Structural Neuroplasticity Effects of Singing in Chronic Aphasia

Aleksi J. Sihvonen,[1,2,3]* Anni Pitkäniemi,[1]* Sini-Tuuli Siponkoski,[1] Linda Kuusela,[4] Noelia Martínez-Molina,[1] Sari Laitinen,[5] Essi-Reetta Särkämö,[6] Johanna Pekkola,[4] Susanna Melkas,[3] Gottfried Schlaug,[7,8] Viljami Sairanen,[4] and Teppo Särkämö[1]

[1]Cognitive Brain Research Unit and Centre of Excellence in Music, Mind, Body and Brain, Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki 00014, Finland, [2]School of Health and Rehabilitation Sciences, Queensland Aphasia Research Centre and UQ Centre for Clinical Research, The University of Queensland, Brisbane QLD 4072, Australia, [3]Department of Neurology, University of Helsinki and Helsinki University Hospital, Helsinki 00029, Finland, [4]HUS Helsinki Medical Imaging Center, Helsinki University Hospital, Helsinki 00029, Finland, [5]Espoo Hospital, Espoo 00029, Finland, [6]Private Choir Conductor, Vantaa 01520, Finland, [7]Department of Neurology, UMass Medical School, Springfield, Massachusetts 01655, and [8]Department of Biomedical Engineering and Institute of Applied Life Sciences, UMass Amherst, Amherst, Massachusetts 01655

Abstract

Singing-based treatments of aphasia can improve language outcomes, but the neural benefits of group-based singing in aphasia are unknown. Here, we set out to determine the structural neuroplasticity changes underpinning group-based singing-induced treatment effects in chronic aphasia. Twenty-eight patients with at least mild nonfluent poststroke aphasia were randomized into two groups that received a 4-month multicomponent singing intervention (singing group) or standard care (control group). High-resolution T1 images and multishell diffusion-weighted MRI data were collected in two time points (baseline/5 months). Structural gray matter (GM) and white matter (WM) neuroplasticity changes were assessed using language network region of interest-based voxel-based morphometry (VBM) and quantitative anisotropy-based connectometry, and their associations to improved language outcomes (Western Aphasia Battery Naming and Repetition) were evaluated. Connectometry analyses showed that the singing group enhanced structural WM connectivity in the left arcuate fasciculus (AF) and corpus callosum as well as in the frontal aslant tract (FAT), superior longitudinal fasciculus, and corticostriatal tract bilaterally compared with the control group. Moreover, in VBM, the singing group showed GM volume increase in the left inferior frontal cortex (Brodmann area 44) compared with the control group. The neuroplasticity effects in the left BA44, AF, and FAT correlated with improved naming abilities after the intervention. These findings suggest that in the poststroke aphasia group, singing can bring about structural neuroplasticity changes in left frontal language areas and in bilateral language pathways, which underpin treatment-induced improvement in speech production.

Key words: aphasia; music; rehabilitation; singing; stroke; structural connectivity

Significance Statement

Understanding the neural underpinnings of improved language outcomes in aphasia is vital. We utilize longitudinal neuroplasticity measures of both gray matter (GM) and white matter (WM) and evaluate their contribution to group-based singing treatment effects in chronic aphasia. The results show that singing intervention induced GM and WM neuroplasticity changes in the left frontal language-related structures, but also in the right hemisphere (WM), correlating with improved naming

abilities. These results shine light on treatment-induced structural changes in chronic aphasia and improve our understanding of aphasia rehabilitation.

Introduction

Poststroke aphasia (PSA) is a common and debilitating consequence of stroke, with 60% of the patients remaining aphasic even 1 year after the stroke (Pedersen et al., 2004). PSA incurs enormous socioeconomical burden to the society (Olesen et al., 2012) and is associated with greater extent of rehabilitation services required (Dickey et al., 2010). Due to these profound consequences, effective interventions of aphasia are of great necessity.

The recovery of language functions after PSA is a complex process involving structural reorganization of the spared neurons within the language network according to the neural and neurocomputational models of PSA recovery (Stefaniak et al., 2020). Yet, research mapping treatment-induced structural changes supporting the reorganization processes that bring about beneficial behavioral change in PSA has remained scarce. Small-scale ($N = 1–8$) within-subject longitudinal studies of PSA have reported structural white matter (WM) changes in the left or right arcuate fasciculus (AF) or inferior longitudinal fasciculus following different types of speech therapy, such as constraint-induced language therapy or anomia treatment (Breier et al., 2011; Van Hees et al., 2014; McKinnon et al., 2017) or singing-based therapy, such as melodic intonation therapy (MIT; Schlaug et al., 2009), as well as in the left frontal regions and anterior corpus callosum following excitatory repetitive transcranial magnetic stimulation (Allendorfer et al., 2012). To date, only a single aphasia study has reported structural changes in a randomized controlled trial (RCT), showing that self-managed spoken word comprehension therapy increased gray matter (GM) or WM density in the left and right temporal regions (Fleming et al., 2020). Taken together, our understanding of PSA treatment-induced structural changes that support the recovery of language function is far from comprehensive.

Aphasia can often lead to depression and decreased social functioning, influencing patients' engagement with the therapy and affecting the outcome (Hilari et al., 2012; Mitchell et al., 2017). Treating associated cognitive impairments and secondary effects of aphasia (depression and social isolation) should also improve communication and, ultimately, quality of life (Doogan et al., 2018). Yet, therapeutic interventions for PSA are implemented traditionally in individual setting, lacking social stimulation. Multifaceted group-based treatments focusing on both language and associated cognitive impairments as well as psychosocial effects of PSA should be of great value and ideal health economically due to limited rehabilitation resources (Doogan et al., 2018). Moreover, as the neuroplasticity changes supporting recovery after stroke can be enhanced by increasing stimulation from the environment (Nithianantharajah and Hannan, 2006; Baroncelli et al., 2010), especially when it involves also a social component (Johansson and Ohlsson, 1996; Venna et al., 2014; Zhang et al., 2021), a combination of increased social, auditory, and cognitive stimulation could provide an avenue to tap into enhanced poststroke neuroplasticity supporting recovery of function in aphasia.

In subacute stroke patients, mere daily vocal music listening (i.e., auditory stimulation) has been shown to improve poststroke language recovery in PSA, increasing the GM volume in the left temporal regions (Sihvonen et al., 2020), strengthening the functional connectivity of the language and default mode networks (Sihvonen et al., 2020, 2021a), and enhancing the structural connectivity of the left frontal aslant tract (FAT) and stimulus-specific activation of its superior frontal termination areas compared with audiobook listening (Sihvonen et al., 2021b). Furthermore, MIT, a singing-based intervention for treating nonfluent aphasia, has been shown to improve connected speech, naming, and repetition (Sparks et al., 1974; Van Der Meulen et al., 2014; Zumbansen et al., 2014) and linking the positive effects to temporal and frontal speech motor areas, either in the left (Belin et al., 1996; Breier et al., 2011) or right (Schlaug et al., 2008; Wan et al., 2014; Tabei et al., 2016) hemisphere. In the healthy older adults, regular singing has recently been linked to enhanced executive function (Pentikäinen et al., 2021; Vetere et al., 2024), frontotemporal auditory functioning (Pentikäinen et al., 2022), structural connectivity (Perron et al., 2021), and structural plasticity in auditory and dorsal speech regions (Perron et al., 2022), suggesting that it may have neuroprotective effects in aging.

*A.J.S. and A.P. contributed equally to this work.

Correspondence should be addressed to Aleksi J. Sihvonen at aleksi.sihvonen@helsinki.fi.

The abovementioned increased social, auditory, and stimulation effects could be further strengthened with group-based singing regimens, where, additionally, the spoken language network is stimulated via producing linguistic and musical information via singing (Loui, 2015) that provide further neural stimulation supporting the recovering brain (Murphy and Corbett, 2009). In this vein, a recently published study has explored the behavioral effect of group-based singing intervention in treating patients with chronic aphasia (Siponkoski et al., 2022). Compared to standard care, singing improved responsive speech (Stockert et al., 2020) as indexed by Western Aphasia Battery (WAB) Naming and Repetition indices as well as enhanced long-term communication abilities and reduced family caregiver burden in questionnaires. However, the specific underlying neuroanatomical mediators of recovery of singing-based treatments for PSA are still unclear (Van Der Meulen et al., 2012).

Here, we set out to determine the structural neuroplasticity benefits of group-based singing in an RCT of 28 patients with chronic PSA from Siponkoski et al. (2022). To do so, we evaluated the structural GM and WM neuroplasticity changes after a 4-month singing intervention using language network region of interest (ROI)-based GM analyses and structural connectometry (Yeh et al., 2016a) based on high-resolution T1 images and multishell diffusion-weighted MRI (DW-MRI) data. Previous studies on MIT (Schlaug et al., 2009; Wan et al., 2014) and vocal music interventions (Sihvonen et al., 2021a, b) have shown improved language outcomes in association with neuroplasticity changes in the bilateral frontal areas. Therefore, we hypothesized that the singing intervention induces neuroplasticity changes especially in the left and right frontal areas within the language network that would underpin the improved naming and repetition abilities.

