Importantly, an interaction effect within the DMN regions was driven by

Fig. 3. Stimuli style and group effects on BOLD responses in core default mode network (DMN) regions. Boxplots show the effect of stimuli style (operatic vs. natural) and group (opera vs. control) on fMRI percent signal change (PSC) in both left (L) and right (R) hemispheric (hemis) DMN regions (pCun/PCC and mPFC ROIs).

a significant difference in the opera singers that was absent in controls. Specifically, opera singers showed weaker negative BOLD responses to operatic compared to natural singing voices, suggesting that their extensive training and active experience with operatic singing have distinctly shaped neural responses to auditory perception related to sound features inherent in their specialised singing style.

Whole brain results in opera singers and non-singers

Differences in brain activation when comparing responses to operatic and natural singing voices were observed in auditory cortical regions in both opera singers and controls, specifically, in bilateral early auditory cortical regions and in the A4 complex of the auditory association cortex, a region associated with processing auditory stimulus complexity[69]. These differences may be attributed to the amplitude and frequency-modulated vibrato characteristic of operatic singing style. In addition to vibrato, a key feature of operatic singing is the spectral envelope, which is harmonically more complex compared to natural singing. The phenomenon of vocal ring in operatic voices, characterized by a clustering of frequency formants around three kHz, psychoacoustically enhances perceived loudness. This acoustic feature allows trained singers to effectively project their voices over the orchestral accompaniment[70,71].

Previous studies have demonstrated that variations in acoustic properties can lead to increased activation within the auditory cortex. For example, harmonic tones produce more activation than pure tones or single frequencies in the right Heschl's gyrus (HG) and bilaterally in the lateral supratemporal planes (STP)[72]. Similarly, frequency-modulated tones elicit more activation than static tones in bilateral HG, the anterolateral part of the lateral supratemporal planes, and the superior temporal sulci[72]. The posterolateral STP, extending into the lateral auditory association cortex, is particularly responsive to frequency modulation[73]. These observations are consistent with the role of posterior non-primary auditory areas in processing spectrotemporally complex sounds[74,75]. Indeed, both amplitude (intensity) and frequency modulation trigger greater activations in these regions compared to unmodulated tones[72]. Given that the intensity of our singing-voice stimuli was normalized, the primary distinctions between the stimuli lay in their spectrotemporal modulations. Therefore, the distinctive acoustic properties of operatic singing voices likely account for the observed style-specific differences in auditory cortex activation. This interpretation aligns with the lack of group difference in our ROI findings, indicating that the enhanced auditory cortical responses to operatic singing were not solely linked to musical expertise. However, only opera singers displayed style-specific activation patterns in cortical regions that are generally associated with the DMN.

ROI results

Building upon the voxel-wise findings, subsequent ROI analyses confirmed the absence of group differences in the auditory cortex. However, a noticeable interaction effect between the group and singing style emerged within the DMN, which was driven by the group of opera singers. Unlike the non-singer controls, opera singers exhibited significant differences in BOLD responses between operatic and natural singing styles within core DMN regions, particularly the mPFC and PCC. This result resonates with prior research demonstrating that music listening can alter the DMN connectivity, depending on the level of self-referential processing the stimuli induce and prior musical experience[76,77], in addition to experience-dependent effects in neural networks associated with auditory and sensorimotor, as well as cognitive and affective processes[5,7,9,11].

The DMN plays critical roles in forming contextual associations, mediating emotional and self-referential processes, fostering autobiographical memory, and facilitating mental state attribution, mind-wandering, and daydreaming[78–80]. According to the triple network model, network switching by the salience network alters the dynamic temporal interactions between DMN and the frontotemporal networks[81,82]. Relevant sensory stimuli can drive external event-driven switching from internally focused mental processes underscored by the DMN, thus suppressing DMN and disrupting the internal narrative. The disengagement of the DMN from cognitive control systems enables allocation of resources to serve external attention-demanding, goal-directed behaviours. Through its characteristic connectivity and coactivation patterns, individual DMN nodes may support specific network functions, as they are not uniformly engaged across cognitive domains. In this context, PCC and mPFC are implicated as forming a core self-referential system within the DMN, with these nodes playing distinct roles in differentiating self from others[82].

