Title: Hemispheric specialization of functions are tuned by conduction velocities of neuronal propagation in large-scale brain networks

Authors: Neeraj Kumar *[1], Dipanjan Roy [1,2], Arpan Banerjee*[1]

Email: neeraj.b15@nbrc.ac.in (N. Kumar) arpan@nbrc.ac.in (A. Banerjee)

Affiliation:
1. Cognitive Brain Dynamics Laboratory, National Brain Research Centre, NH8, Manesar, Gurgaon 122052, India
2. School of Artificial Intelligence and Design, IIT Jodhpur, NH62, Jodhpur, India

Abstract

We propose a tuning-by-delay hypothesis to explain the hemispheric specialization of function that has intrigued psychologists, philosophers, neurologists and neuroscientists for decades. In the auditory domain, speech and melody processing are understood to be lateralized in the left and the right hemispheric brain areas, respectively. Thus, analogous to the notion of pleiotropy where one gene can influence two or three unrelated traits, there exists a shared structural connectome encompassing both hemispheres underlying speech and melody processing, which however gets segregated along left vs right hemispheres, respectively. Here, we demonstrate how empirical observations of hemispheric specialization of speech and melody are shaped by the computational time-scales of information integration in cortical networks. First, we demonstrate that context-specific causal outflow of information emerging from primary auditory cortices (PAC) drives the hemispheric specialization of speech and melody processing when human volunteers listened to *a cappella* songs placed within a delayed match to sample task while electroencephalogram (EEG) were recorded. Second, together with participant specific whole-brain connectome model guided by diffusion weighted imaging, we predicted individual specific lateralization indices of inflow in cortical sources - after rigorous source time series reconstruction of EEG spectra. High levels of accuracy in the prediction of laterality indices in the extended large-scale auditory related regions can be achieved after optimizing conduction speeds of information propagation over a neuro-oscillatory network - a novel way to interpret about the neural mechanisms. We demonstrate that parametric modulation of conduction speeds that effectively controls the transmission delays - a key metric for understanding information processing and control of any biological network. Thus, the

transmission delay in turn, acts as the switch and triggered by the spectro-temporal complexity of the task context to select the geometry of lateralization.

Introduction

Hemispheric specialization of functions has been established as a fundamentally important empirical phenomenon in neuroscience with substantial clinical relevance (Broca, 1865; Sperry, 1974; Geschwind & Galaburda, 1985; Hellige, 1993; Zatorre & Belin, 2001; Toga & Thompson, 2003; Hecht, 2010; Gotts et al., 2013; Kong, et al., 2018; Brederroo et al., 2019; Karolis et al., 2019). While significance of hemispheric specialization from the perspective of efficient information processing is unanimously agreed on (Corbalis 2017; Güntürkün et al., 2020; for auditory domain see Hickok & Poeppel, 2007; Zatorre, 2022), many studies have identified the lack of mechanistic insights in understanding of how functional specialization occurs at the first place (Gotts et al., 2013; Güntürkün & Ströckens 2017; Brederroo et al., 2019; Karolis et al., 2019; Güntürkün et al., 2020). From a theoretical perspective, the conceptual paradigms of *statistical* and *causal complementarity*, and *input asymmetry* have been used to explain the findings such as increased left hemispheric dominance of language processing and right hemispheric dominance of face processing and motor processing (Brederroo et al., 2019; Gotts et al., 2013). Statistical complementarity suggests that lateralization is an outcome of independent statistical sampling (Liu et al., 2009), causal complementarity and input asymmetry, on the other hand, propose lateralization of one function affects the other. For instance, more left lateralized for language processing is indicative of more right lateralized face processing (Gerrits et al., 2019). The input asymmetry can integrate these frameworks by arguing co-lateralization can emerge selectively because of a hierarchical consolidation of ipsilateral processing from low level to high level cortical structures (Woodhead et al., 2011). Thus, higher activity in left occipitotemporal cortex for higher spatial frequency sinusoidal gratings compared to lower frequencies will consolidate leftward bias of word processing.

Earlier studies have indicated that such asymmetry in lateralization schemes can be brought forth by unequal involvement of inter-hemispheric communication for different functions via integrative mechanisms from the vantage point of brain network mechanisms (Gotts et al., 2013, Kumar et al., 2023). Concurrently, observations of specialization in other species having smaller and distinctly different kind of brain topology raises the question whether specialization can solely be explained on the basis of inter-hemispheric transfer, and argue that functional specialization must occur from a set of complex multi-scale processes (Güntürkün et al., 2020). One possible candidate beyond inter-hemispheric transfer may be conduction speeds across the entire connectome that shapes network dynamics which in turn has to encode the spectro-temporal complexity of environmental stimuli (Karolis et al., 2013). Conduction

speeds govern the transmission delays in neural communication among brain network nodes and have been identified as a key variable that facilitates the entry and exit to synchronized oscillatory states of the brain (Jirsa & Ding, 2004; Ghosh et al., 2008; Petkoski & Jirsa 2019). While, several studies have argued that the changes in transmission delays are of structural origin specifically introduced via myelin modulations along axons (Ghosh et al., 2008; Petkoski & Jirsa 2019; Pathak et al., 2022), empirical evidence points to the existence of functionally tuned synaptic delays (Stange-Marten et al., 2017; Byczkowicz et al., 2019). Functional requirements are set up by a complex system of interactions between internally generated network states and external inputs such as arousal (Vinck et al., 2015), or by neuromodulators (Byczkowicz et al., 2019). Here, we propose a - tuning by delay hypothesis - to conceptualize that while the myelin can provide the basic infrastructure of communications such as expressways to go at high speeds, the functional requirement set by internal awareness system and external systems decides the speed on such paths. Synaptic scaling weights, that can be indirectly measured from diffusion weighted imaging (dWI) data (Schirner, et al., 2015) can also guide the whole-brain synchrony captured in fMRI and MEG signals (Cabral et al., 2011; Pathak et al., 2022). Hence, it becomes imperative to test that given an individual's brain structural information, how the task-specific input asymmetries such as speech or melody contexts can be mapped to the conduction velocities with which communication is orchestrated among the auditory network nodes.

Concomitantly, converging evidence suggests that the left primary auditory cortex (PAC) is specifically attuned to rapid temporal changes associated with speech processing, while the right PAC is dominant in processing spectral changes associated with melody processing (Albouy et al., 2020; Zatorre & Belin, 2001). Therefore, the lateralization observed in higher-order cortical areas could likely to originate, at least in part from the PAC. Understanding how this oscillatory information is transmitted to higher-order regions is crucial for comprehensive mapping of the complex dynamics underlying auditory perception and consequently to hemispheric specialization. To address this issue we sought to measure the causal flows of information processing using Granger-Gewke causality (Geweke 1982), from source reconstructed EEG time series on human participants when they listened to *a capella* songs and selectively attended to either the speech or melodic content (Albouy et al., 2020) . Lateralization indices (LI) were computed for the inflow to each source node. Subsequently, to unravel the mechanistic basis of empirical LI patterns, model-predicted LI (referred henceforth as predicted LI) was computed from whole-brain connectome model informed by realistic structural parameters subject-by-subject. The model

involved the Kuramoto oscillator framework (Kuramoto 1984) to capture the phase dynamics of an individual source node (Cabral et al 2011; Petkoski et al., 2018; Pathak et al., 2022) coupled through connection matrix generated from diffusion weighted imaging (dWI) data. We argue that, the transmission delays are set at the gatekeeper node PAC, receiver of auditory sub-cortical output from medial geniculate nucleus (Rauschecker and Tian 2000). We also use a model inversion technique that computes the delays by optimizing a broader parameter space of neural response frequency and synaptic coupling - both within hemispheric and inter-hemispheric, paving the way for individualized prediction of LI required to establish the robustness of the analysis (Seghier & Price, 2018) (see Methods). The frequency-specific tuning of transmission delays was validated by additional auditory steady state (ASSR) recordings, where brain networks are entrained sharply to input frequencies of sounds.

Methods

Participants

Thirty healthy participants (14 males, 16 females, age range 22–34 years old; mean ± SD = 27 ± 3.14) participated in this study. All participants were right-handed, as confirmed by the Edinburgh Handedness Questionnaire with a cut-off score of 60-100, and reported no history of audiological, neurological or psychiatric disorders. All had normal or corrected-to-normal visual acuity. Informed consent was obtained from all participants in a format approved by the Institutional Human Ethics Committee (IHEC) of National Brain Research Centre. All participants were fluent in at least two languages, Hindi and English, with some having knowledge of another language of Indian origin.

Auditory stimuli

We have used two types of auditory stimuli 1.) *a cappella* songs (Albouy et al., 2020) and 2.) amplitude modulated tones to generate auditory steady-state responses (ASSR). The first set of auditory stimuli consisted of 100 *a cappella* songs prepared by Albouy and colleagues, and have been used and described in more detail in their earlier reports (Albouy et al., 2020). Particularly, the a cappella songs were composed of 10 tones, each with an identical rhythm, and ten English sentences composed of 10 syllables each. Sentences were modified to adjust with the rhythm of the melodies. All melodic and sentence materials were combined in all possible combinations to form 100 a cappella songs. The resulting set of 100 songs had a mean duration of 4.1 seconds. The second set of auditory stimuli were pure sine wave tones of frequency 500 Hz, amplitude modulated at 40 Hz. The sound intensity level was maintained at 60 dB SPL for all stimuli.

Paradigm and experimental sessions

Participants underwent two separate sessions: an MRI session in which T1-weighted structural MRI images and diffusion weighted imaging (dWI) data were acquired, and an EEG session during which electrode locations were also obtained using a 3D digitizer. The EEG session always took place on a different day and was scheduled after the MRI session.

The EEG session consisted of 3 blocks. First, the participants underwent 5 minutes of EEG recording in a resting state while fixating on a cross displayed at the center of the screen. Second, after being verbally instructed about the task participants were binaurally presented with pairs of a cappella songs via earphones and were instructed to compare them. To ensure that participants focused their attention on speech or melody content during the presentation of a cappella song, they were engaged in a match-to-sample task. The trial started with a visual cue indicating the relevant domain (sentence or melody) for 1200 milliseconds. Subsequently, the first song was presented, followed by a 1000 millisecond interstimulus interval (ISI), and then the second song was presented. Throughout the entire duration of the songs and ISI, a fixation cross was presented at the center of the screen. Following the presentation of the second song, participants were presented with a question on the screen, requiring them to assess whether the two speech contents were identical or different, irrespective of the melody to which they were sung in 'speech' trials. In the case of 'melody' trials, this process was reversed, with the judgment based solely on the melody, independent of the sentence content. Participants were required to respond by selecting the right arrow key on the keyboard when the domain was the same and the left arrow key for different. The response window lasted for 2000 milliseconds, and participants received no feedback. The total duration of a trial was approximately 12 seconds, and the total duration of the experiment was 20 minutes. There were four types of combinations included: (1) both speech and melody were different in the song pair, (2) speech was the same but the melody was different, (3) speech was different and the melody was the same, and (4) both speech and melody were the same. Each combination had 24 trials, resulting in a total of 96 trials presented. The sentence/melody questions and same/different responses were pseudo-randomly presented, with each stimulus type uniformly distributed over the entire experiment. Presentation software (Neurobehavioral Systems, Berkeley, CA, USA) was used to control stimulus presentation and record participants' responses. Continuous electroencephalography (EEG) was recorded in a noise attenuated isolated room using a BrainVision Recorder acquisition system, which included an actiCHamp module with 63 active channels placed according to the International 10-20 electrode placement system. SuperVisc electrolyte gel (EASYCAP) was used to establish contact between the EEG sensors and the scalp. Continuous EEG data were acquired at a sampling rate of 1 kHz, and the impedance of each sensor was kept below 10 kΩ. The reference electrode was positioned at the vertex (Cz), and the forehead (AFz) electrode was selected as the ground. The electrode locations were obtained with respect to three fiducials at the nasion and left and right preauricular points using a 3D digitizer (Polhemus Inc., Colchester, VT, USA).