Materials and Methods

Subjects and study design. Fifty-four participants with PSA were successfully recruited from the Helsinki region through patient organizations (Helsinki-Uusimaa Stroke Association and Finnish Brain Association) and clinical speech therapists to a registered RCT (ClinicalTrials.gov, NCT03501797). Data collection was performed in two waves, with 33 participants with PSA enrolled and randomized to the trial in Jan. 2018 and 21 participants with PSA in Jan. 2019. Data collection was completed in Dec. 2019. The inclusion criteria were (1) age over 18, (2) Finnish-speaking, (3) time since stroke >6 months, (4) at least mild nonfluent aphasia due to stroke assessed by the Boston Diagnostic Aphasia Examination (BDAE) Aphasia Severity Rating Scale (score ≤4; Goodglass and Kaplan, 1983), (5) normal hearing, (6) no severe cognitive impairment affecting comprehension (e.g., memory disorder or perceptual deficit), (7) no neurological or psychiatric comorbidity or substance abuse, and (8) ability to produce vocal sounds through singing or humming. All participants were interviewed for eligibility by recruiting psychologists (authors A.P. and S-T.S.) and ensured that the patient was able to understand the purpose of the study. The study was conducted in conformance with the Declaration of Helsinki and was approved by the Helsinki University Hospital Ethics Committee. Written informed consent was obtained from all patients and participating caregivers.

Of the full sample ($N = 54$), 33 were eligible at the recruitment stage to undergo MRI and were randomly assigned to two groups stratified for aphasia severity (preliminary BDAE score), family caregiver's participation in training sessions, sex, age, and time since stroke (Fig. 1). The randomization was performed using a random number generator by a researcher not involved in data collection. Outcome measurements including neuropsychological assessment and MRI scans were conducted at baseline (T1) and after the intervention period at the 5-month stage (T2). Additionally, the participants filled out a demographic, musical, and clinical background questionnaire at baseline and also reported at T2 the amount of other rehabilitation received between T1 and T2. In total, 28 patients completed the study from T1 to T2 (singing group $N = 13$, control group $N = 15$) and were included in the analyses (Fig. 1, Table 1). All patients received standard chronic stroke care and rehabilitation throughout the study. There were no significant differences between the groups in the amount of received therapy/rehabilitation (Table 1).

Intervention. The duration of the intervention was 16 weeks, consisting of group training (once a week, 90 min/session, total 24 h) and self-training at home with a tablet computer software (target: three sessions/week, 30 min/session, total 24 h). Each group training session comprised group-based singing (60 min) and adapted group-level MIT (30 min) [see (Siponkoski et al., 2022) for details]. The sessions were arranged at the Aphasia Centre of the Helsinki-Uusimaa Stroke Association and were implemented by a music therapist and choir conductor team. Each patient had the opportunity to invite one caregiver or family member to participate in the sessions with them. Group-based singing included breathing and vocal exercises, vocal improvisation, and group singing with choral elements of 10 songs that were selected and arranged to be suitable for the patients. Group-level MIT comprised production of formulaic phrases utilizing the elements of MIT: intoning the phrases with simple melodic structure, tapping the rhythm with the left hand, and progressing hierarchically from unison production to repetition and from singing to spoken prosody (Schlaug et al., 2009). Home training sessions included self-training with a tablet computer, a headset, and an application developed together with Outloud, a Finnish software company. The application included all songs rehearsed in the training sessions and enabled progressive training with two different auditory models (vocal and instrumental melody) and two different visual aids (visual-kinetic model and visual-text model, that is, seeing the mouth movements of the singer and lyrics on the screen) that could be selected separately or in any combination. The software analyzed key acoustic features of the voice to provide immediate
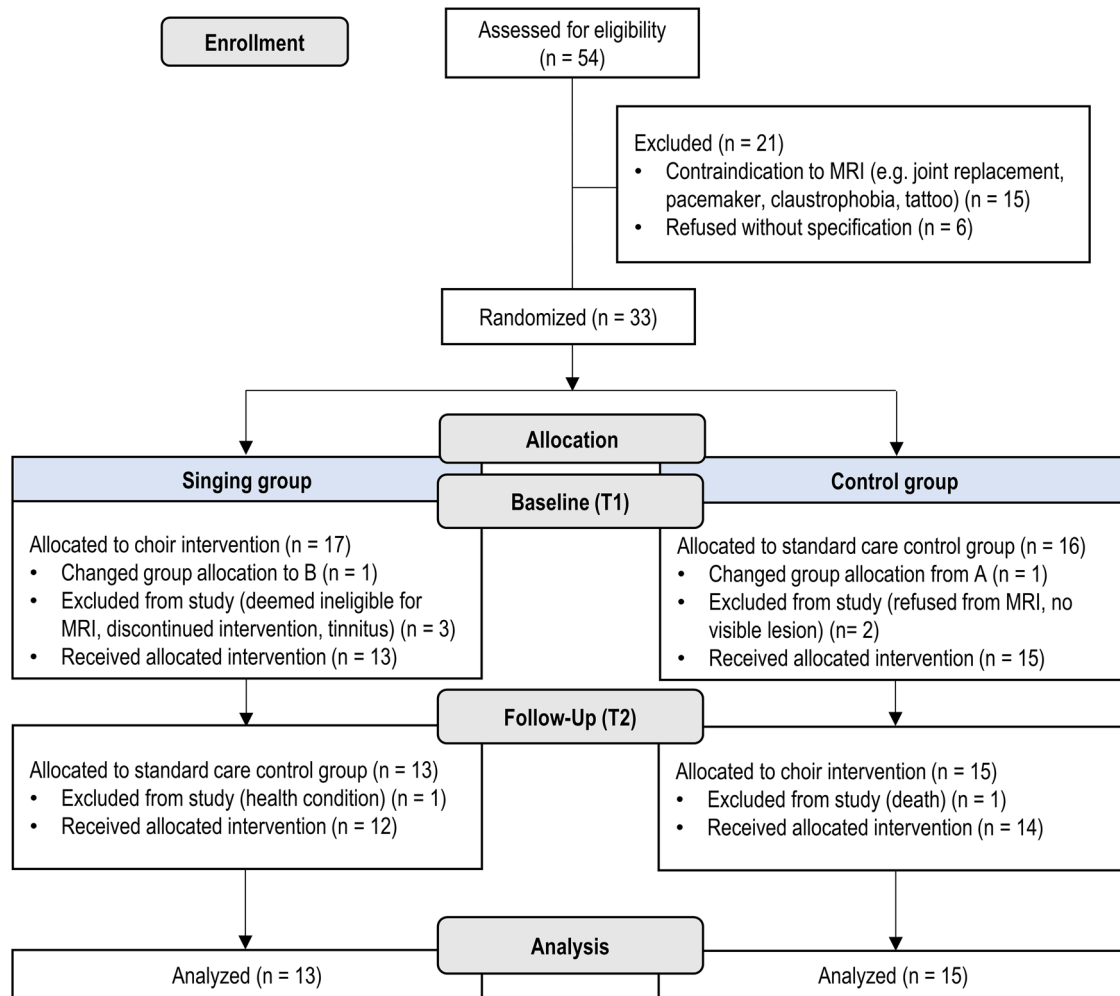
Figure 1. Flowchart.

feedback of the singing performance. The amount of home training was tracked with the app log files and was 10 h 24 min on average for the singing group (SD 6 h 51 min).

Language assessment. The language assessment was conducted by trained psychologists for each patient at all time points, blinded to the group allocation of the participants. At baseline, WAB Aphasia Quotient (Kertesz, 2007), indicating the overall severity level of the aphasia, was calculated from the Spontaneous speech, Repetition, Naming, and Comprehension (estimated based on the Sequential commands subtest) indices. Here, based on the primary spoken language production outcome measure used in the study (Siponkoski et al., 2022), we focused a priori on naming and repetition (as indicated by WAB Naming and Repetition indices; Kertesz, 1982) as the primary aspects of test-assessed language functioning improved by singing intervention due to their strong neurobiological foundation (Geva et al., 2012; Hickok, 2012; Alyahya et al., 2020; Døli et al., 2021), providing commonly used and quantifiable measures of language function that can serve as feasible counterparts of structural brain changes in PSA, thereby also reducing dimensionality in analysis. Additionally, to control for possible group differences at baseline, we evaluated motor speech production (apraxia of speech) using the articulatory agility subtest of BDAE (Goodglass and Kaplan, 1983).

MRI data acquisition and preprocessing. Patients were scanned on a 3T Siemens Skyra scanner at the Department of Radiology of the Helsinki University Hospital. For each patient, high-resolution T1-weighted anatomical images (TR = 1,800 ms; TE = 2.27 ms; TI = 900 ms; field of view = 250×250 mm; voxel size = $1 \times 0.98 \times 0.98$ mm^3) and multishell DW-MRI scan (TR = 5,000 ms, TE = 104 ms, field of view = 240×240 mm, voxel size = $2 \times 2 \times 2$ mm^3, directions = 142, b-max = 2,500 s/mm^2) with 13 nondiffusion-weighted volume and 130 diffusion-weighted volumes (30 volumes with $b = 1,000$ s/mm^2 and 100 volumes with $b = 2,500$ s/mm^2) were acquired.