Listening to music, especially when it evokes introspective thought and self-referential processes, has been linked to increased DMN connectivity, as seen during aesthetic contemplation and engagements with preferred music[76]. This finding aligns with research showing that classical music training can rewire the DMN, leading to enhanced connectivity within this network during resting-state fMRI[83]. Moreover, recent studies such as Mørup et al.[84] and Liao et al.[85] have highlighted the DMN's role in tasks that involve motor coordination and rhythm production, particularly as these tasks become more automatic and internally driven in professional musicians, consistent with the DMN's function in maintaining internal representations and supporting cognitive processes that require less focused attention during music production[76,85,86].

However, in tasks involving free listening to complete musical pieces, the connectivity patterns in DMN-related nodes appear differently for non-musicians and musicians. Non-musicians exhibit higher fMRI connectivity, suggesting introspective engagement, whereas musicians display altered patterns indicative of a more specialized auditory processing approach[77]. Similarly, using electroencephalography, decreased functional connectivity between the ventromedial prefrontal cortex (vmPFC) and the pCun/PCC has been observed in musicians with absolute pitch, as compared to non-musicians, when listening to whole orchestral works[87]. This divergence from our findings could originate from the distinct acoustic stimuli used and their relevance to the specific sensorimotor experiences of the musicians involved. Although speculative, it is possible that naturalistic music listening tasks engage musicians with diverse training backgrounds in more externally focused auditory processing strategies, typically associated with DMN suppression. In contrast, the presence or absence of vibrato—a key feature of operatic and natural singing styles—may consistently evoke autobiographical memories tied to internal representations of embodied sensorimotor experiences of opera singers.

In line with this understanding, recent research has teased apart distinct conceptualizations of self-referential processing within the DMN, highlighting its complexity[79,88]. For instance, activation of the vmPFC and dorsomedial prefrontal cortex has been linked to valenced self-evaluation[89]. The activation of the frontal cluster, which includes the vmPFC, corresponds to the degree of self-relevance of intero- or exteroceptive stimuli[90]. Conversely, the posterior cluster, which encompasses the PCC, is thought to integrate self-referential information into the temporal context and autobiographical memory[91]. These core regions of the DMN are thought to facilitate the integration of incoming sensory inputs with prior internally generated representations to predict and guide perception, learning, and action control in both naturalistic activities and cognitive tasks, suggesting a role of the DMN in maintaining and updating detailed mental models[92,93]. This enables experts to engage in more efficient and automated cognitive processes[94], as evidenced by our study observing increased DMN activity (reduced DMN deactivation) in trained opera singers when they listened to melodies sung in operatic style. Thus, opera singers may integrate key auditory features of operatic singing with their extensive prior knowledge and training. In combination with the increased engagement of auditory cortical areas, the DMN may aid in more efficient processing and integration of training-specific auditory stimuli, highlighting the plasticity of the DMN in adapting to the cognitive demands of expert-level singing ability.

This extensive and continuous learning process, necessary for mastering operatic singing, often requires around 10,000 h of deliberate, goal-oriented practice before the age of 20 years[95]. Such rigorous training integrates these complex skills into an individual's self-model, impacting future self-perceptions and evaluations[96,97]. Our findings demonstrate significantly higher activation within the vmPFC and PCC—key areas associated with self-referential processing—when opera singers listened to vibrato, a hallmark of their trained style. This suggests that the skills necessary to produce vibrato are deeply embedded in the singers' self-models, influencing their perception and engagement with operatically styled auditory cues. Notably, this heightened activation was absent when opera singers listened to natural singing styles, and was not observed in the non-musician controls, reinforcing the unique connection between specialized training and DMN activation.

Conclusion

This study provides novel evidence suggesting that specialized musical sensorimotor training, as exemplified by Western operatic singing, can significantly alter brain responses during the auditory perception of acoustic features that are salient and relevant to one's own singing style, such as vocal vibrato. These findings have pivotal implications for understanding the neural underpinnings of musical training, particularly in the context of vocal music, reinforcing the concept that our perceptions are actively shaped by past experiences and actions[19]. The enhanced BOLD responses observed in the core DMN regions, specifically the vmPFC and the PCC, in opera singers when listening to their own singing style underscore the DMN's crucial role in self-referential processing[82]. These effects suggest that extensive training in operatic singing not only hones musical expertise but also refines acoustically determined self-referential processes, potentially impacting an individual's self-model. This underscores the intriguing interplay between skill acquisition, self-representation, and neural plasticity. Future research should delve deeper into this relationship, extending beyond cross-sectional designs to provide more causal evidence of the effects of long-term singing training on cognition and behaviour.