EEG Preprocessing

Initial EEG data analysis was performed with EEGLAB (Delorme and Makeig, 2004) for removing artifacts from the data, followed by customized MATLAB scripts (www.mathworks.com) and FieldTrip (http://fieldtriptoolbox.org) toolbox for spectral analysis and source analysis. Firstly, using EEGLAB continuous EEG data from each auditory block and resting state were bandpass filtered between 2-48 Hz and independent component analysis was employed to eliminate eyeblinks, muscular, and heartbeat artifacts. The ASSR block trials were epoched for a duration of 1 second from the onset of stimuli, while for the *a cappella* song block, hit trials were identified and segregated into two groups for speech and melody trials and epoched for a duration 4 seconds from the onset of song. In addition, non-overlapping 1 second and 4 second (50% overlapping) time segments from the resting state data were epoched to be utilized as the baseline condition for the ASSR and song blocks, respectively. After detrending the trials, threshold-based artifact rejection was applied at ± 80 μV on trials. Following the preprocessing steps, an average of 97 trials (out of 100) from the ASSR block, 81 trials (out of 96) from the speech block, and 69 trials (out of 96) from the melody block were retained. The preprocessed EEG data were then re-referenced to a common average reference and downsampled to 250 Hz for further processing.

MRI-DWI acquisition and analyses

The T1-weighted structural MRI images were acquired on a Philips Achieva 3.0 T MRI scanner with the following parameters: TR = 8.4 ms, FOV = 250 × 230 × 170, flip angle = 8°, 170 sagittal slices, and voxel size of 1 × 1 × 1 mm. The fiducials were marked at the nasion, left, and right pre-auricular points using Vitamin E capsules to be used in EEG sensor coregistration with MRI for source localization. The initial preprocessing of MRI images and volumetric parcellation for individual subjects were performed using Freesurfer, based on the Desikan- Kilinay atlas (http://surfer.nmr.mgh.harvard.edu/).

Diffusion weighted Imaging (DWI) data was acquired through a single-shot echo planar imaging in 3.0 T Philips Achieva Scanner. The DWI sequence was performed with TR = 8800 ms, TE = 75 ms, FOV = 210 × 210 × 128 mm, flip angle = 9°, matrix size = 104 × 104, slice thickness = 2 mm, no gap, 64 axial slices, and voxel size of 2 × 2 × 2 mm. The diffusion was measured along 32 non-collinear directions using a *b*-value of 1000 s/mm^2, with a b = 0 image included as the first volume of the acquisition.

Image processing and structural connectivity

The processing of diffusion data and T1 MRI data to construct structural connectome (SC) was performed in FSL and ANTS based MRtrix (http://mrtrix.org/), BrainSuite (https://brainsuite.org/), and FreeSurfer. Structural connectivity estimation was performed employing the BATMAN pipeline implemented in MRtrix software. Particularly, probabilistic tractography was performed based on the Constrained Spherical Deconvolution (CSD) algorithm (Tournier et al., 2004, Tournier et al., 2007). The CSD algorithm evidently outperforms the diffusion tensor model (DTI) in regions containing crossing fibers, which DTI cannot capture due to its ellipsoid approach to fiber orientation estimation. The initial preprocessing of diffusion-weighted images included denoising to estimate the spatially varying noise maps (Veraart et al., 2016b, Veraart et al., 2016a), unringing to remove Gibb's ringing artifacts was performed in MRtrix (Kellner et al., 2016) and bias field correction of diffusion images (Tustison et al., 2010). Then MRI images are preprocessed for co-registration with diffusion data. Therefore, firstly, skull-stripping of MR images was performed by the Cortical Surface Extraction Tool (Sandor et al., 1997; Shattuck et al., 2001), followed by correction of gain variation by the Bias Field Corrector (BFC) software (Shattuck et al., 2001), part of the BrainSuite. Particularly, it was used to correct image intensity non-uniformities in magnetic resonance images, which can cause confounding effects in tissue classification. BFC estimates a correction field based on a series of local estimates of the tissue gain variation and corrects for the non-uniformity field by dividing the original image by the estimated tri-cubic B-spline. Thereafter, BrainSuite Diffusion Pipeline (BDP) was used to correct motion and geometric distortions of diffusion images inherent to Echo-Planar Imaging sequences (Bhushan et al., 2012), co-registration of diffusion and anatomical images, and obtaining b-values and b-vectors to be used for further estimation of response functions (Haldar & Leahy 2013; Shattuck et al., 2002). Subsequently, to determine the diffusion orientation within each voxel, a response function was derived using Dhollander's algorithm from representative tissue types of the brain i.e., white matter, gray matter, and cerebrospinal fluid (Dhollander et al., 2016). Thereafter, Fiber Orientation Distributions (FOD) were estimated using the CSD algorithm. Intensity normalization was performed to correct for global intensity differences, and a whole-brain tractography was generated using probabilistic tractography in tandem with biologically plausible Anatomically Constrained Tractography algorithm to generate 20 million tracts seeded in gray matter (Smith et al., 2012). The tractogram was then coregistered onto individual MRI, and a subset of one hundred thousand tracts was filtered out to reduce CSD-based inherent bias in overestimation of longer tracks by Spherical-deconvolution Informed Filtering of Tracts (SIFT) algorithm (Smith et al., 2013).

Finally, the connectome was parcellated by mapping the tracts onto MRI data, defined according to the Desikan Kiliany atlas (Desikan et al., 2006), that was segmented in FreeSurfer (https://surfer.nmr.mgh.harvard.edu/). The resulting 68 x 68 parcellated whole-brain structural connectivity (SC) weights and fiber lengths matrices were representative of the number of streamlines and mean streamline length between each node pair, respectively. Both SC matrices were normalized by the inverse of the two node volumes for scaling the track count. The SC matrix and fiber length matrix were both symmetric matrices and values at the diagonals of structural connectivity and fiber length matrix representing self-connectivity and length with self, respectively, were set at zero.

Source-level analysis of EEG data

To estimate frequency-specific oscillatory source activity, we utilized exact low-resolution brain electromagnetic tomography (eLORETA), a distributed source modeling method that calculates the current source density across the brain volume by minimizing the surface Laplacian component of the spatial filter (Pascual-Marqui, 2007). Notably, eLORETA has the advantage of not requiring any assumptions about the number of underlying sources and is highly effective in controlling false positives during source detection (Halder et al., 2019). Therefore, we performed source analysis using eLORETA with the FieldTrip toolbox (Oostenveld et al., 2011).

For forward modelling, MRI data from each participant was segmented into three tissue types, and meshes were created with 3000, 2000, and 1000 vertices for brain, skull, and scalp, respectively. The boundary element method (BEM) was employed to generate the realistic volume conduction model using the OpenMEEG package (Fuchs et al., 2001; Gramfort et al., 2010). Subsequently, the EEG channels were aligned with the BEM model based on 3 fiducials present in both Polhemus data and MRI. Thereafter, leadfields were created for 3 orthogonal directions using the sensor positions and the volume conduction model. The location of vertices within the leadfield was defined based upon the participant-specific T1 MRI previously parcellated according to Desikan-Killiany atlas in FreeSurfer (Desikan et al., 2006), with an average resolution of 4 mm.

Subsequently, the cross-spectral for each frequency band (2-48 Hz) were computed from the preprocessed sensor level data over the entire time window for each condition for ASSR, speech, melody, and baseline conditions. Frequency-specific spatial filters were computed from combination of the forward model and cross-spectral matrix for each condition. Separate spatial filters were constructed for ASSR, speech, and

melody conditions, with 1-second and 4-second rest trials serving as baseline conditions, respectively. To reduce dimensionality, principal component analysis was performed on the spatial filters of vertices belonging to individual parcels, resulting in the spatial filters of 68 regions of interest. The trial-wise sensor-level cross-spectra were then projected to the spatial filter, obtaining subject and trial-wise distribution of source level Fourier coefficients across the brain volume for each condition.

Granger causality analysis

From the source we obtained the reconstructed Fourier coefficients of entire frequency range. Thereafter, subject-wise multivariate Granger causality (GC) was calculated in the spectral domain to determine the causal outflow from PAC to whole-brain (Brovelli et al., 2004; Dhamala et al., 2008; Granger, 1969). To calculate the pairwise Granger causality (GC) from PAC to X regions at frequency (f), a nonparametric spectral matrix factorization was performed using Wilson's algorithm. This factorization of the cross-spectral density allowed for the determination of the transfer function H and the noise covariance matrix $\sum$. Subsequently, the pairwise GC can be expressed as follows:

$$I_{PAC\rightarrow X} = \ln \frac{S_{xx}(f)}{S_{xx}(f) - \left(\Sigma_{yy} - \frac{\Sigma_{xy}^2}{\Sigma_{xx}} |H_{xy}(f)|^2\right)} \quad (1)$$

where $S_{xx}(f)$ represents the total spectral power (auto-spectrum), while the denominator represents the intrinsic power of the “effect” signal. The intrinsic power is calculated by subtracting the causal contribution from the total power (Geweke, 1982, Dhamala et al 2008).

The GC spectra of speech and melody condition were divided into respective frequency bands, i.e., delta (2-4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–48 Hz) bands. Significant frequency bands were identified from non-parametric statistical testing (Brovelli et al., 2004). Particularly, 1000 permuted data sets were generated by independently shuffling trials between resting condition and auditory conditions. Thereafter, GC was computed, and the maximum GC value was selected over the frequency range from each permuted data set (Ding et al., 2006). Subsequently, a null distribution consisting of all GC values was constructed from the shuffled data set. GC bands in unshuffled data were considered statistically significant when the observed GC value reached beyond the 99th quantile value ($p = 0.01$) of the null distribution. The multiple comparison problem was handled by Bonferroni correction.

Since the key hypothesis of the manuscript was to investigate the input to the auditory network which is gated by left and right PAC, we restricted our analysis of GC to only the causal outflows from bilateral PAC to other nodes of the auditory network.

Laterality analysis

To evaluate the extent of asymmetry in outflow from the bilateral primary auditory cortices (PAC) during auditory conditions, we calculated the Laterality Indices (LI) of the inflow to each area (Kumar et al., 2023). The LI for each node of the auditory network was then calculated as the difference between the total GC inflow (from bilateral PAC) to the right region and the left region, divided by the sum of the inflows to both regions.

$$LI = \frac{RH - LH}{RH + LH} \quad (2)$$

LI values range from +1 to -1, indicating complete right dominance and left dominance, respectively, while a value of 0 suggests a bilaterally symmetrical response. Statistical significance was of the LI distribution across subjects were determined by computing 95% confidence intervals.