Table 1. Demographic and clinical characteristics of the patients

	All $N = 28$	Singing group $N = 13$	Control group $N = 15$	p-value
Demographical				
Age (years)	64.7 (8.4)	64.1 (8.8)	65.2 (8.3)	0.731 (t)
Sex (male/female)	14/14	7/6	7/8	0.705 (χ^2)
Handedness (right/left)	23/5	12/1	11/4	0.191 (χ^2)
Education level[a]	3.0 (1.3)	3.2 (1.4)	2.9 (1.3)	0.492 (U)
Music background				
Choir singing years	3.3 (9.3)	3.8 (12.4)	2.8 (6.0)	0.389 (U)
Singing lessons years	0.6 (3.5)	0.0 (0.0)	1.2 (4.6)	0.352 (U)
Instrument lessons years	0.8 (2.3)	0.2 (0.6)	1.3 (3.0)	0.325 (U)
Clinical				
Time from injury (months)	78.1 (82.0)	69.7 (77.9)	85.4 (87.3)	0.982 (U)
Lesion size (cm^3)	89.3 (85.3)	119.0 (75.6)	63.5 (87.1)	**0.023 (U)**
Type of stroke (ischemic/hemorrhagic)	20/8	11/2	9/6	0.150 (χ^2)
AQ score (mild or moderate/severe)[b]	20/8	8/5	12/3	0.281 (χ^2)
BDAE verbal agility score	7.7 (4.8)	6.5 (5.2)	8.8 (4.2)	0.339 (U)
Received rehabilitation (T1–T2)				
Speech therapy[c]	7.3 (10.8)	7.6 (11.6)	6.9 (10.4)	0.717 (U)
Physical therapy[c]	6.9 (12.3)	8.5 (15.7)	5.5 (8.8)	1.000 (U)
Occupational therapy[c]	2.2 (5.7)	3.9 (8.0)	0.7 (1.5)	0.525 (U)
Neuropsychological rehabilitation[c]	0.3 (1.2)	0.0 (0.0)	0.6 (1.6)	0.555 (U)

Data are mean (SD) unless otherwise reported. Bold values denote statistical significance at $p < 0.05$. t, t test; χ^2, chi-square test; U, Mann–Whitney U test; AQ, aphasia quotient; BDAE, Boston Diagnostic Aphasia Examination.
[a]Education level according to the UNESCO International Standard Classification of Education: range 1 (primary education) to 6 (doctoral or equivalent level).
[b]Aphasia severity based on the AQ score: 0–50 = severe, 51–100 = mild/moderate.
[c]Data are mean (SD) in hours between T1 and T2.

MRI data were preprocessed using the Statistical Parametric Mapping software (SPM12, Wellcome Department of Cognitive Neurology, UCL, www.fil.ion.ucl.ac.uk/spm/) under MATLAB 9.4.0. To achieve optimal normalization of MRI images containing stroke lesions, cost function masking (CFM) was applied (Brett et al., 2001). This exact approach has been widely used in stroke patients (Crinion et al., 2007; Ripollés et al., 2012; Sihvonen et al., 2020) and prevents postregistration lesion shrinkage and out-of-brain distortions. The CFMs were defined separately in each time point by creating precise binary masks of the lesions by manually delineating them to the individual T1 images slice-by-slice by authors A.J.S. and N.M-M. using the MRIcron software package (http://people.cas.sc.edu/rorden/mricron/index.htm). Lesion masks were verified by a neuroradiologist (Fig. 2). T1 images and lesion masks were then reoriented according to the anterior commissure and segmented using unified segmentation (Ashburner and Friston, 2005) with medium regularization and SPM12 IXI tissue probability maps. Individual lesion maps were used to apply CFM during the preprocessing. Due to large lesion sizes, damaged voxels were masked out to achieve accurate segmentation and spatial normalization. The segmented GM probability maps were then modulated, resampled to $2 \times 2 \times 2\ mm^3$ voxel size, and normalized to Montreal Neurological Institution (MNI) space, together with the binarized lesion maps. Residual interindividual variability was reduced by smoothing the GM probability maps using an isotropic spatial filter (FWHM = 6 mm). Lastly, the segmented GM probability maps were visually inspected for segmentation errors and distortions to ensure optimal segmentation.

DW-MRI data preprocessing and reconstruction. First, the DW-MRI data were denoised for thermal noise with the MP-PCA method (Veraart et al., 2016) using a denoise tool from MRTrix3 (https://www.mrtrix.org/; Tournier et al., 2019) and corrected for Gibbs ringing based on local subvoxel shifts (Kellner et al., 2016). The b-table was checked by an automatic quality control routine to ensure its accuracy (Schilling et al., 2019). The DW-MRI data were reconstructed in the MNI space using DSI Studio (http://dsi-studio.labsolver.org, version April 7, 2021) and q-space diffeomorphic reconstruction (Yeh and Tseng, 2011) that allows the construction of spin distribution functions (Yeh et al., 2010). Normalization to the MNI space provides a direct way to analyze, for example, group differences. Normalization was carried out using the anisotropy map of each participant, and a diffusion sampling length ratio of 1.25 was used. The data output was resampled to 2 mm isotropic resolution. The quality of the normalization was inspected using the R^2 values denoting goodness of fit ($R^2 > 0.6$) between the participant's anisotropy map and template. Furthermore, each participant's forceps major and minor were inspected and used as an anatomical benchmark to confirm the normalization quality (Hula et al., 2020). The restricted diffusion was quantified using restricted diffusion imaging (Yeh et al., 2017) and quantitative anisotropy (QA) was extracted as the local connectome fingerprint (Yeh et al., 2016b) and used in the connectometry analysis. QA has been shown to outperform traditional fractional anisotropy (FA) by being more specific to individual's connectivity patterns (Yeh et al., 2016b) and less susceptible to the partial volume effect of crossing fibers and free water as well as to provide better resolution in tractography (Yeh et al., 2013).
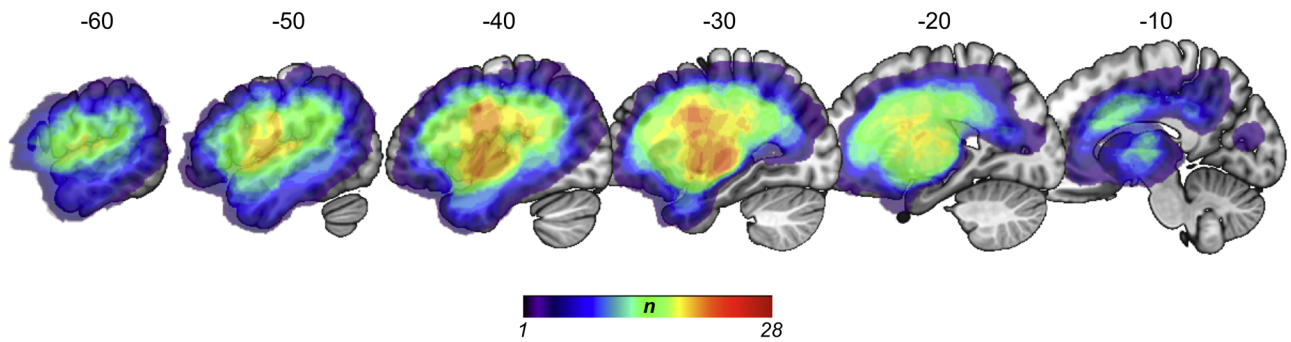
Figure 2. Lesion overlap map of all patients ($N = 28$). The n-value represents the number of patients with a lesion to a specific voxel.

Regions of interest. As we expected the intervention to induce treatment-related and activity-dependent neuroplasticity effects in brain regions activated by the singing intervention (Murphy and Corbett, 2009), we focused the analyses on neural structures related to both singing and aphasia recovery. To do this, a probabilistic human brain atlas, Brainnetome atlas (https://atlas.brainnetome.org/; Fan et al., 2016), explicitly accommodating intersubject variability in anatomy, was used to define the ROIs. Four regions from each hemisphere were derived from the Brainnetome atlas: Brodmann area (BA) 44, BA45, ventral premotor cortex (vPMC), and posterior middle temporal gyrus (pMTG). All of these areas have been implicated in singing [BA44/45 (Kleber et al., 2013; Zarate, 2013; Marchina et al., 2023), vPMC (Callan et al., 2006; Kleber et al., 2013; Marchina et al., 2023), pMTG (Whitehead and Armony, 2018; Marchina et al., 2023)] or in supporting recovery in PSA [BA44/45 (Crinion and Price, 2005; Turkeltaub et al., 2011; Hartwigsen and Saur, 2019; Stefaniak et al., 2021), vPMC (Saur et al., 2006; Seghier et al., 2014a), pMTG (Crinion and Price, 2005; Griffis et al., 2017)]. BA44 and BA45 (together known as Broca's area) were investigated separately given their established differentiation in function (Gough et al., 2005) and connectivity (Anwander et al., 2007).

Connectometry analysis. Connectometry (Yeh et al., 2016a) analyses were carried out using DSI Studio (http://dsi-studio.labsolver.org, version April 7, 2021). Connectometry is a reasonably new statistical method that includes mapping and analysis of local connectomes, that is, the degree of connectivity between adjacent voxels within a WM tract defined by the density of the diffusing spins. As a result, connectometry identifies the segments of WM fiber bundles that exhibit significant association with the study variable, here group over time. Unlike traditional FA-based structural connectome analyses, which identify differences in the mean values for the whole WM tract or using voxel-based FA values, connectometry uses QA, a measure based on the diffusion orientation distribution function (ODF), to track only the segment of the fiber bundle that exhibits significant association with the study variable or group difference. To do this, DW-MRI data are reconstructed into a standard template space (MNI) onto a local connectome matrix from the studied sample. Study-relevant variables or group information are then associated with this local connectome matrix to identify local connectomes expressing significant associations with the variable of interest. Using diffusion ODF-based measure (QA) for resolving multiple fibers, these local connectomes are then tracked along the core pathway of a fiber bundle using a fiber tracking algorithm within the Human Connectome Project tractography atlas (HCP-1065) based on 1,065 subjects and compared with a null distribution of coherent associations using permutation statistics. In summary, connectometry analyzes significant QA associations with a variable of interest or QA differences between two groups along the pathways themselves as compared with mean FA in a voxel or representing a whole tract. The analysis then outputs the significant segments of the connectome and tracts that were significantly associated between the group difference and the study variable. As the DW-MRI data are reconstructed into standard space and tracking is based on template, it also minimizes bias and variability induced by manual tracking in which, for example, slightly increasing the size of ROIs used to dissect tracts drastically changes the resulting number of streamlines and the volume they occupy, inducing significant variability within protocols and across subjects (Rheault et al., 2020; Schilling et al., 2021). The minimum length is set by voxel threshold (here 20 voxels).