Data availability

Data supporting the findings of this study are available from the corresponding authors on reasonable request.

References

1. Penhune, V. B. Sensitive periods in human development: Evidence from musical training. *Cortex* **47**, 1126–1137 (2011).
2. Platz, F., Kopiez, R., Lehmann, A. C. & Wolf, A. The influence of deliberate practice on musical achievement: a meta-analysis. *Front. Psychol.* **5**, (2014).
3. Kraus, N. & Chandrasekaran, B. Music training for the development of auditory skills. *Nat. Rev. Neurosci.* **11**, 599–605 (2010).
4. Pantev, C., Paraskevopoulos, E., Kuchenbuch, A., Lu, Y. & Herholz, S. C. Musical expertise is related to neuroplastic changes of multisensory nature within the auditory cortex. *Eur. J. Neurosci.* **41**, 709–717 (2015).
5. Saari, P., Burunat, I., Brattico, E. & Toiviainen, P. Decoding musical training from dynamic processing of musical features in the brain. *Sci. Rep.* **8**, 708 (2018).
6. Brattico, E. *et al.* It's sad but i like it: The neural dissociation between musical emotions and liking in experts and laypersons. *Front. Hum. Neurosci.* **9**, 676 (2015).
7. Särkämö, T. Music for the ageing brain: Cognitive, emotional, social, and neural benefits of musical leisure activities in stroke and dementia. *Dementia* **17**, 670–685 (2018).
8. Herholz, S. C. & Zatorre, R. J. Musical training as a framework for brain plasticity: Behavior, function, and structure. *Neuron* **76**, 486–502 (2012).
9. Reybrouck, M. & Brattico, E. Neuroplasticity beyond sounds: Neural adaptations following long-term musical aesthetic experiences. *Brain Sci.* **5**, 69–91 (2015).
10. Zamorano, A. M. *et al.* Singing training predicts increased insula connectivity with speech and respiratory sensorimotor areas at rest. *Brain Res.* **1813**, 148418 (2023).
11. Zamorano, A. M., Cifre, I., Montoya, P., Riquelme, I. & Kleber, B. Insula-based networks in professional musicians: Evidence for increased functional connectivity during resting state fMRI. *Hum. Brain Mapp.* **38**, 4834–4849 (2017).
12. Leipold, S., Klein, C. & Jäncke, L. Musical expertise shapes functional and structural brain networks independent of absolute pitch ability. *J. Neurosci.* **41**, 2496–2511 (2021).
13. Klein, C., Liem, F., Hänggi, J., Elmer, S. & Jäncke, L. The "silent" imprint of musical training. *Hum. Brain Mapp.* **37**, 536–546 (2016).
14. Münte, T. F., Altenmüller, E. & Jäncke, L. The musician's brain as a model of neuroplasticity. *Nat Rev Neurosci* **3**, 473–478 (2002).
15. Koelsch, S., Vuust, P. & Friston, K. Predictive processes and the peculiar case of music. *Trends Cognit. Sci.* **23**, 63–77 (2019).
16. Krishnan, S. *et al.* Beatboxers and guitarists engage sensorimotor regions selectively when listening to the instruments they can play. *Cereb. Cortex* **28**, 4063–4079 (2018).
17. Lahav, A., Saltzman, E. & Schlaug, G. Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *J. Neurosci.* **27**, 308–314 (2007).
18. Ohl, F. W. & Scheich, H. Learning-induced plasticity in animal and human auditory cortex. *Curr. Opin. Neurobiol.* **15**, 470–477 (2005).
19. Buzsáki, G., Peyrache, A. & Kubie, J. Emergence of cognition from action. *Cold Spring Harb. Symp. Quant. Biol.* **79**, 41–50 (2014).
20. Hommel, B., Müsseler, J., Aschersleben, G. & Prinz, W. The theory of event coding (TEC): A framework for perception and action planning. *Behav. Brain Sci.* **24**, 849–878 (2001).
21. Prinz, W. A Common coding approach to perception and action. in *Relationships between perception and action: Current approaches* (eds. Neumann, O. & Prinz, W.) 167–201 (Springer, Berlin, Heidelberg, 1990). https://doi.org/10.1007/978-3-642-75348-0_7.
22. Vuust, P., Heggli, O. A., Friston, K. J. & Kringelbach, M. L. Music in the brain. *Nat. Rev. Neurosci.* **23**, 287–305 (2022).
23. Vuust, P. & Witek, M. A. G. Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music. *Front. Psychol.* **5**, (2014).
24. Margulis, E. H., Mlsna, L. M., Uppunda, A. K., Parrish, T. B. & Wong, P. C. M. Selective neurophysiologic responses to music in instrumentalists with different listening biographies. *Hum. Brain Mapp.* **30**, 267–275 (2009).
25. Pantev, C., Roberts, L. E., Schulz, M., Engelien, A. & Ross, B. Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport* **12**, 169–174 (2001).
26. Pantev, C. *et al.* Increased auditory cortical representation in musicians. *Nature* **392**, 811–814 (1998).
27. Pantev, C. & Herholz, S. C. Plasticity of the human auditory cortex related to musical training. *Neurosci. Biobehav. Rev.* **35**, 2140–2154 (2011).
28. Shahin, A. J., Roberts, L. E., Chau, W., Trainor, L. J. & Miller, L. M. Music training leads to the development of timbre-specific gamma band activity. *NeuroImage* **41**, 113–122 (2008).
29. Dichter, B. K., Breshears, J. D., Leonard, M. K. & Chang, E. F. The control of vocal pitch in human laryngeal motor cortex. *Cell* **174**, 21-31.e9 (2018).
30. Mürbe, D., Zahnert, T., Kuhlisch, E. & Sundberg, J. Effects of professional singing education on vocal vibrato—a longitudinal study. *J. Voice* **21**, 683–688 (2007).
31. Sundberg, J. 6—The perception of singing. in *The Psychology of Music (Second Edition)* (ed. Deutsch, D.) 171–214 (Academic Press, San Diego, 1999). https://doi.org/10.1016/B978-012213564-4/50007-X.
32. Sundberg, J. Perceptual aspects of singing. *J. Voice* **8**, 106–122 (1994).