Network model of neural activity

We aimed to model the propagation of oscillatory activity from primary auditory cortices to specialized regions of the brain during speech, melody and ASSR conditions using a simplified computational model, retaining the essence of key physiological parameters. The propagation of oscillation can be conceptualized as unidirectional interactions originating in PAC to form a functional neural network. Such networks communicate via phase-based synchronization (Kuramoto, 1984) and were shown to be the processing units of cognition (Cabral et al., 2011; Pathak et al., 2022). We employ the Kuramoto model to simulate the source time series of each phase oscillator equivalent to phase dynamics of one parcel (Kuramoto, 1984; Cabral et al., 2011; Pathak et al., 2022). Hence, a network of 68 coupled Kuramoto oscillators was simulated, following Desikan-Killiany parcellation scheme, also comprising bilateral primary auditory cortices (PAC). Each other oscillators had intrinsic noise (η), which was randomly drawn from a distribution with a mean of zero and a standard deviation of one, reflecting the natural variability of oscillatory dynamics in the brain. The remaining two oscillators situated in the primary auditory cortices and have fixed intrinsic frequencies that correspond to their oscillatory responses. This framework allowed the unidirectional propagation of entertainment from PAC to the specific nodes and

consequent enhancement of spectral power of those nodes at the frequency of PAC stimulation (Fig. 1; Right panel). To make the model bio-physiologically realistic, information about structural connectivity was incorporated in the model by having the coupling between oscillators determined by strength of anatomical connections (fibre thickness in dWI, see section "Image processing and structural connectivity"). The time scale separation for processing of "fast" speech and "slow" melody could be achieved by a propagation time-delay (τ) in the model such that the current phase $\theta(t)$ is dependent on its interaction with the past phase $\theta(t-\tau)$. Inadvertently, τ can be conceptualized to be emerging from a joint contribution of the functional processing in the auditory network and the constraints posed by myelination, although the former is the only quantity that can change in a participant and in a session specific way. Thus, the dynamics of each oscillator $\dot{\theta}_n$ are governed by the following equation:

$$\dot{\theta}_n = k\sum_{p=1}^{N} C_{np}\sin\left(\theta_p(t-\tau_{np}) - \theta_n(t)\right) + \eta, \forall n = 1,2 \ldots (N-2), p \in [1, N], p \neq n \qquad (3)$$

The coupling strength matrix C_{np} was normalized between 0 and 1 such that the maximum strength among connections was 1 and, k is the global mean coupling strength used to scale all the coupling strengths; τ_{np} represents transmission delay for the propagation of information between two nodes. Thus, $\tau_{np} = \frac{D_{np}}{v}$, where for a bio-physiologically realistic communication speed (v) is $5-20\,m/s$ in the adult primate brain (Ghosh et al., 2008), mean τ ranged from 3.4 to 17 when D_{np} was fibre length between nodes *n* and *p*. Therefore, the dynamics of phase (θ) at any node is a function of its anatomical strength, distance with other nodes and propagation speed v, which is set be the nature of functional processing. While equation (3) describes the dynamics of 66 non-auditory nodes, the phase dynamics of bilateral auditory cortices (auditory nodes) are defined as,

$$\dot{\theta}_A = \omega_A \qquad (4)$$

where $\omega_A = 2\pi * f_A$, (f_A) representing the frequency of auditory nodes.

Participant-wise model fitting

The subject-by-subject laterality indices obtained from the empirical Granger-causality estimation were fitted with the LI's computed from analyzing synthetic neural time-series generated by the network model (equations 3, 4). The system of differential equations (equation 3,4) was numerically integrated using the Euler integration method for 250000-time points with a step size (dt) of 0.0001 representative of 25 seconds duration. We took the sine of (θ) obtained at each node which represents a simulation of neural time series at EEG source level. Subsequently, we calculated the power spectral density from 68 nodes. For each combination of frequency (f), global coupling (k), and time delay (τ) model simulation, simulated lateralization indices were calculated from power spectral density obtained from each 68 parcels (equation 2). We ran an optimization algorithm to obtain a set ($f_{opt}, k_{opt}, \tau_{opt}$), at which the Pearson correlation between simulated and empirical LI reaches a global maximum. Hence, a model inversion was achieved for the following ranges of each parameter, $f_{opt} \in [1,48]$; $k_{opt} \in [1,20]$; $\tau_{opt} \in [3.4,17]$ in steps of 0.4. The frequency window of interest was guided by the observation window of current study, the range of k_{opt} and τ_{opt} chosen has been shown earlier to represent biophysically relevance (Ghosh et al., 2008; Cabral et al 2011).

We hypothesize, to process a specific auditory task, phase-based communication in an oscillatory network can be achieved by adjusting the combination of these free parameters ($f_{opt}, k_{opt}, \tau_{opt}$). While the fitting was constrained by maximizing Pearson correlation between empirical LI and model generated LI, frequency-specificity of these optimal parameters model need to be studied to ensure that the simulated neural responses align with the observed empirical frequency-specific enhancements in Granger causality (GC) outflow from the bilateral primary auditory cortices (PAC). Hence, the model inversion was carried out independently for the major frequency bands, delta (2-4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–48 Hz) for speech and melody and also specifically at 40 Hz for ASSR condition. Importantly, it was required to ensure the frequency specificity of the model i.e., the significant distribution f_{opt} across subjects should fall into the corresponding frequency range. For instance, the confidence interval of f_{opt} to predict empirical theta lateralization should fall into the theta range and so on. Hence, to establish a valid comparison between empirical EEG frequency and f_{opt} that can establish the frequency selectivity of the model, we also correlated the empirical LI across a range of probable model frequencies ($f_{simrange}$) and estimated the 95 % confidence bound of resultant f_{opt}. Since, the outflow from PAC during speech and melody existed in the broadband frequency range rather than a

single frequency unlike the model simulation, making the comparative assessments between model and empirical results were not straightforward resulting to spurious non-unique values of frequencies and other parameters during the model inversion process. Hence, 40 Hz ASSR condition where the signal displays a salient entraining frequency becomes a further validation ground to assess whether the model can be frequency-selective. Outflow from bilateral PAC was shown earlier to be present only at 40 Hz during 40 Hz ASSR response (Kumar et al., 2023). Hence, for the purpose of current study, demonstration of maximum proximity of model generated LI to empirical LI concomitant with matching of model estimated frequency and ASSR frequency gives us the confidence about the model estimation process when a complex range of frequencies are present in signal, e.g., during speech and melody processing. Furthermore, there could be different number of target regions in each frequency bands across different conditions. Hence, for practical reasons and to standardize the comparative analysis across conditions, empirical distribution of LI values across the entire brain volume (34 values) are correlated. Thus, the combination of $f_{opt}, k_{opt}, \tau_{opt}$ that yielded maximum correlation with empirical LI was selected for each subject. Note that the k_{opt} and τ_{opt} are valid only if the f_{opt} lies within the same range of empirical frequency band.

Furthermore, to confirm the accuracy of the model in utilizing structural connectivity (SC) to guide predictions, a control condition was implemented. In this condition, the weights and fiber lengths of the SC matrices for each subject were shuffled and used for the connectome model simulation. Specifically, the weights and fiber lengths of the SC matrices were randomly reassigned while maintaining the same number of connections and nodes. Comparative analyses were then performed between the shuffled and empirical 40 Hz ASSR condition, focusing on the assessment of the frequency specificity of the model. In particular, the model's ability to accurately predict the frequency-specific responses, hence at 40 Hz, was examined. The presence of frequency-specific enhancement (f_{opt}) in the shuffled and unshuffled conditions was compared, with particular emphasis on the distribution of f_{opt} in each condition. By contrasting the shuffled and unshuffled conditions, this assessment aimed to determine the extent to which the model's predictions were reliant on the intact organization of the structural connections in the brain. Notably, the fact that each task state such as processing speech / melody requires idiosyncratic distribution of these free parameters given the specific nature of information processing. Additionally, these neurophysiologically relevant free parameters would not be identical in every subject owing to the inter-

subject variability (Seghier & Price, 2018). Subject-specific prediction enables not only contribute to the robustness of the analysis but also enables attributing the functional variability to the differences in individual structural connectivity. Nevertheless, another attribute of population studies is the existence of central tendency in the distribution of free parameter across subjects. Hence, assessing the normality of k_{opt} and τ_{opt} across subjects was required in this framework to evaluate the suitability of assuming a normal distribution for these parameters. A normal distribution indicates a predictable and stable pattern of parameter values, which is desirable for interpreting and generalizing the results of our model. Hence, we assessed the extent of prediction for individual regions and its inter-subject variability as the empirical source analysis revealed that only specific regions exhibited frequency-specific enhancement in Granger causality (GC) outflow from the bilateral PAC. We employed a linear regression analysis, wherein the empirical LI values of individual regions across subjects were regressed against the corresponding simulated LI values. The goodness of fit was evaluated by assessing the coefficient of determination (R-squared) and the significance of the regression coefficients. The LI values of a particular region were selected from the spatial distribution of LI that yielded maximum correlation a f_{opt}, k_{opt}, and τ_{opt}. Finally, we examined the global coupling (k_{opt}) and mean delay (τ_{opt}) between the speech and melody conditions across all frequencies via paired t-test. This analysis allowed us to explore whether the brain exhibited distinct functional mechanisms to process auditory signals of various spatiotemporal complexities, e.g., speech, melody and rhythmic tonal sounds.

Results

First, we present the results from empirical EEG data analysis during three auditory conditions, speech, melody and 40 Hz ASSR. The laterality indices (LI) of frequency-specific net Granger causality outflow from bilateral primary auditory cortices were estimated for each task condition. Second, we present the computational modelling results, where bilateral primary auditory cortices were stimulated in-silico, at frequencies ranging from 1 to 48 Hz. The entrainment propagated throughout the entire brain constructed via coupling brain areas (using Desikan-Kiliany parcellation) with strengths derived from diffusion weighted imaging derived structural connectivity (see Methods, as well as Schirner, et al 2015; Pathak et al 2022). Model parameters - a range of realistic conduction velocities that influences the effective time-scales of processing information (scales inversely with time delays in the oscillator model, τ) and global synaptic coupling values (k) among the nodes, were predicted using from best fits of empirical LIs using model generated LIs. Parameters estimated from the model inversion were subsequently used for inferring about the underlying neural mechanisms of functional hemispheric lateralization.

Causal outflow of information from primary auditory cortices maps auditory input asymmetries

Cross-spectral matrices at source level were reconstructed using eLORETA at parcels identified by Desikan Kiliany atlas for each participant during speech, melody, ASSR and resting conditions. Multivariate Granger causality analysis revealed significant ($p<0.01$) causal outflow from bilateral primary auditory cortices to other parcels at distinct frequency ranges: delta, theta, beta and gamma during both speech and melody processing (Fig. 2). The spatially averaged causal outflow across all significantly active sources were similar in the alpha frequency for speech and melody perception but differed across delta, beta and gamma (Fig. 2; Right panel). During the ASSR condition, the significant outflow from bilateral PAC were observed specifically at 40 Hz, whereas speech and melody processing exhibited distinct frequency bands with some sources showing responses at multiple frequency bands (see Fig. 2; left panel and Table 1 for details). Overall, the sources included superior temporal, middle temporal lobe, precentral, Broca's areas (pars opercularis and pars triangularis), superior frontal, supramarginal and insula. Source in the lateral orbitofrontal and parahippocampal cortex were exclusive to ASSR. Notably, the left pars opercularis showed significant activation across all frequency bands during speech conditions, indicating its involvement in left hemisphere dominance during language processing. On the other hand,

the right pars triangularis exhibited significant activation across all frequency bands specifically during melody conditions, suggesting its role in right hemispheric dominance of melody processing (Fig. 2).