A statistical model utilizing nonparametric Spearman rank-based correlation was built to consider the nonlinear effect of the group (singing group vs control group, control group vs singing group) and the longitudinal change of QA. In other words, the model compared the longitudinal (T2 > T1) significant QA changes across the structural connectome between the groups to evaluate possible treatment-related neuroplasticity changes and whether they were larger in the singing group or in the control group. Due to the group difference observed in lesion size, it was included as a nuisance variable in the analyses (Table 1). The eight selected ROIs were used as seeding regions. Local connectomes with T-score exceeding 2 were selected (Hula et al., 2020) and tracked using a deterministic fiber tracking algorithm (Yeh et al., 2013) to obtain correlational tractography. The tracks were filtered by topology-informed pruning (Yeh et al., 2019) with four iterations, and a length threshold of 20-voxel distance was used to identify significant tracts. Bootstrap resampling with 10,000 randomized permutations was used to obtain the null distribution of the track length and estimate the false discovery rates (FDRs). The alpha level was set to $p_{FDR} < 0.05$.

GM neuroplasticity analysis. The GM volume in each of the eight ROIs in two time points (T1 and T2) was extracted for all participants using SPM12 and exported to SPSS (IBM SPSS Statistics for Windows, v.27.0.: IBM, https://www.ibm.com/products/spss-statistics). A multivariate ANOVA across groups (the independent variable), including age, sex, education, total intracranial volume, and lesion size as nuisance variables (Barnes et al., 2010; Table 1), was calculated for the GM volume change (the dependent variable) in T2 > T1 (treatment-related neuroplasticity changes vs control). The total intracranial volume and lesion size did not correlate significantly ($r = 0.163$; $p = 0.407$). FDR correction was applied to control for multiple comparisons ($N = 8$), and the alpha level was set to $p_{FDR} < 0.05$.

Additional voxel-based morphometry (VBM) analysis was performed using SPM12 (Wellcome Department of Cognitive Neurology, UCL, www.fil.ion.ucl.ac.uk/spm/) under MATLAB 9.4.0 to evaluate the voxel-wise GM changes in association with the intervention. To assess the longitudinal differences in GM volume changes over time between the singing group and the control group, GM difference images were first calculated with ImCalc by subtracting each patient's GM probability map at baseline from the 5-month follow-up GM probability map. Then, the longitudinal individual preprocessed GM images were submitted to second-level independent sample t test analyses with group (singing group and control group) as factor and age, sex, education, time from stroke, total intracranial volume, and lesion size as nuisance variables. Three different t tests were calculated: (1) an unrestricted whole-brain analysis, (2) a voxel-wise analysis focused within the language network based on meta-analysis (https://neurosynth.org/analyses/terms/language/), and (3) a voxel-wise analysis focused within the study-specific ROIs. All results were thresholded at an uncorrected $p < 0.005$ threshold at the voxel level, and standard SPM family-wise error rate (FWE) cluster-level correction based on random field theory with a $p_{FWE} < 0.05$ was used. Only clusters surviving FWE-corrected $p < 0.05$ at the cluster level are reported.

Data availability. Anonymized data reported in this manuscript are available from the corresponding author upon reasonable request and subject to approval by the appropriate regulatory committees and officials. We have reported how we determined our sample size, all data exclusions (if any), all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

Results

First, by using two separate univariate ANOVAs, we confirmed that the singing and control groups did not differ in baseline WAB Naming ($p_{FDR} = 0.585$) and WAB Repetition ($p_{FDR} = 0.585$) scores (Table 2). Then, using two separate univariate ANOVAs, we evaluated whether the singing intervention was associated with language function improvements in the current sample patients with chronic PSA from the original trial (Siponkoski et al., 2022). Due to the group difference (Table 1), the analysis was adjusted for lesion size. The ANOVAs showed that the singing group significantly improved in WAB Naming [$F_{(1,25)} = 10.98$; $p_{FDR} = 0.006$; $\eta_p^2 = 0.305$] but not in WAB Repetition [$F_{(1,25)} = 1.24$; $p_{FDR} = 0.277$; $\eta_p^2 = 0.047$] compared with the control group between T1 and T2 (ΔT2–T1).

Intervention-induced WM neuroplasticity changes

First, using a nonparametric Spearman correlational model, we confirmed that the singing and control groups did not differ in baseline QA in any WM tract ($p_{FDR} = 0.140–1.000$).

Compared to the control group, the singing group showed greater increase in QA in the left AF ($p_{FDR} = 0.005$), FAT ($p_{FDR} = 0.004$), superior longitudinal fasciculus (SLF: $p_{FDR} = 0.005$), and corticostriatal tract ($p_{FDR} = 0.004$), and in the right FAT ($p_{FDR} = 0.004$), SLF ($p_{FDR} = 0.004$) and corticostriatal tract ($p_{FDR} = 0.004$) as well as in the the corpus callosum ($p_{FDR} = 0.005$) in ΔT2–T1 (Fig. 3).

Intervention-induced GM neuroplasticity changes

First, using separate univariate ANOVAs, we confirmed that the singing and control groups did not differ in baseline GM volume in any of the chosen ROIs ($p_{FDR} = 0.822–0.867$).

In the multivariate ANOVA evaluating treatment-related GM changes (ΔT2–T1), there was a statistically significant difference in the longitudinal change in the GM volume between the groups [$F_{(8,19)} = 3.759$; $p = 0.020$; Wilk's $\Lambda = 0.285$; $\eta_p^2 = 0.715$]. Therefore, eight separate univariate ANOVAs (i.e., one for each ROI) were performed. These revealed that the GM volume change (ΔT2–T1) of the left BA44 [$F_{(1,19)} = 20.785$, $p_{FDR} = 0.008$; $\eta_p^2 = 0.522$; singing group$_{mean} = 61$ mm^3, control group$_{mean} = -140$ mm^3], left BA45 [$F_{(1,19)} = 9.772$, $p_{FDR} = 0.024$; $\eta_p^2 = 0.340$; singing group$_{mean} = 28$ mm^3, control group$_{mean} = -73$ mm^3], and left vPMC [$F_{(1,19)} = 8.607$, $p_{FDR} = 0.024$; $\eta_p^2 = 0.312$; singing group$_{mean} = 57$ mm^3, control group$_{mean} = -195$ mm^3] showed greater increase in the singing group than in the control group (A > B; Fig. 4). ANOVAs for the other ROIs were not significant after FDR correction.

In addition to the a priori defined ROI-based analysis, longitudinal treatment-related GM volume changes were also assessed using three different voxel-wise analyses: (1) an unrestricted whole-brain analysis, (2) a voxel-wise analysis focused within the language network based on meta-analysis (https://neurosynth.org/analyses/terms/language/), and (3) a voxel-wise analysis focused within the study-specific ROIs. Similar to the ROI-based analysis, age, sex, education, total intracranial volume, and lesion size were added as nuisance variables. In all three longitudinal VBM analyses, the GM

Table 2. Spoken language and neuroplasticity analysis outcomes

Measure	Group	T1 mean (SD)	T2 mean (SD)	ΔT2 > T1 mean (SD)	Baseline diff. (p_{FDR}-value)	ΔT2 > T1 (p_{FDR}-value)
Spoken language outcomes						
WAB Naming	Singing	5.0 (3.6)	5.7 (3.8)	0.7 (0.7)	0.585	**0.006**
	Control	7.4 (3.6)	7.4 (3.5)	0.0 (0.5)		
WAB Repetition	Singing	5.4 (3.6)	5.7 (3.4)	0.3 (0.7)	0.585	0.277
	Control	7.6 (3.4)	7.6 (3.3)	0.0 (0.4)		
Neuroplasticity outcomes						
GMV left BA44[a]	Singing	747 (568)	807 (586)	61 (108)	0.945	**0.008**
	Control	854 (387)	713 (427)	−140 (141)		
GMV right BA44[a]	Singing	1,619 (159)	1,628 (278)	9 (99)	0.718	0.587
	Control	1,798 (336)	1,743 (324)	−55 (149)		
GMV left BA45[a]	Singing	387 (304)	414 (306)	28 (43)	0.945	**0.024**
	Control	409 (240)	336 (263)	−73 (78)		
GMV right BA45[a]	Singing	873 (159)	891 (173)	19 (61)	0.720	0.759
	Control	924 (201)	913 (164)	−11 (110)		
GMV left vPMC[a]	Singing	884 (567)	941 (607)	57 (203)	0.718	**0.024**
	Control	1,106 (400)	913 (460)	−195 (151)		
GMV right vPMC[a]	Singing	1,783 (293)	1,775 (309)	−8 (105)	0.833	0.876
	Control	1,827 (448)	1,810 (370)	−21 (172)		
GMV left pMTG[a]	Singing	1,103 (718)	1,090 (724)	−13 (76)	0.718	0.268
	Control	1,383 (430)	1,431 (591)	47 (56)		
GMV right pMTG[a]	Singing	2,205 (533)	2,242 (551)	37 (137)	0.718	0.777
	Control	2,498 (503)	2,513 (477)	14 (74)		
WM left AF[b]	Singing	0.10 (0.04)	0.11 (0.05)	0.01 (0.01)	0.480	**0.005**
	Control	0.11 (0.04)	0.11 (0.05)	0.00 (0.02)		
WM left FAT[b]	Singing	0.13 (0.03)	0.14 (0.04)	0.01 (0.02)	0.523	**0.004**
	Control	0.14 (0.05)	0.13 (0.06)	−0.01 (0.02)		
WM right FAT[b]	Singing	0.19 (0.03)	0.20 (0.03)	0.01 (0.01)	0.815	**0.004**
	Control	0.20 (0.03)	0.19 (0.02)	0.00 (0.02)		
WM left SLF[b]	Singing	0.10 (0.05)	0.11 (0.05)	0.01 (0.01)	0.480	**0.005**
	Control	0.11 (0.04)	0.11 (0.05)	0.00 (0.02)		
WM right SLF[b]	Singing	0.21 (0.03)	0.22 (0.03)	0.01 (0.01)	0.480	**0.004**
	Control	0.20 (0.03)	0.20 (0.03)	0.00 (0.01)		
WM left CS[b]	Singing	0.09 (0.04)	0.10 (0.04)	0.01 (0.01)	0.480	**0.004**
	Control	0.10 (0.04)	0.10 (0.05)	0.00 (0.01)		
WM right CS[b]	Singing	0.16 (0.04)	0.17 (0.04)	0.01 (0.01)	0.555	**0.004**
	Control	0.17 (0.03)	0.17 (0.02)	0.00 (0.01)		
WM CC[b]	Singing	0.14 (0.03)	0.15 (0.03)	0.01 (0.01)	0.480	**0.005**
	Control	0.15 (0.03)	0.15 (0.03)	−0.01 (0.01)		