33. Anand, S., Wingate, J. M., Smith, B. & Shrivastav, R. Acoustic parameters critical for an appropriate vibrato. *J. Voice* **26**(820), e19-820.e25 (2012).
34. Dromey, C., Holmes, S. O., Hopkin, J. A. & Tanner, K. The effects of emotional expression on vibrato. *J. Voice* **29**, 170–181 (2015).
35. Scherer, K. R., Sundberg, J., Fantini, B., Trznadel, S. & Eyben, F. The expression of emotion in the singing voice: Acoustic patterns in vocal performance. *J. Acoust. Soc. Am.* **142**, 1805–1815 (2017).
36. Kleber, B., Zeitouni, A. G., Friberg, A. & Zatorre, R. J. Experience-dependent modulation of feedback integration during singing: Role of the right anterior insula. *J. Neurosci.* **33**, 6070–6080 (2013).
37. Kleber, B., Veit, R., Birbaumer, N., Gruzelier, J. & Lotze, M. The brain of opera singers: Experience-dependent changes in functional activation. *Cereb. Cortex* **20**, 1144–1152 (2010).
38. Kleber, B., Birbaumer, N., Veit, R., Trevorrow, T. & Lotze, M. Overt and imagined singing of an Italian aria. *NeuroImage* **36**, 889–900 (2007).
39. Zarate, J. M. The neural control of singing. *Front. Hum. Neurosci.* **7**, 237 (2013).
40. Zarate, J. M. & Zatorre, R. J. Experience-dependent neural substrates involved in vocal pitch regulation during singing. *NeuroImage* **40**, 1871–1887 (2008).
41. Zatorre, R., Delhommeau, K. & Zarate, J. Modulation of auditory cortex response to pitch variation following training with microtonal melodies. *Front. Psychol.* **2**, (2012).
42. Finkel, S. *et al.* Intermittent theta burst stimulation over right somatosensory larynx cortex enhances vocal pitch-regulation in nonsingers. *Hum. Brain Mapp.* **40**, 2174–2187 (2019).
43. Guenther, F. H. *Neural Control of Speech* (The MIT Press, USA, 2016).
44. Serino, A. Peripersonal space (PPS) as a multisensory interface between the individual and the environment, defining the space of the self. *Neurosci. Biobehav. Rev.* **99**, 138–159 (2019).
45. Price, C. J. A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage* **62**, 816–847 (2012).
46. Cohen, A. J., Levitin, D. J. & Kleber, B. Brain mechanisms underlying singing. In *The Routledge companion to interdisciplinary studies in singing volume I: development* (eds Russo, F. A. *et al.*) 79–86 (Routledge, UK, 2020).
47. Kleber, B. & Zarate, J. M. The Neuroscience of Singing. In *The Oxford Handbook of Singing* (eds Welch, G. F. *et al.*) (Oxford University Press, Oxford, New York, 2014).
48. Coffey, E. B. J., Herholz, S. C., Scala, S. & Zatorre, R. J. Montreal Music History Questionnaire: a tool for the assessment of music-related experience in music cognition research. in *The Neurosciences and Music IV: Learning and Memory, Conference. Edinburgh, UK* (2011).
49. Lévêque, Y. & Schön, D. Listening to the human voice alters sensorimotor brain rhythms. *PLOS ONE* **8**, e80659 (2013).
50. Lévêque, Y. & Schön, D. Modulation of the motor cortex during singing-voice perception. *Neuropsychologia* **70**, 58–63 (2015).
51. Brainard, D. H. The psychophysics toolbox. *Spat. Vis.* **10**, 433–436 (1997).
52. Pelli, D. G. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* **10**, 437–442 (1997).
53. Setsompop, K. *et al.* Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magn. Reson. Med.* **67**, 1210–1224 (2012).
54. Cauley, S. F., Polimeni, J. R., Bhat, H., Wald, L. L. & Setsompop, K. Interslice leakage artifact reduction technique for simultaneous multislice acquisitions. *Magn. Reson. Med.* **72**, 93–102 (2014).
55. Pang, H.-S., Lim, J. & Lee, S. Discrete Fourier transform-based method for analysis of a vibrato tone. *J. New Music Res.* **49**, 307–319 (2020).