Lateralization of outflows from bilateral PAC to the other significantly active parcels for different auditory conditions are captured by computing laterality indices for each parcel. The comprehensive list of regions, along with their corresponding 95% confidence intervals and mean LI are provided in Table 2. Certain regions exhibited significant lateralization to either side; where both the upper and lower confidence intervals lie completely on either side of zero (Table 2). Particularly, during ASSR the right hemispheric dominance was present in superior temporal gyrus (STG) with mean LI = 0.13 and lower and upper confidence interval (CI) at 95% significance - [0.08, 0.18]; lateral orbitofrontal (LI = 0.31, CI = [0.17, 0.44]) and parahippocampal (LI = 0.19, CI = [0.05, 0.32]). Analogously, for the speech condition, the left hemispheric dominance in beta band power was present in STG (LI = -0.03, CI = [-0.04, -0.02]), precentral (PrC) (LI = -0.03, CI = [-0.05, -0.01]), pars opercularis (LI = -0.07, CI = [-0.10, -0.05]) and in gamma band power for STG (LI = -0.13, CI = [-0.15, -0.10]), PrC (LI = -0.16, CI = [-0.24, -0.08]), pars opercularis (LI = -0.22, CI = [-0.31, -0.13]). For the melody condition, right hemispheric dominance of beta band was present in STG (LI = 0.10, CI = [0.06, 0.14]), PrC (LI = 0.11, 95% CI = [0.03, 0.18]), pars triangularis (LI = 0.20, CI = [0.05, 0.35]) and at gamma band, in STG (LI = 0.17, CI = [0.12, 0.22]), PrC (LI = 0.17, CI = [0.04, 0.31]), pars triangularis (LI = 0.12, CI = [0.03, 0.22]).

Estimation of transmission delays, synaptic scaling and neurally mapped input frequencies

We employed the Kuramoto phase oscillator network (Kuramoto, 1984) coupled by heterogeneous coupling derived from diffusion weighted images to model neural oscillations observed in the whole-brain connectome (see Methods for details). Since our interest is only in the causal outflow from PAC, the input node was only PAC for the simulations (see Methods for details, equations 3 and 4). Subsequently, the steady state oscillatory responses in each node of this network stems from the inflow of communication from PAC. Hence, the laterality indices (LI) can be computed using the power spectral density at each parcel. By maximizing the correlations between the simulated and empirical LI's we can undertake a model inversion to estimate transmission delays (τ_{opt}), global coupling (k_{opt}) and neurally mapped input frequencies (f_{opt}). Subsequently, the following issues were addressed. First, we take the auditory steady state responses tuned sharply at a specific entraining input frequency to validate if the network model can predict the LI's, tuned at the known frequency, while, readjusting transmission delays and

global coupling (Fig. 3). Second, after establishing the validity of the auditory network model, we investigate the complex multifrequency EEG response for speech and melody processing and estimate the transmission delays and global coupling while ensuring frequency selectivity in band specific (delta, theta, beta and gamma) responses and maximizing the LIs (Fig. 3). The estimated parameters can then be used to understand the mechanisms of processing speech and melody by a commonly shared structural auditory network.

Validation of auditory network model using ASSR

Since, ASSR involves a sharp frequency tuned response, the proposed auditory network model parameters when fitted with a constraining factor of maximizing LI, should be able to detect the frequency of tuned neural oscillations around 40 Hz. Using a realistic parameter search space, $f_{opt} \in [1,48]$, $k_{opt} \in [1,20]$, $\tau_{opt} \in [3.417]$, frequency-selectivity of the simulated auditory network was estimated by first constructing a null distribution with 95% confidence bounds for f_{opt}. The control null distribution was created by shuffling the SC matrices, such that each connection weight is essentially a random rational number between 0 and 1, and using this randomized SC to generate neural time series that was used to fit the empirical 40 Hz ASSR. No frequency-specificity was observed from the resultant synthetic data, f_{opt} was distributed in the whole $f_{simrange}$. The maximum correlation ranged between 0.4 to 0.7 across all participants (Fig. 3; upper right panel). In contrast, when using the unshuffled SC matrices, the empirical 40 Hz ASSR showed frequency specificity, with a 95% confidence interval of [39.4, 40.9] Hz and maximum correlation values ranging from 0.79 to 0.96 (both > 0.7), thus, confirming the model's ability to predict empirical responses. The 95% confidence intervals for *k* and τ were [2.5, 3.5] Hz and [1.5, 2.2] Hz respectively. The distributions of both k_{opt} and k_{opt} followed a normal distribution, assessed using Lilliefors test with k-stat values of 0.24 ($p < 0.001$) and 0.16 ($p = 0.043$) and τ, respectively (Fig. 3; upper left panel).

Mechanisms underlying lateralization of causal outflow in beta and gamma frequencies during speech and melody processing

During speech and melody conditions, no frequency-specificity could be predicted by the auditory network model in the delta and theta frequency bands (Fig. 3; bottom panels). In the delta range, the 95%

confidence intervals of f_{opt} for speech and melody were [17.6, 29.8] Hz and [15.7, 26.9] Hz, respectively. Similarly, in the theta range, the 95% confidence intervals for speech and melody were [19.2, 30.8] Hz and [16.0, 28.1] Hz, respectively. Therefore, the k_{opt} and τ_{opt} values for these frequency ranges are not considered reliable. On the contrary, frequency-specificity was observed in beta and gamma frequency bands (see Fig. 3). For speech conditions, the 95% confidence interval of f_{opt} in beta band was [17.2, 23.7] Hz, and for melody [17.6, 22.8] Hz. In the gamma band, the 95% confidence interval of f_{opt} was [31.0, 41.7] Hz for speech, and [31.1, 37.4] Hz for melody. These CIs indicated that the model correctly captured the frequency-specific activity of empirical data within the beta and gamma bands. The values of k_{opt} and τ_{opt}also showed normal distributions for these frequency bands, indicating a smooth and gradual transition in the model's behavior required for a consistent and reliable model performance. The complete list of 95% confidence intervals for frequency, k_{opt}, and τ_{opt} as well as the results of the normality tests, are present in Table 3.

The paired t-tests showed the presence of overall differences in neural processing characteristics between speech and melody stimuli (Fig. 5A). The findings revealed a significant difference in the k parameter [$t = 3.05$, $p = 0.003$], indicating distinct global coupling between the two conditions. Moreover, there was a significant difference in the tau parameter [$t = 11.32$, $p < 0.0001$], suggesting variations in temporal delays during the processing of speech and melody stimuli (Fig. 5A). These analyses shed light on the global neural dynamics and temporal characteristics associated with different auditory stimuli.

Individual-specific prediction of functional lateralization in auditory processing networks

We conducted linear regression analyses in the beta and gamma frequency ranges to evaluate the extent of prediction for individual regions found significant in the empirical data. Most regions significantly predicted lateralization in the empirical data, with the exception of the supramarginal gyrus during speech condition in the gamma band. Figure 4 A show the respective prediction score for the individual brain region. Overall, the prediction scores for the ASSR condition were greater than the speech and melody condition. When both the upper and lower confidence intervals of a region's laterality index (LI) for both the simulated and empirical data fell on the same side of zero, we considered the model to successfully predict the side of lateralization for that region. Consistent with the empirical conditions, several regions

exhibited lateralization. Particularly, during the ASSR condition, right hemispheric dominance was predicted by the superior temporal gyrus (STG), parahippocampal cortex, and lateral orbitofrontal cortex (Fig. 4B). During speech and melody conditions, the STG predicted left hemispheric dominance during speech and right hemispheric dominance during both beta and gamma frequencies. The precentral gyrus predicted the side of lateralization during speech beta and gamma, and in melody during beta frequencies. Additionally, the pars opercularis predicted left hemispheric dominance during speech gamma conditions.

LI prediction from auditory network model follows the auditory processing hierarchy

Interestingly, we observed a systematic decrease in R-squared values for individual regions along a specific axis of information flow. When arranging the regions in a specific order, namely superior temporal gyrus (STG), middle temporal gyrus (MTG), insula, precentral gyrus, Broca's area, and supramarginal gyrus, for both speech and melody conditions, we found that the degree of prediction exhibited a consistent linear decrease in the gamma frequency band (Fig. 5B). This pattern was observed in both the speech condition [r^2 = 0.93, p = 0.001] and the melody condition [r^2 = 0.93, p = 0.001]. Note that, similar analysis in the beta frequency band was not feasible due to the limited number of regions. However, in the beta range the highest prediction accuracy was present in STG for speech and PrC for melody condition.

Discussion

The present study introduces a combination of empirical findings and theoretical understanding by how a potential functional mechanism - tuning of large-scale brain networks by modulating conduction speeds can coordinate the processing of spectro-temporal complexity of auditory perceptual objects. The outcome of such delay-tuned neurodynamics leads to the empirical observation of hemispheric lateralization for speech and melody. Our study reported a common frequency-specific outflow of information communication from PAC in both speech and melody conditions, reaching to distinct lateralized regions reflecting the domain-specific organization of the auditory system. This enhances the well-established hemispheric dominance of left hemisphere for speech, right hemisphere for melody, and tonal objects such as ASSR, elucidating the important role of PAC guided causal flow in hemispheric specializations. To further understand the origins of this flow patterns and subsequent lateralization, we utilize a structurally-guided frequency-specific neural model to predict an individual's lateralization index (LI) from their brain responses. The novel approach of whole-brain connectome modeling to explain task induced electrophysiological response during speech, melody, and ASSR conditions led to a prediction of LI of an individual subject in homologous brain areas at a high level of accuracy. Thus, the second notable aspect of current study is the individual level parametrization of the whole-brain connectome model to capture variability in brain structure across humans. By integrating individual diffusion weighted imaging (dWI) data into our neural model, we ensure subject-specific variability of the structural constraints are considered to explain the inter-subject variability of neural dynamics (Seghier & Price, 2018). We validate the frequency-specificity of the model and characterize the neuro-physiologically relevant conduction velocity and global coupling characteristics of each auditory condition ASSR, speech and melody. Thirdly, from a systematic investigation of the model predicted parameter space of conduction speeds, we found a dissociation of network transmission delays for speech and melody processing, stemming from the input asymmetry in spectro-temporal complexity of these two task contexts. This supports the existence of a tuning-by-delay mechanism in cortical networks for routing contextual information from identical stimuli into two hemispherically distinct homologous areas. In fact, such tuning-by-delay mechanism can explain not only the conundrum of lateralization of speech and melody but also in other domains such as auditory-motor coordination where right hand (controlled by the left hemisphere) seems more agile to fast finger tappings to auditory beat compared to left hand (controlled primarily by right hemisphere) which can follow slower frequency tappings more accurately (Pflug et al., 2019). Finally, even though there was a specific focus on the role of the primary auditory cortices (PAC) and their outflow

in governing hemispheric specialization, the modelling tool itself can be tweaked to predict other functional and effective connectivity measures as well and hence have tremendous application scenarios. In following sections, we discuss these issues in more detail.

Hierarchical and domain-specific organization of the auditory system

Several researchers have established that the auditory connectome is hierarchically organized (Kaas & Hackett, 2000; Rauschecker & Tian, 2000) with the PAC being the first cortical structure to receive ascending auditory inputs exhibiting a rich repertoire of oscillatory activity (Gourévitch et al., 2020; Hackett, 2011). Albouy and colleagues demonstrated the sensitivity of the left PAC to speech features and the right PAC to melody features (Albouy et al., 2020). On the other hand, there is substantial evidence of left hemispheric specialization for speech and right hemispheric specialization for music (Zatorre & Belin, 2001; Zatorre et al., 2002). In our study, the laterality profile of inflow in higher-order auditory regions from bilateral PAC complies with well-established hemispheric dominance of left hemisphere for speech, right hemisphere for melody and ASSR (Fig. 2 and Fig. 4B), suggesting observed hemispheric specializations could be attributed to the feed-forward flow from PAC. This contributes to our understanding of the hierarchical organization of auditory processing, as it reveals the flow of oscillatory information from lower-level sensory regions to higher-order cognitive regions. The oscillatory activity in PAC and the neural markers of its coordination with higher order areas seems to be specific to the nature of the auditory stimuli (Fig. 2). These specific communication dynamics across different regions are required for binding of disparate features, the segmentation of auditory streams, and the formation of hierarchical representations (Kösem & van Wassenhove, 2017; Plakke & Romanski, 2014).