Data are mean (SD) unless otherwise reported. Bold values denote statistical significance at $p < 0.05$. P-values are FDR corrected.
[a]Volume in mm^3.
[b]Normalized QA.

volume increased more in the singing group than in the control group in one cluster comprising the left BA44 and vPMC (peak MNI coordinate = −56, 16, 18; Extended Data Figure 4-1). Both within language network (cluster size = 512 voxels, T-value = 6.43, p_{FWE} = 0.038) and within study-specific ROIs (cluster size = 167 voxels, T-value = 6.45, p_{FWE} = 0.041) analyses were statistically significant, but the GM volume change failed to reach statistical significance in the unrestricted whole-brain analysis (cluster size = 366 voxels; T-value = 5.96; p_{FWE} = 0.263).

Brain–behavior relationships

To investigate treatment-induced brain–behavior relationships, we first evaluated the longitudinal QA change associations with improved naming, as the functional restoration in PSA mostly relies upon the structural remodeling of the injured networks (Stefaniak et al., 2020). Therefore, a nonparametric Spearman correlational model (T2 > T1) was built to derive the correlational tractography within the significant WM findings evaluating the longitudinal change of QA correlated with the improvement in WAB Naming. In ΔT2 > T1, increased QA in the left FAT (p_{FDR} = 0.009) and AF (p_{FDR} = 0.016) correlated with improved naming (Fig. 3).

Evaluation of the relationship between improved naming scores and GM volume changes were restricted to the left BA44 to which both left AF and FAT are frontally terminated (Thiebaut de Schotten et al., 2012). Partial correlation (Pearson's, one-tailed, controlling for the same nuisance variables as in the initial GM volume analysis, i.e., age, sex, education, total intracranial volume, and lesion size) using SPSS showed that increased GM volume (ΔT2–T1) in the left BA44 correlated with improved naming abilities ($r = 0.371$; $p = 0.044$; Fig. 4).
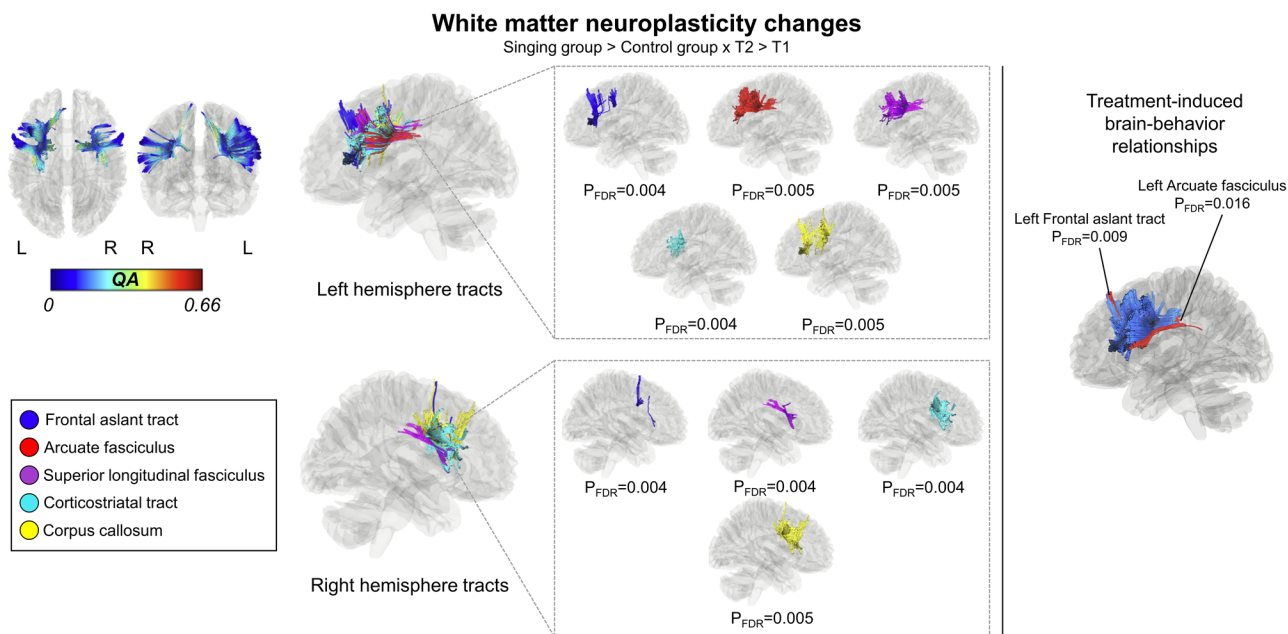
Figure 3. Treatment-induced WM neuroplasticity changes. Connectometry results displaying the significant segments of the tracts with longitudinal QA increases significantly associated with singing group versus control group between T1 and T2 (ΔT2–T1; left) and longitudinal QA change correlation with improved naming (right). FDR, false discovery rate; L, left; QA, quantitative anisotropy; R, right.
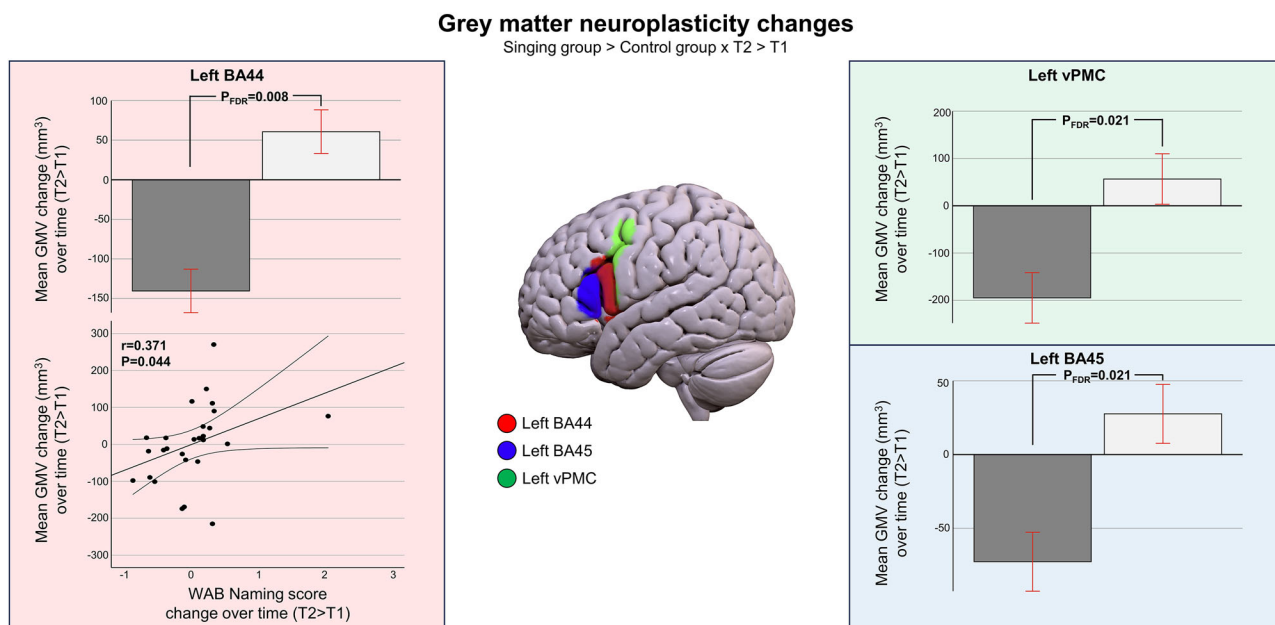

Figure 4. Treatment-induced GM neuroplasticity changes. Longitudinal GM volume increases (singing group > control group) in T2 > T1 and longitudinal GM volume change correlation with improved naming. Additional exploratory voxel-wise analyses are reported in Extended Data Figure 4-1. Bar plots for mean group GM volume changes are shown: bar, mean; error bar, standard error of mean. BA, Brodmann area; FDR, false discovery rate; vPMC, ventral premotor cortex.