56. Fischl, B. *et al.* Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron* **33**, 341–355 (2002).
57. Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S. J. & Turner, R. Movement-related effects in fMRI time-series. *Magn. Reson. Med.* **35**, 346–355 (1996).
58. Fair, D. A. *et al.* Correction of respiratory artifacts in MRI head motion estimates. *Neuroimage* **208**, 116400 (2020).
59. Power, J. D. *et al.* Distinctions among real and apparent respiratory motions in human fMRI data. *Neuroimage* **201**, 116041 (2019).
60. Whitfield-Gabrieli, S., Nieto-Castanon, A., & Ghosh, S. Artifact detection tools (ART). Cambridge, MA. Release Version, 7(19), 11 (2011). https://www.nitrc.org/projects/artifact_detect/.
61. Bullmore, E. T. *et al.* Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Trans. Med. Imaging* **18**, 32–42 (1999).
62. Whitfield-Gabrieli, S. & Nieto-Castanon, A. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect.* **2**, 125–141 (2012).
63. Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).
64. Kong, R. *et al.* Individual-specific areal-level parcellations improve functional connectivity prediction of behavior. *Cereb. Cortex* **31**, 4477–4500 (2021).
65. R Core Team (2020).—European Environment Agency. https://www.eea.europa.eu/mobile/data-and-maps/indicators/oxygen-consuming-substances-in-rivers/r-development-core-team-2006.
66. Singmann, H. *et al.* Afex: Analysis of Factorial Experiments R package version 1.2–1. https://CRAN.R-project.org/package=afex. (2023).
67. Lenth, R. V. *et al.* Emmeans: Estimated marginal means, aka least-squares means. (2023).
68. Hothorn, T., Bretz, F. & Westfall, P. Simultaneous inference in general parametric models. *Biom. J.* **50**, 346–363 (2008).
69. Güçlütürk, Y., Güçlü, U., van Gerven, M. & van Lier, R. Representations of naturalistic stimulus complexity in early and associative visual and auditory cortices. *Sci. Rep.* **8**, 3439 (2018).
70. Sundberg, J. Level and center frequency of the singer's formant. *J. Voice* **15**, 176–186 (2001).
71. Titze, I. R. & Story, B. H. Acoustic interactions of the voice source with the lower vocal tract. *J. Acoust. Soc. Am.* **101**, 2234–2243 (1997).
72. Hart, H. C., Palmer, A. R. & Hall, D. A. Amplitude and frequency-modulated stimuli activate common regions of human auditory cortex. *Cereb. Cortex* **13**, 773–781 (2003).
73. Hall, D. A. *et al.* Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* **12**, 140–149 (2002).
74. Hall, D. A., Hart, H. C. & Johnsrude, I. S. Relationships between human auditory cortical structure and function. *Audiol. Neurotol.* **8**, 1–18 (2003).
75. Zatorre, R. J., Belin, P. & Penhune, V. B. Structure and function of auditory cortex: music and speech. *Trends Cognit. Sci.* **6**, 37–46 (2002).
76. Reybrouck, M., Vuust, P. & Brattico, E. Brain connectivity networks and the aesthetic experience of music. *Brain Sci.* **8**, 107 (2018).
77. Alluri, V. *et al.* Connectivity patterns during music listening: Evidence for action-based processing in musicians. *Hum. Brain Mapp.* **38**, 2955–2970 (2017).
78. Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R. & Buckner, R. L. Functional-anatomic fractionation of the brain's default network. *Neuron* **65**, 550–562 (2010).
79. Frewen, P. *et al.* Neuroimaging the consciousness of self: Review, and conceptual-methodological framework. *Neurosci. Biobehav. Rev.* **112**, 164–212 (2020).
80. Bar, M. The proactive brain: Memory for predictions. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 1235–1243 (2009).