Although the oscillatory patterns observed during speech and melody conditions are similar, characterized by frequency-specific outflows from the bilateral PAC, the specific regions involved in receiving these oscillations differ between speech and melody conditions. This suggests that the observed oscillatory patterns do not encode generic processing of acoustic signal dynamics but instead capture specific linguistic or musical features. Our experimental paradigm ensures the selective attention of participants to the speech or melody aspects of ecologically valid auditory stimuli (*a cappella* songs), thereby eliminating confounding effects of signal properties. The involvement of distinct regions in processing these acoustic environments indicates specialized and domain-specific organization within the auditory system (Angulo-Perkins & Concha, 2019; Peretz et al., 2015; Scott & McGettigan, 2013; Xie et al., 2018). Furthermore, the presence of homologous areas in the opposite hemisphere, such as the superior temporal

gyrus (STG), motor cortex, and Broca's area, suggests a bilateral representation and processing of auditory stimuli. Homologous areas share an evolutionary history and tend to perform similar functions, necessitating communication to coordinate their respective functions (Agcaoglu et al., 2018; Karolis et al., 2019; Wan et al., 2022). This parallel processing mechanism enables the brain to handle the complementary nature of speech and melody, two most important auditory communication skills. By distributing the cognitive load, the brain enables simultaneous processing when speech and melody are presented together, as in the case of ecologically valid *a cappella* songs (Angulo-Perkins & Concha, 2019). The identified respondent regions of the primary auditory cortices (PACs) in our study during both speech and melody conditions are consistent with previous literature and widely recognized as auditory-related regions (Giraud & Poeppel, 2012; Hickok & Poeppel, 2007; Morillon et al., 2010). The regions in the left hemisphere are known for processing various linguistic aspects, including phonological analysis, syntactic structure, semantic processing, and articulation. Conversely, the right hemisphere is implicated in processing non-linguistic aspects of auditory stimuli, such as melodic and tonal structures, pitch variations, temporal patterns, and rhythmic patterns in music (Zatorre & Belin, 2001; Zatorre & Gandour, 2008). Additionally, the ability of a single region to receive input at multiple frequencies highlights the brain's capacity to decode and integrate distinct aspects of oscillatory activity for coherent representation of auditory percept. This multiscale hierarchy integrates motor, syntactic, and working memory information, contributing to the construction of a comprehensive auditory scene. Our study establishes that functional integration at these cognitive levels can be identified using EEG activity with high temporal resolution, that opens up the path for studying many functions in detail.

Role of structural constraints in lateralized network dynamics

Structural symmetry of the brain, particularly the patterns of anatomical connectivity between brain regions, plays a crucial role in shaping the processing of information in brain networks (Honey et al., 2007; Sporns, 2010). In the auditory domain, auditory functions of speech and melody is postulated to be hinged in structural connectome (Mišić et al., 2018). In order to dig deeper into the mechanistic principles of such functional asymmetries, we compared the prediction of LI using brain network topology from real human brain to that of LI predicted from a network from a brain where connectivity is organized randomly, but frequencies, delay and coupling can be tuned (Fig. 3). The neural model incorporating subject specific structural connectivity (SC) while tuning for delay, brain response frequencies and couplings achieved the highest prediction accuracy in the 40 Hz ASSR condition (Fig. 3). The high model accuracy sheds

light to the fact that brain responses can be entrained to an oscillatory response mode that captures the collective dynamics of large-scale network via modulation of time delays (functional parameter) for a given structural connectivity (participant specific parameter). The high degree of predictability of LIs in homologous brain areas during ASSR response (ρ = [0.79 0.96]) gives us the candidates that the model inversion techniques will also be capable of detecting the response frequencies, coupling and transmission delays in situations where spectro-temporal complexity increases. Thus, all in all, the usage of biologically realistic network topology in the neurodynamic model expanded the capability of the models (Sanz Leon et al., 2013; Schirner, et al 2015) and also underlines the importance of structural connectivity shaping mode dynamics of auditory processing. Furthermore, high prediction in motor regions in beta frequencies could be attributed to the "hard-wiring" of auditory–motor interaction as reported earlier (Morillon et al., 2010) and further consolidates the role of structural asymmetries in shaping up function. Further the linear decrease in prediction accuracy across hierarchical brain regions (Fig. 5B) further support the general view that auditory processing areas follows a posterior to anterior hierarchy from simple to complex processing from primary sensory to complex cognitive areas. Although some recent advances in whole-brain modelling suggests geometry is also an important factor (Pang et al., 2023), however, such is currently out of scope of the present model but certainly remains a candidate for future explorations. Finally, the synaptic scaling used by brain regions can have structural origins owing to fibre densities among regions and functional origins measured via functional and effective connectivity. In our model, while the connection weights (C_{ij}) are estimated from fibre densities global coupling is a parameter that uniformly scales all functional associations such as in high and low attention scenarios, thus regulating synchronization strength. By modulating these factors, the brain can finely adjust the degree of coordination between neural signals across different areas.

Tuning by delay hypothesis for hemispheric lateralization

An important to note that brain structure is static in an observational window of sessions and days while functions are not. While basic features of auditory processing can be explained solely based on the network properties of structural connectivity, more abstract and higher-level features of the environment, such as meaning, context, and emotion require a more complex and dynamic interplay between different regions of the brain (Bauer et al., 2020; Berger et al., 2019; Di & Biswal, 2019). These findings hint that the higher-level processing or other factors beyond phase-based interactions contribute to propagation of delta and theta oscillations (Pandey et al., 2022). On the other hand, the processing of meaning in language

requires the coordination of neural activity between regions involved in phonetic, lexical, and semantic processing, as well as regions involved in attention, working memory, and executive control (Kazanina & Tavano, 2023). Similarly, the processing of emotion in music requires the coordination of neural activity between regions involved in auditory processing, pitch perception, and emotional valence (Gnanateja et al., 2022; Wang et al., 2023; Yurgil et al., 2020). Although, the distinct functional requirements are supported by the structural connectivity between these regions, nonetheless, brain accomplishes such higher cognitive processing by precisely manipulating the timing and coordination of neural signals through two key factors: conduction velocities and synaptic scaling (Pathak et al., 2022; Petkoski & Jirsa 2019). Conduction velocities, which controls the transmission delay of neural impulses travel along nerve fibers, play a pivotal role in governing the temporal aspects of neural communication. (Ghosh et al., 2008; Pflug et al., 2019; Pathak et al., 2022). Varied conduction velocities facilitate the emergence of distinct time delays, thereby shaping the arrival and integration of neural information across disparate regions (Cariani & Baker, 2022). This temporal coding ensures the precise synchronization necessary for intricate processing tasks (Ibrahim et al., 2021; Petkoski et al., 2018). Hence, despite the underlying static structural connectivity within the brain that changes over the time-scale of months or years, modulation by conduction velocities and coupling strength enables the generation of a diverse range of time delays and patterns of information propagation. Our simulation model revealed that differences in conduction velocities contribute to the functional lateralization observed in speech and melody processing (Fig 5). The faster conduction velocities in the left hemisphere, specialized for speech, enable the precise analysis of temporal changes associated with speech processing, while the slower conduction velocities meaning greater time window allows to capture fine spectral resolution of melody in the right hemisphere. These dichotomies could be attributed to enhanced myelination thereby enhanced transmission speed in the left hemisphere than right hemisphere (Seldon, 1981, 1982) which has also been reported for motor coordination (Pflug et al., 2019). Most importantly observed differences in conduction velocities between speech and melody stimuli in our study align with previous findings that speech exhibits rapid temporal changes and music shows fast spectral changes (Zatorre & Belin, 2001). The role of dichotomous or distinct categories within a system could be the causes of development of separation of cognitive processes or functional specialization observed in dichotomies. More over as suggested earlier, brain's parallel processing mechanism allows it to effectively handle the complementary aspects of speech and melody, which are two crucial forms of auditory communication. By distributing the cognitive load, the brain enables simultaneous processing when speech and melody are combined, as seen in ecologically valid a

cappella songs. Somewhat subtle, yet an important point to distinguish is while myelination may set in place the overall network tuning frequencies within a hemisphere, the functional input asymmetry decides the spatial geometry (left vs right) of lateralization.

In summary, we highlight the dichotomy or dualistic nature of the auditory system interactions at distinct levels. In the case of acoustic signals, viz., *a cappella* songs, the auditory system utilizes two divergent but related (temporal and spectral) cues embedded in a common signal to dissociate perceptual experiences in the form of speech and melody, highlighting the dichotomy of auditory system. In our empirical study, we show the presence of common oscillatory outflow in homologous brain regions and hints at "spatial" dichotomy. The common oscillatory signatures observed in the brain may represent shared underlying neural generators and gives rise to distinct hemispheric dominance. Furthermore, the variations in conduction velocities result in different temporal processing characteristics, enabling the specialized processing of speech and melody in their respective hemispheres highlighting a "temporal" dichotomies. Thus, our findings highlight the interplay between functional dynamics and structural connectivity in shaping auditory processing and importance of dissociation of spatial and temporal scales of processing. Hence, the present study contributes to our understanding of hemispheric specialization and provides an overarching mechanistic explanation of speech, melody, and tonal auditory perception in a common framework. Speech processing involves the encoding of fast temporal changes of auditory stimuli. On the other hand, the melodic processing of musical notes unfolds in slower time-scale to bind temporally far away events. Thus, the perception of speech and melody is determined by the organization of oscillatory activity in the PAC, as well as the subsequent propagation dynamics to downstream auditory networks along the input asymmetries which ultimately leads to hemispheric specialization.