Discussion

This study set out to determine the structural GM and WM benefits of group-based singing in PSA rehabilitation and their correlation to improved language outcome. Our novel main findings were that group-based singing enhanced structural WM connectivity in both the left and right hemispheres within the language network and the GM volume in the left language-related frontal areas compared with the control group. The left frontal neuroplasticity effects correlated with improved naming abilities. The present study provides the first evidence on the neural benefits of group-based singing

that supports language recovery in PSA and extends previous results on the effects of music-based interventions in stroke rehabilitation, including MIT (Schlaug et al., 2009; Marchina et al., 2023) and music listening (Sihvonen et al., 2017a, 2020, 2021b). These results are important in (1) providing evidence of treatment-induced structural changes in chronic PSA, (2) improving our understanding of chronic PSA rehabilitation, and (3) determining targets and mediators of music-based rehabilitation strategies (Sihvonen et al., 2017a).

In aphasia, recovery relies on the ability of the spared neurons to remodel the injured network (Kiran et al., 2019). The recovery processes in PSA exploit activity-dependent neuroplasticity mechanisms (Murphy and Corbett, 2009) within the language network (Stefaniak et al., 2020, 2021), that is, increased stimulation through iterative utilization of language processes, and therefore activating the language network supports the recovering brain by increasing, for example, dendritic spine density and neurotrophic factor levels (Carmichael et al., 2001; Nithianantharajah and Hannan, 2006). In the context of neural stimulation, music-based interventions, such as singing, are feasible tools to promote language network recovery in PSA (Särkämö and Sihvonen, 2018). First, simply listening to vocal music has been demonstrated to activate language-related regions of the brain, even in the case of acute stroke (Sihvonen et al., 2017b). Moreover, daily listening to vocal music poststroke has been linked to improved language recovery in PSA, evidenced by improved language skills and verbal memory and increased GM volume within the left temporal regions (Sihvonen et al., 2020). The vocal music listening intervention also enhances the functional connectivity within the language and default mode networks (Sihvonen et al., 2020, 2021a), and strengthens the structural connectivity of the left FAT and stimulus-specific activation of its superior frontal termination areas compared with audiobook listening (Sihvonen et al., 2021b). Compared to mere vocal music listening, choir singing should be superior in administering neural stimulation and providing more fertile environment for recovery as it incorporates multiple elements such as the production of words through singing, physiological effects of singing, the experience of singing with others, the perception of sung music, social interaction, and the learning of new songs and lyrics. In theory, these combined factors create a more conducive environment for neural stimulation and recovery (Murphy and Corbett, 2009).

Second, singing and speech share core neuronal circuitry within the left hemisphere (Pitkäniemi et al., 2023). Singing also binds linguistic and musical information into a unified representation and naturally increases connectedness between syllables and words and, in this respect, resembles connected spoken language production. Compared to speaking, singing engages bilateral language-related frontotemporal areas more extensively (Callan et al., 2006; Schön et al., 2010) and requires multiple neural circuits to operate in concert (Loui, 2015). The classical notion in neurology is that even patients with severe PSA can retain the ability to sing lyrics of familiar songs (Johnson and Graziano, 2015). Sung information is also accessible to patients with PSA who have been shown to repeat and recall more words when singing than when speaking (Racette et al., 2006; Leo et al., 2018). This evidence suggests that singing can provide an avenue for language rehabilitation in PSA. Indeed, singing-based interventions such as MIT have been shown to improve language recovery in non-fluent PSA (Sparks et al., 1974; Van Der Meulen et al., 2014; Zumbansen et al., 2014) with associated positive neuroplasticity effects reported in language-related frontotemporal areas bilaterally (Belin et al., 1996; Schlaug et al., 2008; Breier et al., 2011; Wan et al., 2014; Tabei et al., 2016).

In accordance with the abovementioned, our current findings revealed that group-based singing enhanced WM connectivity in both hemispheres, but with left-hemispheric dominance. Treatment-related changes correlating with improved naming abilities comprised the left AF and FAT, damage of which has been associated with speech production outcome in PSA (Fridriksson et al., 2013; Alyahya et al., 2020). Changes were also observed in the corpus callosum, SLF, and corticostriatal tract bilaterally. These observations might lend to two distinct mechanisms, that is, treatment-related changes within the language network and in the shared structures between the language and domain-general networks. According to the neurocomputational model of PSA recovery, initially damaged AF/SLF undergoes plasticity-related changes during the recovery (Stefaniak et al., 2020). The corpus callosum has been shown to play a critical role in language comprehension in integrating prosodic and syntactic information (Sammler et al., 2010), and its treatment-related changes after singing-based treatments, combining melody, rhythm, and linguistic information, are reasonable findings. In contrast, treatment-related plasticity changes in the corticostriatal tracts might reflect more domain-general network effects as corticostriatal systems have been shown to play a domain-general regulatory role in language operations (Copland et al., 2021). Moreover, the proposed neuroanatomical model supporting singing center on the left AF/SLF but also includes the left FAT as well as ventral tracts (Loui, 2015). The present findings conform with this model and the previous neuroanatomical evidence on the core neuronal circuitry underpinning singing of words in aphasia (Pitkäniemi et al., 2023) as well as with the previous small-scale PSA treatment-induced WM findings (Schlaug et al., 2009; Breier et al., 2011; Allendorfer et al., 2012; Van Hees et al., 2014).

Group-based singing-induced GM plasticity changes that correlated with improved naming abilities were observed in the left BA44 where the left AF/SLF and FAT cortically terminate (Thiebaut de Schotten et al., 2012). The intervention group showed longitudinally slightly increased GM volume in the left BA44, whereas the control group showed a decline in that area. This most likely owes to the accelerated brain atrophy rate after stroke, which is 2–4 times greater than in healthy controls (Brodtmann et al., 2020, 2021; Salah Khlif et al., 2022). Lesioned areas have been shown to lead to further neuronal decay, even in the chronic poststroke stage, with a median rate of 1,590 mm^3 per year (Seghier et al., 2014b). In comparison, the mean left BA44 GM volume decrease in the control group in the present study was 140 mm^3 in 5 months. Similarly, WM neurodegeneration in ipsi- and contralesional tracts continues to be greater in

stroke survivors compared with the healthy population (Egorova-Brumley et al., 2023). In a recent study on patients with PSA, most patients showed evidence of lesion expansion and that it was associated with further declining language performance (Johnson et al., 2023). The poststroke brain atrophy rates may serve as biomarkers reflecting treatment response for interventions to reduce poststroke secondary degeneration and vascular cognitive impairment (Brodtmann et al., 2021). For these reasons, interventions, including singing-based ones, might not only increase GM volume but also prevent further brain atrophy in PSA. However, future longitudinal studies with larger sample sizes are needed to elucidate whether the improved functional outcomes are underpinned by possible neuroprotective neuroplasticity changes that prevent GM atrophy and WM neurodegeneration poststroke, if not increase GM volumes and WM structural connectivity.

The observed relationship between the treatment-induced improvement of naming and neuroplasticity in the left frontal GM and WM is well in line with the classic models of word production (Indefrey and Levelt, 2004; Hickok, 2012) in which the speech motor processes of syllabification, phonetic encoding, and articulation are attributed to largely to these regions and pathways. The singing intervention may support this process by slowing the rate of word production and increasing the connectedness between syllables/words through continuous voicing and melodic intonation (Wan et al., 2010). Moreover, treatment-related modulation of the left-hemispheric cortical activity in Broca's area and the premotor cortex has been associated with improved naming in PSA (Fridriksson, 2010).

The present study has some potential limitations. First, the singing intervention comprised multiple components, and differentiating between the efficacies of individual treatment components is not possible based on the current data. Second, the sample size is relatively small and lacked an active control matched for dose of the intervention, limiting the generalizability of the findings. The sample size also affected the sensitivity of the additional voxel-wise analyses, where both focused analyses provided statistically significant results paralleling the a priori ROI-based analysis, but the unrestricted whole-brain analysis did not. Yet, the results from all the three voxel-wise analyses were similar and conformed with the ROI analysis. However, the beneficial effects of singing-based interventions might not be restricted to the left frontal regions, and future studies with larger sample sizes utilizing whole-brain analyses are warranted. Yet, this study is the largest multimodal RCT to date on treatment-induced neuroplasticity changes in PSA. While the present results need to be replicated in future larger studies, they are encouraging in providing us evidence of health economically promising multifaceted PSA treatment bringing about beneficial neuroplasticity change.

In conclusion, the present results suggest that the positive effects of singing on chronic PSA recovery are underpinned by structural GM and WM reorganization, mainly within the left frontal areas. Clinically, together with previous behavioral results on positive effects of singing in chronic PSA (Siponkoski et al., 2022), this evidence suggests that group-based singing is a feasible tool to promote language network reorganization and recovery in PSA.

References

Allendorfer JB, Storrs JM, Szaflarski JP (2012) Changes in white matter integrity follow excitatory rTMS treatment of post-stroke aphasia. Restor Neurol Neurosci 30:103–113.

Alyahya RSW, Halai AD, Conroy P, Lambon MA (2020) A unified model of post-stroke language deficits including discourse production and their neural correlates. Brain 143:1541–1554.

Anwander A, Tittgemeyer M, Von Cramon DY, Friederici AD, Knösche TR (2007) Connectivity-based parcellation of Broca's area. Cereb Cortex 17:816–825.

Ashburner J, Friston KJ (2005) Unified segmentation. Neuroimage 26:839–851.

Barnes J, Ridgway GR, Bartlett J, Henley SMD, Lehmann M, Hobbs N, Clarkson MJ, MacManus DG, Ourselin S, Fox NC (2010) Head size, age and gender adjustment in MRI studies: a necessary nuisance? Neuroimage 53:1244–1255.