81. Sridharan, D., Levitin, D. J. & Menon, V. A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proc. Natl. Acad. Sci.* **105**, 12569–12574 (2008).
82. Menon, V. 20 years of the default mode network: A review and synthesis. *Neuron* **0**, (2023).
83. Belden, A. *et al.* Improvising at rest: Differentiating jazz and classical music training with resting state functional connectivity. *NeuroImage* **207**, 116384 (2020).
84. Mårup, S. H., Kleber, B. A., Møller, C. & Vuust, P. When direction matters: Neural correlates of interlimb coordination of rhythm and beat. *Cortex* **172**, 86–108 (2024).
85. Liao, Y.-C. *et al.* Inner sense of rhythm: percussionist brain activity during rhythmic encoding and synchronization. *Front. Neurosci.* **18**, (2024).
86. Loui, P. Rapid and flexible creativity in musical improvisation: review and a model. *Ann. N. Y. Acad. Sci.* **1423**, 138–145 (2018).
87. Brauchli, C., Leipold, S. & Jäncke, L. Diminished large-scale functional brain networks in absolute pitch during the perception of naturalistic music and audiobooks. *NeuroImage* **216**, 116513 (2020).
88. Bao, Z., Howidi, B., Burhan, A. M. & Frewen, P. Self-referential processing effects of non-invasive brain stimulation: A systematic review. *Front. Neurosci.* **15**, (2021).
89. Fingelkurts, A. A., Fingelkurts, A. A. & Kallio-Tamminen, T. Selfhood triumvirate: From phenomenology to brain activity and back again. *Conscious. Cognit.: Int. J.* **86**, 103031 (2020).
90. Sui, J. & Gu, X. Self as object: Emerging trends in self research. *Trends Neurosci.* **40**, 643–653 (2017).
91. Na, C.-H., Jütten, K., Forster, S. D., Clusmann, H. & Mainz, V. Self-referential processing and resting-state functional MRI connectivity of cortical midline structures in glioma patients. *Brain Sci.* **12**, 1463 (2022).
92. Stawarczyk, D., Bezdek, M. A. & Zacks, J. M. Event representations and predictive processing: The role of the midline default network core. *Top. Cognit. Sci.* **13**, 164–186 (2021).
93. Smallwood, J. *et al.* The default mode network in cognition: a topographical perspective. *Nat. Rev. Neurosci.* **22**, 503–513 (2021).
94. Vatansever, D., Menon, D. K. & Stamatakis, E. A. Default mode contributions to automated information processing. *Proc. Natl. Acad. Sci.* **114**, 12821–12826 (2017).
95. Breivik, G. The role of skill in sport. *Sport, Eth. Philos.* **10**, 222–236 (2016).
96. Ravn, S. Embodied learning in physical activity: Developing skills and attunement to interaction. *Front. Sports Act. Living* **4**, (2022).
97. Newen, A. The embodied self, the pattern theory of self, and the predictive mind. *Front. Psychol.* **9**, (2018).

Acknowledgements

Author contributions

Conception and design: P.H., B.K., Z.V., J.Z.; Analysis and interpretation: A.B., P.H., Á.N., Z.V.; Data collection: A.B., P.H., Á.N.; Writing the article: A.B., P.H., B.K., Á.N., Z.V., J.Z.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.B. or Z.V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.