References

Agcaoglu, O., Miller, R., Damaraju, E., Rashid, B., Bustillo, J., Cetin, M. S., Van Erp, T. G. M., McEwen, S., Preda, A., Ford, J. M., Lim, K. O., Manoach, D. S., Mathalon, D. H., Potkin, S. G., & Calhoun, V. D. (2018). Decreased hemispheric connectivity and decreased intra- and inter- hemisphere asymmetry of resting state functional network connectivity in schizophrenia. *Brain Imaging and Behavior*, *12*(3), 615–630. https://doi.org/10.1007/s11682-017-9718-7

Albouy, P., Benjamin1, L., Morillon, B., & Zatorre, R. J. (2020). Distinct sensitivity to spectrotemporal modulation supports brain asymmetry for speech and melody. *Science*, *367*(6481), 1043–1047. https://doi.org/10.1126/science.aaz3468

Angulo-Perkins, A., & Concha, L. (2019). Discerning the functional networks behind processing of music and speech through human vocalizations. *PLoS ONE*, *14*(10). https://doi.org/10.1371/journal.pone.0222796

Avena-Koenigsberger, A., Misic, B., & Sporns, O. (2017). Communication dynamics in complex brain networks. *Nature Reviews Neuroscience*, *19*(1), 17–33. https://doi.org/10.1038/nrn.2017.149

Bauer, A. K. R., Debener, S., & Nobre, A. C. (2020). Synchronisation of Neural Oscillations and Cross-modal Influences. In *Trends in Cognitive Sciences* (Vol. 24, Issue 6, pp. 481–495). Elsevier. https://doi.org/10.1016/j.tics.2020.03.003

Berezutskaya, J., Freudenburg, Z. V., Güçlü, U., van Gerven, M. A., & Ramsey, N. F. (2017). Neural tuning to low-level features of speech throughout the perisylvian cortex. Journal of Neuroscience, 37(33), 7906-7920. https://doi.org/10.1523/JNEUROSCI.0238-17.2017

Berger, B., Griesmayr, B., Minarik, T., Biel, A. L., Pinal, D., Sterr, A., & Sauseng, P. (2019). Dynamic regulation of interregional cortical communication by slow brain oscillations during working memory. *Nature Communications*, *10*(1), 1–11. https://doi.org/10.1038/s41467-019-12057-0

Broca, P. (1865). Sur le siège de la faculté du langage articulé. Bulletins et Mémoires de la Société d'Anthropologie de Paris, 6(1), 377-393. https://doi.org/10.3406/bmsap.1865.9495

Brovelli, A., Ding, M., Ledberg, A., Chen, Y., Nakamura, R., & Bressler, S. L. (2004). Beta oscillations in a large-scale sensorimotor cortical network: Directional influences revealed by Granger causality. *Proceedings of the National Academy of Sciences*, *101*(26), 9849–9854. https://doi.org/10.1073/pnas.0308538101

Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, *10*(3), 186–198. https://doi.org/10.1038/nrn2575

Byczkowicz, N., Eshra, A., Montanaro, J., Trevisiol, A., Hirrlinger, J., Kole, M. H., ... & Hallermann, S. (2019). HCN channel-mediated neuromodulation can control action potential velocity and fidelity in central axons. Elife, 8, e42766.https://doi.org/10.7554/eLife.42766

Cabral, J., Hugues, E., Sporns, O., & Deco, G. (2011). Role of local network oscillations in resting-state functional connectivity. *NeuroImage*, *57*(1), 130–139. https://doi.org/10.1016/j.neuroimage.2011.04.010

Cariani, P., & Baker, J. M. (2022). Time Is of the Essence: Neural Codes, Synchronies, Oscillations, Architectures. *Frontiers in Computational Neuroscience*, *16*, 898829. https://doi.org/10.3389/FNCOM.2022.898829/BIBTEX

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Dhamala, M., Rangarajan, G., & Ding, M. (2008). Analyzing information flow in brain networks with nonparametric Granger causality. *NeuroImage*, *41*(2), 354–362. https://doi.org/10.1016/j.neuroimage.2008.02.020

Di, X., & Biswal, B. B. (2019). Toward Task Connectomics: Examining Whole-Brain Task Modulated Connectivity in Different Task Domains. *Cerebral Cortex*, *29*(4), 1572–1583. https://doi.org/10.1093/cercor/bhy055

Ding, M., Chen, Y., & Bressler, S. L. (2006). Granger Causality: Basic Theory and Application to Neuroscience. *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*, *February*, 437–460. https://doi.org/10.1002/9783527609970.ch17

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2015). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*(1), 158–164. https://doi.org/10.1038/nn.4186

Eggermont, J. J. (1998). Representation of spectral and temporal sound features in three cortical fields of the cat. Similarities outweigh differences. Journal of Neurophysiology, 80(5), 2743-2764. https://doi.org/10.1152/jn.1998.80.5.2743

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474–480. https://doi.org/10.1016/j.tics.2005.08.011

Fuchs, M., Wagner, M., & Kastner, J. (2001). Boundary element method volume conductor models for

EEG source reconstruction. *Clinical Neurophysiology*, *112*(8), 1400–1407. https://doi.org/10.1016/S1388-2457(01)00589-2

Gerrits, R., Van der Haegen, L., Brysbaert, M., & Vingerhoets, G. (2019). Laterality for recognizing written words and faces in the fusiform gyrus covaries with language dominance. Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 117, 196–204. https://doi.org/10.1016/j.cortex.2019.03.010

Geschwind, N., & Galaburda, A. M. (1985). Cerebral lateralization: Biological mechanisms, associations, and pathology: I. A hypothesis and a program for research. Archives of neurology, 42(5), 428-459. https://doi.org/10.1001/archneur.1985.04060050026008

Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, *77*(378), 304–313. https://doi.org/10.1080/01621459.1982.10477803

Ghosh, A., Rho, Y., McIntosh, A. R., Kötter, R., & Jirsa, V. K. (2008). Noise during Rest Enables the Exploration of the Brain's Dynamic Repertoire. *PLoS Computational Biology*, *4*(10), e1000196. https://doi.org/10.1371/journal.pcbi.1000196

Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. In *Nature Neuroscience* (Vol. 15, Issue 4, pp. 511–517). Nature Publishing Group. https://doi.org/10.1038/nn.3063

Gnanateja, G. N., Devaraju, D. S., Heyne, M., Quique, Y. M., Sitek, K. R., Tardif, M. C., Tessmer, R., & Dial, H. R. (2022). On the Role of Neural Oscillations Across Timescales in Speech and Music Processing. *Frontiers in Computational Neuroscience*, *16*, 872093. https://doi.org/10.3389/FNCOM.2022.872093/BIBTEX

Gourévitch, B., Martin, C., Postal, O., & Eggermont, J. J. (2020). Oscillations in the auditory system and their possible role. In *Neuroscience and Biobehavioral Reviews* (Vol. 113, pp. 507–528). Pergamon. https://doi.org/10.1016/j.neubiorev.2020.03.030

Gotts, S. J., Jo, H. J., Wallace, G. L., Saad, Z. S., Cox, R. W., & Martin, A. (2013). Two distinct forms of functional lateralization in the human brain. Proceedings of the National Academy of Sciences, 110(36), E3435-E3444. https://doi.org/10.1073/pnas.1302581110

Gramfort, A., Papadopoulo, T., Olivi, E., & Clerc, M. (2010). OpenMEEG: Opensource software for quasistatic bioelectromagnetics. *BioMedical Engineering Online*, *9*(1), 1–20. https://doi.org/10.1186/1475-925X-9-45/TABLES/2

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, *37*(3), 424. https://doi.org/10.2307/1912791

Güntürkün, O., Ströckens, F., & Ocklenburg, S. (2020). Brain lateralization: a comparative perspective. Physiological reviews. https://doi.org/10.1152/physrev.00006.2019

Güntürkün, O., & Ocklenburg, S. (2017). Ontogenesis of lateralization. Neuron, 94(2), 249-263. https://doi.org/10.1016/j.neuron.2017.02.045

Hackett, T. A. (2011). Information flow in the auditory cortical network. *Hearing Research*, *271*(1–2), 133–146. https://doi.org/10.1016/j.heares.2010.01.011

Halder, T., Talwar, S., Jaiswal, A. K., & Banerjee, A. (2019). Quantitative evaluation in estimating sources underlying brain oscillations using current source density methods and beamformer approaches. *ENeuro*, *6*(4). https://doi.org/10.1523/ENEURO.0170-19.2019

Hecht, D. (2010). Depression and the hyperactive right-hemisphere. Neuroscience research, 68(2), 77-87. https://doi.org/10.1016/j.neures.2010.06.013

Hellige, J. B. (2001). Hemispheric asymmetry: What's right and what's left (Vol. 6). Harvard University Press.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. In *Nature Reviews Neuroscience* (Vol. 8, Issue 5, pp. 393–402). Nature Publishing Group. https://doi.org/10.1038/nrn2113

Honey, C. J., Kötter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(24), 10240–10245. https://doi.org/10.1073/pnas.0701519104

Ibrahim, M. M., Kamran, M. A., Mannan, M. M. N., Jung, I. H., & Kim, S. (2021). Lag synchronization of coupled time-delayed FitzHugh–Nagumo neural networks via feedback control. *Scientific Reports*, *11*(1), 1–15. https://doi.org/10.1038/s41598-021-82886-x

Jirsa, V. K., & Ding, M. (2004). Will a large complex system with time delays be stable?. Physical review letters, 93(7), 070602. https://doi.org/10.1103/PhysRevLett.93.070602

Kaas, J. H., & Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences*, *97*(22), 11793–11799. https://doi.org/10.1073/pnas.97.22.11793

Karolis, V. R., Corbetta, M., & Thiebaut de Schotten, M. (2019). The architecture of functional lateralisation and its relationship to callosal connectivity in the human brain. *Nature*

Communications, *10*(1), 1–9. https://doi.org/10.1038/s41467-019-09344-1

Kazanina, N., & Tavano, A. (2023). What neural oscillations can and cannot do for syntactic structure building. *Nature Reviews Neuroscience*, *24*(2), 113–128. https://doi.org/10.1038/s41583-022-00659-5

Kösem, A., & van Wassenhove, V. (2017). Distinct contributions of low- and high-frequency neural oscillations to speech comprehension. *Language, Cognition and Neuroscience*, *32*(5), 536–544. https://doi.org/10.1080/23273798.2016.1238495

Kong, X. Z., Mathias, S. R., Guadalupe, T., ENIGMA Laterality Working Group, Glahn, D. C., Franke, B., ... & Karolinska Schizophrenia Project (KaSP) consortium. (2018). Mapping cortical brain asymmetry in 17,141 healthy individuals worldwide via the ENIGMA Consortium. Proceedings of the National Academy of Sciences, 115(22), E5154-E5163. https://doi.org/10.1073/pnas.1718418115

Kumar, N., Jaiswal, A., Roy, D., & Banerjee, A. (2023). Effective networks mediate right hemispheric dominance of human 40 Hz auditory steady-state response. *Neuropsychologia*, *184*, 108559. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2023.108559

Kuramoto, Y. (1984). *Chemical Oscillations, Waves, and Turbulence* (Vol. 19). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-69689-3

Liu, H., Stufflebeam, S. M., Sepulcre, J., Hedden, T., & Buckner, R. L. (2009). Evidence from intrinsic activity that asymmetry of the human brain is controlled by multiple factors. Proceedings of the national academy of sciences, 106(48), 20499-20503. https://doi.org/10.1073/pnas.0908073106

Meena, C., Hens, C., Acharyya, S., Haber, S., Boccaletti, S., & Barzel, B. (2023). Emergent stability in complex network dynamics. Nature Physics, 1-10. https://doi.org/10.1038/s41567-023-02020-8

Mišić, B., Betzel, R.F., Griffa, A., De Reus, M.A., He, Y., Zuo, X.N., Van Den Heuvel, M.P., Hagmann, P., Sporns, O. and Zatorre, R.J. (2018). Network-based asymmetry of the human auditory system. Cerebral Cortex, 28(7), 2655-2664. https://doi.org/10.1093/cercor/bhy101

Morillon, B., Lehongre, K., Frackowiak, R. S. J., Ducorps, A., Kleinschmidt, A., Poeppel, D., & Giraud, A. L. (2010). Neurophysiological origin of human brain asymmetry for speech and language. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(43), 18688–18693. https://doi.org/10.1073/pnas.1007189107

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence*

and Neuroscience, *2011*, 156869. https://doi.org/10.1155/2011/156869

Pandey, P., Miyapuram, K. P., & Lomas, D. (2022). *Non-Linear Features of β Brain Rhythms Predict Listener-Specific Neural Signature in Naturalistic Music Listening*. https://doi.org/10.36227/TECHRXIV.20502423.V1

Pang, J. C., Aquino, K. M., Oldehinkel, M., Robinson, P. A., Fulcher, B. D., Breakspear, M., & Fornito, A. (2023). Geometric constraints on human brain function. Nature, 1-9. https://doi.org/10.1038/s41586-023-06098-1