Baroncelli L, Braschi C, Spolidoro M, Begenisic T, Sale A, Maffei L (2010) Nurturing brain plasticity: impact of environmental enrichment. Cell Death Differ 17:1092–1103.

Belin P, Van Eeckhout P, Zilbovicius M, Remy P, François C, Guillaume S, Chain F, Rancurel G, Samson Y (1996) Recovery from nonfluent aphasia after melodic intonation therapy: a PET study. Neurology 47:1504–1511.

Breier JI, Juranek J, Papanicolaou AC (2011) Changes in maps of language function and the integrity of the arcuate fasciculus after therapy for chronic aphasia. Neurocase 17:506–517.

Brett M, Leff AP, Rorden C, Ashburner J (2001) Spatial normalization of brain images with focal lesions using cost function masking. Neuroimage 14:486–500.

Brodtmann A, et al. (2021) Neurodegeneration over 3 years following ischaemic stroke: findings from the cognition and neocortical volume after stroke study. Front Neurol 12:754204.

Brodtmann A, Khlif MS, Egorova N, Veldsman M, Bird LJ, Werden E (2020) Dynamic regional brain atrophy rates in the first year after ischemic stroke. Stroke 51:183–192.

Callan DE, Tsytsarev V, Hanakawa T, Callan AM, Katsuhara M, Fukuyama H, Turner R (2006) Song and speech: brain regions involved with perception and covert production. Neuroimage 31:1327–1342.

Carmichael ST, Wei L, Rovainen CM, Woolsey TA (2001) New patterns of intracortical projections after focal cortical stroke. Neurobiol Dis 8:910–922.

Copland DA, Brownsett S, Iyer K, Angwin AJ (2021) Corticostriatal regulation of language functions. Neuropsychol Rev 31:472–494.

Crinion J, Price CJ (2005) Right anterior superior temporal activation predicts auditory sentence comprehension following aphasic stroke. Brain 128:2858–2871.

Crinion J, Ashburner J, Leff A, Brett M, Price C, Friston K (2007) Spatial normalization of lesioned brains: performance evaluation and impact on fMRI analyses. Neuroimage 37:866–875.

Dickey L, Kagan A, Lindsay MP, Fang J, Rowland A, Black S (2010) Incidence and profile of inpatient stroke-induced aphasia in Ontario, Canada. Arch Phys Med Rehabil 91:196–202.

Døli H, Andersen Helland W, Helland T, Specht K (2021) Associations between lesion size, lesion location and aphasia in acute stroke. Aphasiology 35:745–763.

Doogan C, Dignam J, Copland D, Leff A (2018) Aphasia recovery: when, how and who to treat? Curr Neurol Neurosci Rep 18:90.

Egorova-Brumley N, Dhollander T, Khan W, Khlif MS, Ebaid D, Brodtmann A (2023) Changes in white matter microstructure over 3 years in people with and without stroke. Neurology 100: E1664–E1672.

Fan L, et al. (2016) The human Brainnetome atlas: a new brain atlas based on connectional architecture. Cereb Cortex 26:3508–3526.

Fleming V, et al. (2020) Efficacy of spoken word comprehension therapy in patients with chronic aphasia: a cross-over randomised controlled trial with structural imaging. J Neurol Neurosurg Psychiatry 92:418–424.

Fridriksson J (2010) Preservation and modulation of specific left hemisphere regions is vital for treated recovery from anomia in stroke. J Neurosci 30:11558–11564.

Fridriksson J, Guo D, Fillmore P, Holland A, Rorden C (2013) Damage to the anterior arcuate fasciculus predicts non-fluent speech production in aphasia. Brain 136:3451–3460.

Geva S, Baron JC, Jones PS, Price CJ, Warburton EA (2012) A comparison of VLSM and VBM in a cohort of patients with post-stroke aphasia. NeuroImage Clin 1:37–47.

Goodglass H, Kaplan E (1983) *Boston diagnostic aphasia examination (BDAE)*, Ed 2. Philadelphia, PA, USA: Lea & Febiger.

Gough PM, Nobre AC, Devlin JT (2005) Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. J Neurosci 25:8010–8016.

Griffis JC, Nenert R, Allendorfer JB, Vannest J, Holland S, Dietz A, Szaflarski JP (2017) The canonical semantic network supports residual language function in chronic post-stroke aphasia. Hum Brain Mapp 38:1636–1658.

Hartwigsen G, Saur D (2019) Neuroimaging of stroke recovery from aphasia – insights into plasticity of the human language network. Neuroimage 190:14–31.

Hickok G (2012) Computational neuroanatomy of speech production. Nat Rev Neurosci 13:135–145.

Hilari K, Needle JJ, Harrison KL (2012) What are the important factors in health-related quality of life for people with aphasia? A systematic review. Arch Phys Med Rehabil 93:S86–S95.

Hula WD, Panesar S, Gravier ML, Yeh FC, Dresang HC, Dickey MW, Fernandez-Miranda JC (2020) Structural white matter connectometry of word production in aphasia: an observational study. Brain 143:2532–2544.

Indefrey P, Levelt WJM (2004) The spatial and temporal signatures of word production components. Cognition 92:101–144.

Johansson BB, Ohlsson AL (1996) Environment, social interaction, and physical activity as determinants of functional outcome after cerebral infarction in the rat. Exp Neurol 139:322–327.

Johnson JK, Graziano AB (2015) Some early cases of aphasia and the capacity to sing. Prog Brain Res 216:73–89.

Johnson L, Newman-Norlund R, Teghipco A, Rorden C, Bonilha L, Fridriksson J, Author C (2023) Progressive lesion necrosis is related to increasing aphasia severity in chronic stroke. medRxiv 41:2023.06.13.23291362.

Kellner E, Dhital B, Kiselev VG, Reisert M (2016) Gibbs-ringing artifact removal based on local subvoxel-shifts. Magn Reson Med 76:1574–1581.

Kertesz A (1982) *Western aphasia battery*, Ed 1. San Antonio, TX: The Psychological Corporation.

Kertesz A (2007) Western aphasia battery-revised (WAB-R).

Kiran S, Meier EL, Johnson JP (2019) Neuroplasticity in aphasia: a proposed framework of language recovery. J Speech Lang Hear Res 62:3973–3985.

Kleber B, Zeitouni AG, Friberg A, Zatorre RJ (2013) Experience-dependent modulation of feedback integration during singing: role of the right anterior insula. J Neurosci 33:6070–6080.

Leo V, Sihvonen AJ, Linnavalli T, Tervaniemi M, Laine M, Soinila S, Särkämö T (2018) Sung melody enhances verbal learning and recall after stroke. Ann N Y Acad Sci 1423:296–307.

Loui P (2015) A dual-stream neuroanatomy of singing. Music Percept 32:232–241.

Marchina S, Norton A, Schlaug G (2023) Effects of melodic intonation therapy in patients with chronic nonfluent aphasia. Ann N Y Acad Sci 1519:173–185.

McKinnon ET, Fridriksson J, Glenn GR, Jensen JH, Helpern JA, Basilakos A, Rorden C, Shih AY, Spampinato MV, Bonilha L (2017) Structural plasticity of the ventral stream and aphasia recovery. Ann Neurol 82:147–151.

Mitchell AJ, Sheth B, Gill J, Yadegarfar M, Stubbs B, Yadegarfar M, Meader N (2017) Prevalence and predictors of post-stroke mood disorders: a meta-analysis and meta-regression of depression, anxiety and adjustment disorder. Gen Hosp Psychiatry 47:48–60.

Murphy TH, Corbett D (2009) Plasticity during stroke recovery: from synapse to behaviour. Nat Rev Neurosci 10:861–872.

Nithianantharajah J, Hannan AJ (2006) Enriched environments, experience-dependent plasticity and disorders of the nervous system. Nat Rev Neurosci 7:697–709.

Olesen J, Gustavsson A, Svensson M, Wittchen HU, Jönsson B (2012) The economic cost of brain disorders in Europe. Eur J Neurol 19:155–162.

Pedersen PM, Vinter K, Olsen TS (2004) Aphasia after stroke: type, severity and prognosis: the Copenhagen aphasia study. Cerebrovasc Dis 17:35–43.

Pentikäinen E, Kimppa L, Makkonen T, Putkonen M, Pitkäniemi A, Salakka I, Paavilainen P, Tervaniemi M, Särkämö T (2022) Benefits of choir singing on complex auditory encoding in the aging brain: an ERP study. Ann N Y Acad Sci 1514:82–92.

Pentikäinen E, Pitkäniemi A, Siponkoski ST, Jansson M, Louhivuori J, Johnson JK, Paajanen T, Särkämö T (2021) Beneficial effects of choir singing on cognition and well-being of older adults: evidence from a cross-sectional study. PLoS One 16:e0245666.

Perron M, Theaud G, Descoteaux M, Tremblay P (2021) The frontotemporal organization of the arcuate fasciculus and its relationship with speech perception in young and older amateur singers and nonsingers. Hum Brain Mapp 42:3058–3076.

Perron M, Vaillancourt J, Tremblay P (2022) Amateur singing benefits speech perception in aging under certain conditions of practice: behavioural and neurobiological mechanisms. Brain Struct Funct 227:943–962.

Pitkäniemi A, Särkämö T, Siponkoski S-T, Brownsett SLE, Copland DA, Sairanen V, Sihvonen AJ (2023) Hodological organization of spoken language production and singing in the human brain. Commun Biol 6:779.

Racette A, Bard C, Peretz I (2006) Making non-fluent aphasics speak: sing along!. Brain 129:2571–2584.