Pascual-Marqui, R. D. (2007). *Discrete, 3D distributed, linear imaging methods of electric neuronal activity. Part 1: exact, zero error localization*. http://arxiv.org/abs/0710.3341

Pathak, A., Roy, D., & Banerjee, A. (2022). Whole-brain network models: from physics to bedside. Frontiers in Computational Neuroscience, 16, 866517. https://doi.org/10.3389/fncom.2022.866517

Peretz, I., Vuvan, D., Lagrois, M. É., & Armony, J. L. (2015). Neural overlap in processing music and speech. In *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 370, Issue 1664, p. 20140090). The Royal Society. https://doi.org/10.1098/rstb.2014.0090

Petkoski, S., Palva, J. M., & Jirsa, V. K. (2018). Phase-lags in large scale brain synchronization: Methodological considerations and in-silico analysis. *PLoS Computational Biology*, *14*(7), e1006160. https://doi.org/10.1371/journal.pcbi.1006160

Petkoski, S., & Jirsa, V. K. (2019). Transmission time delays organize the brain network synchronization. Philosophical Transactions of the Royal Society A, 377(2153), 20180132. https://doi.org/10.1098/rsta.2018.0132

Pflug, A., Gompf, F., Muthuraman, M., Groppa, S., & Kell, C. A. (2019). Differential contributions of the two human cerebral hemispheres to action timing. elife, 8, e48404. https://doi.org/10.7554/eLife.48404

Picton, T. W. (2013). Hearing in Time: Evoked Potential Studies of Temporal Processing. *Ear & Hearing*, *34*, 385–401. https://doi.org/10.1097/AUD.0b013e31827ada02

Plakke, B., & Romanski, L. M. (2014). Auditory connections and functions of prefrontal cortex. *Frontiers in Neuroscience*, *8 JUL*. https://doi.org/10.3389/fnins.2014.00199

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in Cognitive Sciences*, *1*(2), 56–61. https://doi.org/10.1016/S1364-6613(97)01008-5

Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of “what” and “where” in auditory cortex. In *Proceedings of the National Academy of Sciences of the United States of America*

(Vol. 97, Issue 22, pp. 11800–11806). Proc Natl Acad Sci U S A. https://doi.org/10.1073/pnas.97.22.11800

Roberts, J. A., Gollo, L. L., Abeysuriya, R. G., Roberts, G., Mitchell, P. B., Woolrich, M. W., & Breakspear, M. (2019). Metastable brain waves. Nat Commun 10: 1056. https://doi.org/10.1038/s41467-019-08999-0

Sanz Leon, P., Knock, S. A., Woodman, M. M., Domide, L., Mersmann, J., McIntosh, A. R., & Jirsa, V. (2013). The Virtual Brain: a simulator of primate brain network dynamics. Frontiers in neuroinformatics, 7, 10. https://doi.org/10.3389/fninf.2013.00010

Schirner, M., Rothmeier, S., Jirsa, V. K., McIntosh, A. R., & Ritter, P. (2015). An automated pipeline for constructing personalized virtual brains from multimodal neuroimaging data. NeuroImage, 117, 343-357. https://doi.org/10.1016/j.neuroimage.2015.03.055

Scott, S. K., & McGettigan, C. (2013). Do temporal processes underlie left hemisphere dominance in speech perception? *Brain and Language*, *127*(1), 36–45. https://doi.org/10.1016/j.bandl.2013.07.006

Seghier, M. L., & Price, C. J. (2018). Interpreting and Utilising Intersubject Variability in Brain Function. In *Trends in Cognitive Sciences* (Vol. 22, Issue 6, pp. 517–530). Elsevier. https://doi.org/10.1016/j.tics.2018.03.003

Seldon, H. L. (1981). Structure of human auditory cortex. I. Cytoarchitectonics and dendritic distributions. *Brain Research*, *229*(2), 277–294. https://doi.org/10.1016/0006-8993(81)90994-X

Seldon, H. L. (1982). Structure of human auditory cortex. III. Statistical analysis of dendritic trees. *Brain Research*, *249*(2), 211–221. https://doi.org/10.1016/0006-8993(82)90055-5

Sperry, R. W. (1974) Lateral specialization in the surgically separated hemispheres. In: F. Schmitt and F. Worden (Eds.), Neurosciences Third Study Program, Ch. I, Vol. 3, pp. 5- 19. Cambridge: MIT Press.

Sporns, O. (2010). Analysis and Function of Large-Scale Brain Networks. *Society for Neuroscience Short Course III*, 1–68. https://www.sfn.org/~/media/SfN/Documents/Short Courses/2010 Short Course III/Short Course Book III.ashx

Sporns, O. (2011). The human connectome: a complex network. *Annals of the New York Academy of Sciences*, *1224*(1), 109–125. https://doi.org/10.1111/j.1749-6632.2010.05888.x

Stam, C. J., Nolte, G., & Daffertshofer, A. (2007). Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Human Brain Mapping*, *28*(11), 1178–1193. https://doi.org/10.1002/HBM.20346

Stange-Marten, A., Nabel, A. L., Sinclair, J. L., Fischl, M., Alexandrova, O., Wohlfrom, H., ... & Grothe,

B. (2017). Input timing for spatial processing is precisely tuned via constant synaptic delays and myelination patterns in the auditory brainstem. Proceedings of the National Academy of Sciences, 114(24), E4851-E4858. https://doi.org/10.1073/pnas.1702290114

Toga, A. W., & Thompson, P. M. (2003). Mapping brain asymmetry. Nature Reviews Neuroscience, 4(1), 37-48. https://doi.org/10.1038/nrn1009

Vinck, M., Batista-Brito, R., Knoblich, U., & Cardin, J. A. (2015). Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. Neuron, 86(3), 740-754. https://doi.org/10.1016/j.neuron.2015.03.028

Wan, B., Bayrak, Ş., Xu, T., Schaare, H. L., Bethlehem, R. Al, Bernhardt, B. C., & Valk, S. L. (2022). Heritability and cross-species comparisons of human cortical functional organization asymmetry. *ELife*, *11*. https://doi.org/10.7554/ELIFE.77215

Wang, Y., Siu, T. S. C., & Cheung, H. (2023). Effect of music emotion on mu and beta oscillations. *Psychology of Music*. https://doi.org/10.1177/03057356221145960

Weineck, K., Wen, O. X., & Henry, M. J. (2022). Neural synchronization is strongest to the spectral flux of slow music and depends on familiarity and beat salience. Elife, 11, e75515. https://doi.org/10.7554/eLife.75515

Xie, H., Calhoun, V. D., Gonzalez-Castillo, J., Damaraju, E., Miller, R., Bandettini, P. A., & Mitra, S. (2018). Whole-brain connectivity dynamics reflect both task-specific and individual-specific modulation: A multitask study. In *NeuroImage* (Vol. 180, pp. 495–504). Academic Press. https://doi.org/10.1016/j.neuroimage.2017.05.050

Woodhead, Z. V. J., Wise, R. J. S., Sereno, M., & Leech, R. (2011). Dissociation of sensitivity to spatial frequency in word and face preferential areas of the fusiform gyrus. Cerebral Cortex, 21(10), 2307–2312. https://doi.org/10.1093/cercor/bhr008

Yurgil, K. A., Velasquez, M. A., Winston, J. L., Reichman, N. B., & Colombo, P. J. (2020). Music Training, Working Memory, and Neural Oscillations: A Review. *Frontiers in Psychology*, *11*, 478420. https://doi.org/10.3389/FPSYG.2020.00266/BIBTEX

Zatorre, R J, & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex (New York, N.Y. : 1991)*, *11*(10), 946–953. https://doi.org/10.1093/cercor/11.10.946

Zatorre, Robert J. (2022). Hemispheric asymmetries for music and speech: Spectrotemporal modulations and top-down influences. In *Frontiers in Neuroscience* (Vol. 16, p. 1075511). Frontiers Media S.A. https://doi.org/10.3389/fnins.2022.1075511

Zatorre, Robert J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, *6*(1), 37–46. https://doi.org/10.1016/S1364-6613(00)01816-7

Zatorre, Robert J, & Gandour, J. T. (2008). Neural specializations for speech and pitch: Moving beyond the dichotomies. In *Philosophical Transactions of the Royal Society B: Biological Sciences* Vol. 363, Issue 1493, pp. 1087–1104. https://doi.org/10.1098/rstb.2007.2161

Figure captions

Figure 1: Overview of the methodology: The figure illustrates the pipeline of study, incorporating both empirical and theoretical analyses. The left panel depicts the empirical analysis, highlighting the experimental setup and data processing steps involved in recording neural activity using electroencephalography (EEG) while participants selectively attended to speech or melody stimuli. For visualization purpose, after source localization only ASSR analysis was represented. The right panel represents the theoretical analysis, showcasing the averaged structural connectivity (SC) network that was derived from diffusion weighted imaging (dWI) data. This network served as a constraint for a neural dynamic model to simulate the frequency-specific outflow from the primary auditory cortices from which lateralization indices (LI) can be calculated for specific brain areas.

Figure 2: Frequency-specific outflow and target regions of bilateral PAC
The left panel displays the averaged Granger causality spectra, representing outflow from the bilateral primary auditory cortices (PAC) to the target regions. The right panel presents the activation maps of the target regions, averaged for the delta, theta, beta, and gamma frequency bands. For the ASSR condition, only the 40 Hz frequency is represented, reflecting the frequency-specific neural response observed at this frequency.

Figure 3: Prediction accuracy of frequency simulation (f (Hz)), global coupling (k), and global delay ($<\tau>$): Distributions of prediction accuracy of free parameters ASSR, ASSR with shuffled SC, speech, and melody conditions. The x-axis represents the different simulation parameters, while the y-axis represents the prediction accuracy. The first column in each condition demonstrates the effectiveness of the frequency simulation in accurately predicting the neural responses. The last panel shows the correlation coefficients of respective predictions.

Figure 4: Lateralization prediction accuracy of individual regions: (A) Scatter plot shows the linear fitting of empirical LI (Y-axis) against predicted LI (X-axis). Each dot represents the LI of one subject. The analysis focuses on beta and gamma frequencies for speech and melody conditions, as well as 40 Hz for the ASSR condition. (B) Confidence intervals for LI, the confidence intervals for the LI during empirical (blue) and predicted (orange) analyses. The shaded background indicates cases where both empirical and predicted LIs fall on either side of zero, indicating accurate prediction of the lateralization direction.

Figure 5: (A) Distribution of global coupling and global delay for beta and gamma frequencies in speech and melody conditions. The distributions are visualized using kernel smoothing with a bandwidth of 0.3 to enhancing the clarity of the patterns. (B) Decrease in prediction accuracy along the auditory hierarchy in gamma range in the speech and melody conditions. (C) Prediction accuracy in homologous auditory areas segregate along sensory processing hierarchies for speech and melody processing.