Rheault F, Poulin P, Valcourt Caron A, St-Onge E, Descoteaux M (2020) Common misconceptions, hidden biases and modern challenges of dMRI tractography. J Neural Eng 17:011001.

Ripollés P, Marco-Pallares J, de Diego-Balaguer R, Miro J, Falip M, Juncadella M, Rubio F, Rodriguez-Fornells A (2012) Analysis of automated methods for spatial normalization of lesioned brains. Neuroimage 60:1296–1306.

Salah Khlif M, Egorova-Brumley N, Bird LJ, Werden E, Brodtmann A (2022) Cortical thinning 3 years after ischaemic stroke is associated with cognitive impairment and APOE ε4. NeuroImage Clin 36:103200.

Sammler D, Kotz SA, Eckstein K, Ott DVM, Friederici AD (2010) Prosody meets syntax: the role of the corpus callosum. Brain 133:2643–2655.

Särkämö T, Sihvonen AJ (2018) Golden oldies and silver brains: deficits, preservation, learning, and rehabilitation effects of music in ageing-related neurological disorders. Cortex 109:104–123.

Saur D, Lange R, Baumgaertner A, Schraknepper V, Willmes K, Rijntjes M, Weiller C (2006) Dynamics of language reorganization after stroke. Brain 129:1371–1384.

Schilling KG, Yeh FC, Nath V, Hansen C, Williams O, Resnick S, Anderson AW, Landman BA (2019) A fiber coherence index for quality control of B-table orientation in diffusion MRI scans. Magn Reson Imaging 58:82–89.

Schilling KG, et al. (2021) Tractography dissection variability: what happens when 42 groups dissect 14 white matter bundles on the same dataset? Neuroimage 243:118502.

Schlaug G, Maechina S, Norton A (2008) From singing to speaking: why singing may lead to recovery of expressive language function in patients with Broca's aphasia. Music Percept 25:315–323.

Schlaug G, Marchina S, Norton A (2009) Evidence for plasticity in white-matter tracts of patients with chronic Broca's aphasia undergoing intense intonation-based speech therapy. Ann N Y Acad Sci 1169:385–394.

Schön D, Gordon R, Campagne A, Magne C, Astésano C, Anton JL, Besson M (2010) Similar cerebral networks in language, music and song perception. Neuroimage 51:450–461.

Seghier ML, Bagdasaryan J, Jung DE, Price CJ (2014a) The importance of premotor cortex for supporting speech production after left capsular–putaminal damage. J Neurosci 34:14338–14348.

Seghier ML, Ramsden S, Lim L, Leff AP, Price CJ (2014b) Gradual lesion expansion and brain shrinkage years after stroke. Stroke 45:877–879.

Sihvonen AJ, et al. (2020) Vocal music enhances memory and language recovery after stroke: pooled results from two RCTs. Ann Clin Transl Neurol 7:2272–2287.

Sihvonen AJ, Särkämö T, Leo V, Tervaniemi M, Altenmüller E, Soinila S (2017a) Music-based interventions in neurological rehabilitation. Lancet Neurol 16:648–660.

Sihvonen AJ, Särkämö T, Ripollés P, Leo V, Saunavaara J, Parkkola R, Rodríguez-Fornells A, Soinila S (2017b) Functional neural changes associated with acquired amusia across different stages of recovery after stroke. Sci Rep 7:11390.

Sihvonen AJ, Pitkäniemi A, Leo V, Soinila S, Särkämö T (2021a) Resting-state language network neuroplasticity in post-stroke music listening: a randomized controlled trial. Eur J Neurosci 54:7886–7898.

Sihvonen AJ, Ripollés P, Leo V, Saunavaara J, Parkkola R, Rodríguez-Fornells A, Soinila S, Särkämö T (2021b) Vocal music listening enhances poststroke language network reorganization. eNeuro 8:ENEURO.0158-21.2021.

Siponkoski S-T, et al. (2022) Efficacy of a multicomponent singing intervention on communication and psychosocial functioning in chronic aphasia: a randomized controlled crossover trial. Brain Commun 5:fcac337.

Sparks R, Helm N, Albert M (1974) Aphasia rehabilitation resulting from melodic intonation therapy. Cortex 10:303–316.

Stefaniak JD, Halai AD, Lambon Ralph MA (2020) The neural and neurocomputational bases of recovery from post-stroke aphasia. Nat Rev Neurol 16:43–55.

Stefaniak JD, Alyahya RSW, Lambon Ralph MA (2021) Language networks in aphasia and health: a 1000 participant activation likelihood estimation meta-analysis. Neuroimage 233:117960.

Stockert A, Wawrzyniak M, Klingbeil J, Wrede K, Kümmerer D, Hartwigsen G, Kaller CP, Weiller C, Saur D (2020) Dynamics of language reorganization after left temporo-parietal and frontal stroke. Brain 143:844–861.

Tabei KI, Satoh M, Nakano C, Ito A, Shimoji Y, Kida H, Sakuma H, Tomimoto H (2016) Improved neural processing efficiency in a chronic aphasia patient following melodic intonation therapy: a neuropsychological and functional MRI study. Front Neurol 7:148.

Thiebaut de Schotten M, Dell'Acqua F, Valabregue R, Catani M (2012) Monkey to human comparative anatomy of the frontal lobe association tracts. Cortex 48:82–96.

Tournier JD, Smith R, Raffelt D, Tabbara R, Dhollander T, Pietsch M, Christiaens D, Jeurissen B, Yeh CH, Connelly A (2019) MRtrix3: a fast, flexible and open software framework for medical image processing and visualisation. Neuroimage 202:116137.

Turkeltaub PE, Messing S, Norise C, Hamilton RH (2011) Are networks for residual language function and recovery consistent across aphasic patients? Neurology 76:1726–1734.

Van Der Meulen I, Van De Sandt-Koenderman ME, Ribbers GM (2012) Melodic intonation therapy: present controversies and future opportunities. Arch Phys Med Rehabil 93:S46–S52.

Van Der Meulen I, Van De Sandt-Koenderman WME, Heijenbrok-Kal MH, Visch-Brink EG, Ribbers GM (2014) The efficacy and timing of melodic intonation therapy in subacute aphasia. Neurorehabil Neural Repair 28:536–544.

Van Hees S, McMahon K, Angwin A, De Zubicaray G, Read S, Copland DA (2014) Changes in white matter connectivity following therapy for anomia post stroke. Neurorehabil Neural Repair 28:325–334.

Venna VR, Xu Y, Doran SJ, Patrizz A, McCullough LD (2014) Social interaction plays a critical role in neurogenesis and recovery after stroke. Transl Psychiatry 4:e351.

Veraart J, Novikov DS, Christiaens D, Ades-aron B, Sijbers J, Fieremans E (2016) Denoising of diffusion MRI using random matrix theory. Neuroimage 142:394–406.

Vetere G, Williams G, Ballard C, Creese B, Hampshire A, Palmer A, Pickering E, Richards M, Brooker H, Corbett A (2024) The relationship between playing musical instruments and cognitive trajectories: analysis from a UK ageing cohort. Int J Geriatr Psychiatry 39:e6061.

Wan CY, Rüber T, Hohmann A, Schlaug G (2010) The therapeutic effects of singing in neurological disorders. Music Percept 27:287–295.

Wan CY, Zheng X, Marchina S, Norton A, Schlaug G (2014) Intensive therapy induces contralateral white matter changes in chronic stroke patients with Broca's aphasia. Brain Lang 136:1–7.

Whitehead JC, Armony JL (2018) Singing in the brain: neural representation of music and voice as revealed by fMRI. Hum Brain Mapp 39:4913–4924.

Yeh FC, Tseng WYI (2011) NTU-90: a high angular resolution brain atlas constructed by q-space diffeomorphic reconstruction. Neuroimage 58:91–99.

Yeh FC, Wedeen VJ, Tseng WYI (2010) Generalized q-sampling imaging. IEEE Trans Med Imaging 29:1626–1635.

Yeh FC, Verstynen TD, Wang Y, Fernández-Miranda JC, Tseng WYI (2013) Deterministic diffusion fiber tracking improved by quantitative anisotropy. PLoS One 8:e80713.

Yeh FC, Badre D, Verstynen T (2016a) Connectometry: a statistical approach harnessing the analytical potential of the local connectome. Neuroimage 125:162–171.

Yeh FC, Vettel JM, Singh A, Poczos B, Grafton ST, Erickson KI, Tseng WYI, Verstynen TD (2016b) Quantifying differences and similarities in whole-brain white matter architecture using local connectome fingerprints. PLoS Comput Biol 12:e1005203.

Yeh FC, Liu L, Hitchens TK, Wu YL (2017) Mapping immune cell infiltration using restricted diffusion MRI. Magn Reson Med 77:603–612.

Yeh FC, Panesar S, Barrios J, Fernandes D, Abhinav K, Meola A, Fernandez-Miranda JC (2019) Automatic removal of false connections in diffusion MRI tractography using topology-informed pruning (TIP). Neurotherapeutics 16:52–58.

Zarate JM (2013) The neural control of singing. Front Hum Neurosci 7:237.

Zhang X, Liu JY, Liao WJ, Chen XP (2021) Differential effects of physical and social enriched environment on angiogenesis in male rats after cerebral ischemia/reperfusion injury. Front Hum Neurosci 15:622911.

Zumbansen A, Peretz I, Hébert S (2014) The combination of rhythm and pitch can account for the beneficial effect of melodic intonation therapy on connected speech improvements in Broca's aphasia. Front Hum Neurosci 8:592.