Figure 1

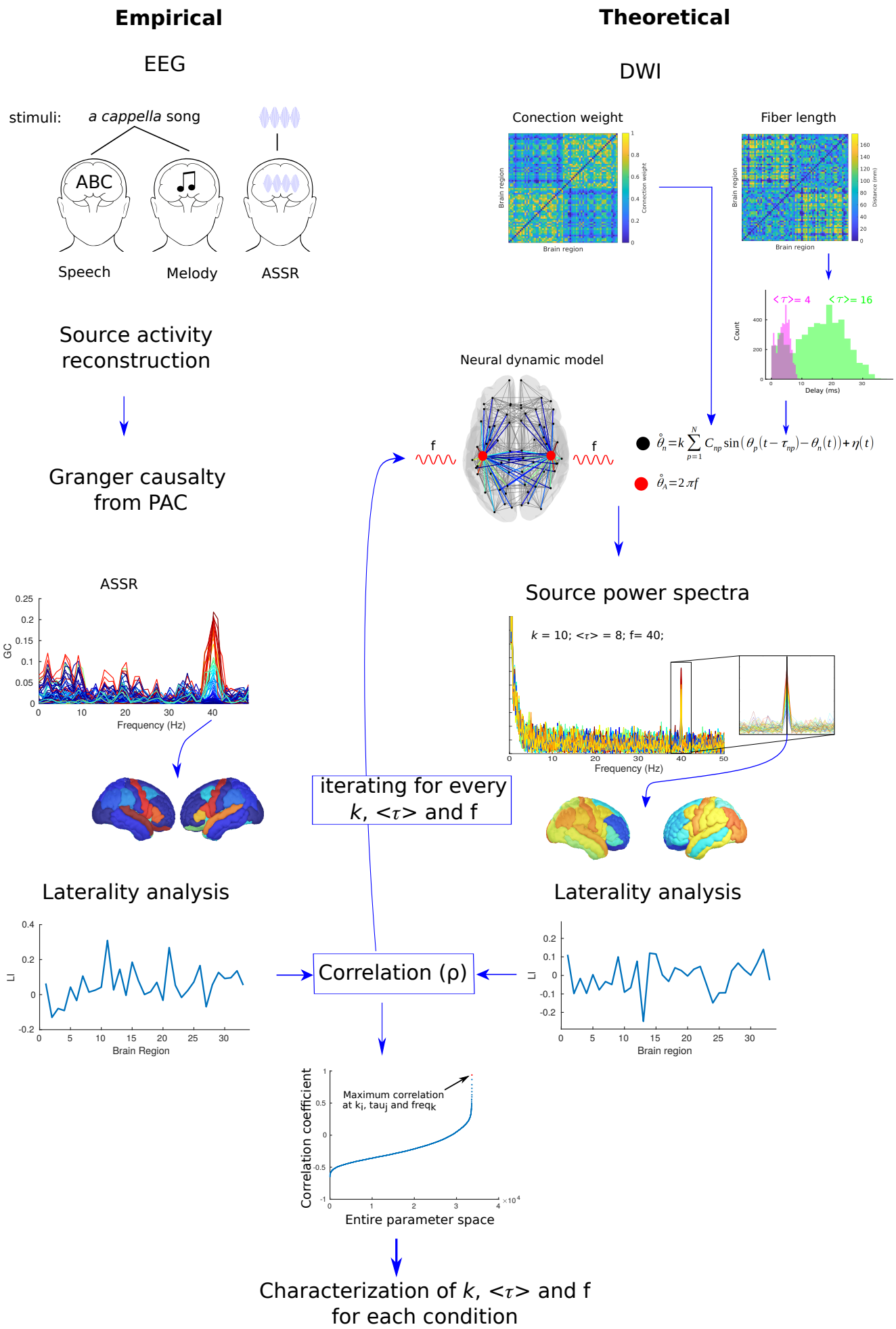
Empirical
Theoretical
EEG
DWI
stimuli:
a cappella song
ABC
Speech
Melody
ASSR
Conection weight
Fiber length
Brain region
Connection weight
Distance (mm)
<τ>= 4
<τ>= 16
Count
Delay (ms)
Source activity reconstruction
Neural dynamic model
f
f
$\dot{\theta}_n = k\sum_{p=1}^{N} C_{np}\sin(\theta_p(t-\tau_{np})-\theta_n(t))+\eta(t)$
$\dot{\theta}_A = 2\pi f$
Granger causalty from PAC
ASSR
GC
Frequency (Hz)
Source power spectra
k = 10; <τ> = 8; f= 40;
Frequency (Hz)
iterating for every k, <τ> and f
Laterality analysis
Laterality analysis
LI
Brain Region
Correlation (ρ)
LI
Brain region
Correlation coefficient
Maximum correlation at k_i, tau_j and $freq_k$
Entire parameter space
Characterization of k, <τ> and f for each condition

Figure 2

Inflow from bilateral PAC

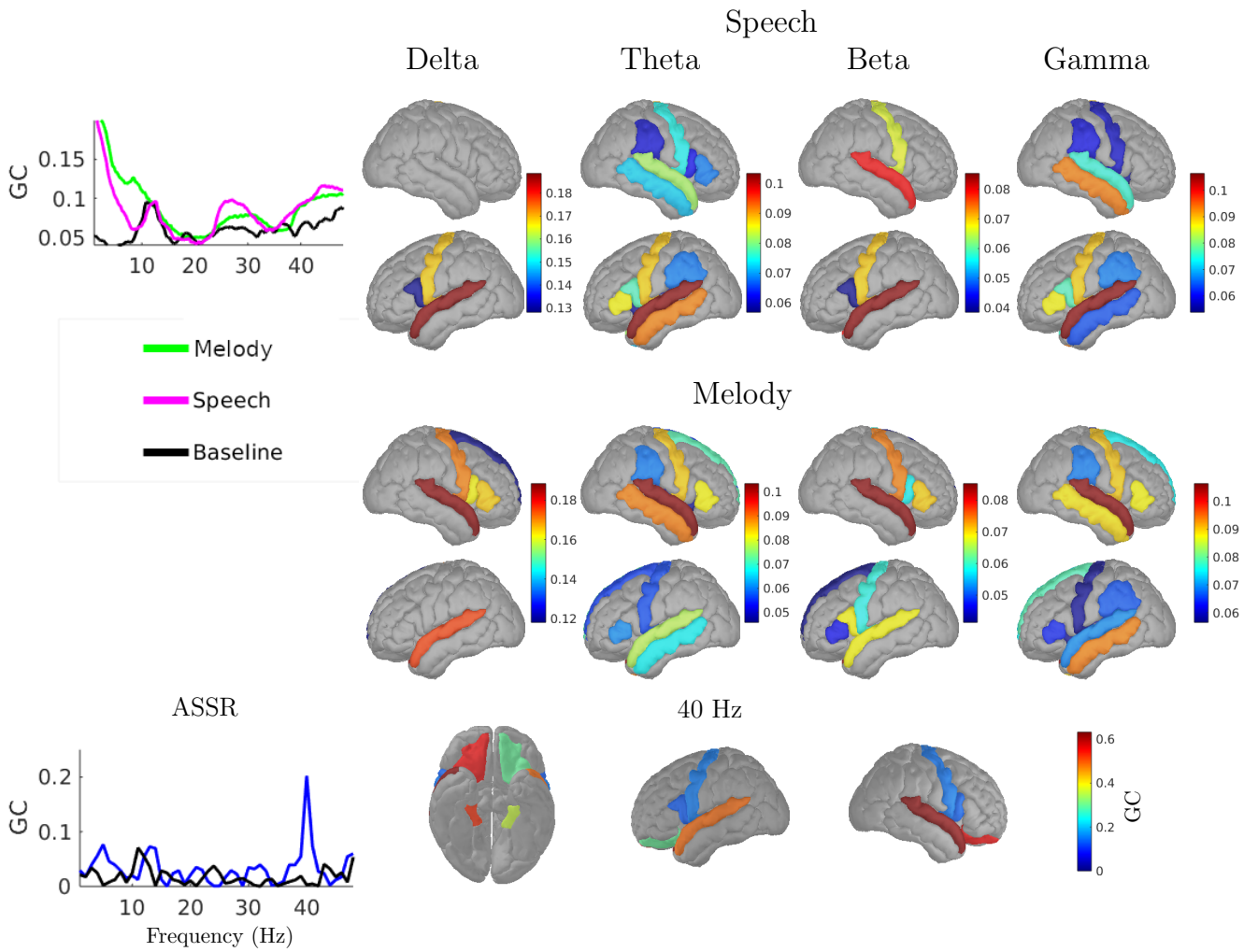

Figure 3

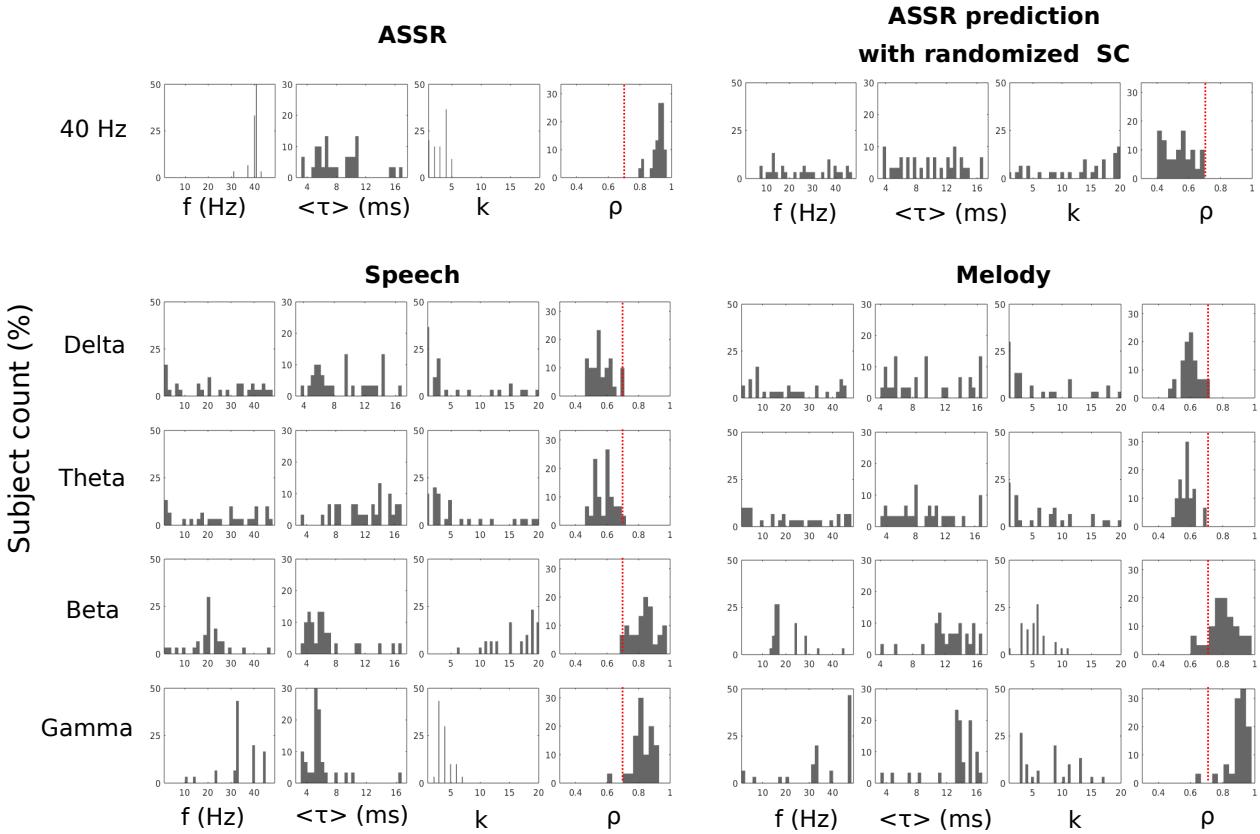
ASSR
ASSR prediction with randomized SC
40 Hz
Speech
Melody
Delta
Theta
Beta
Gamma
Subject count (%)
f (Hz)
<τ> (ms)
k
ρ

Figure 4

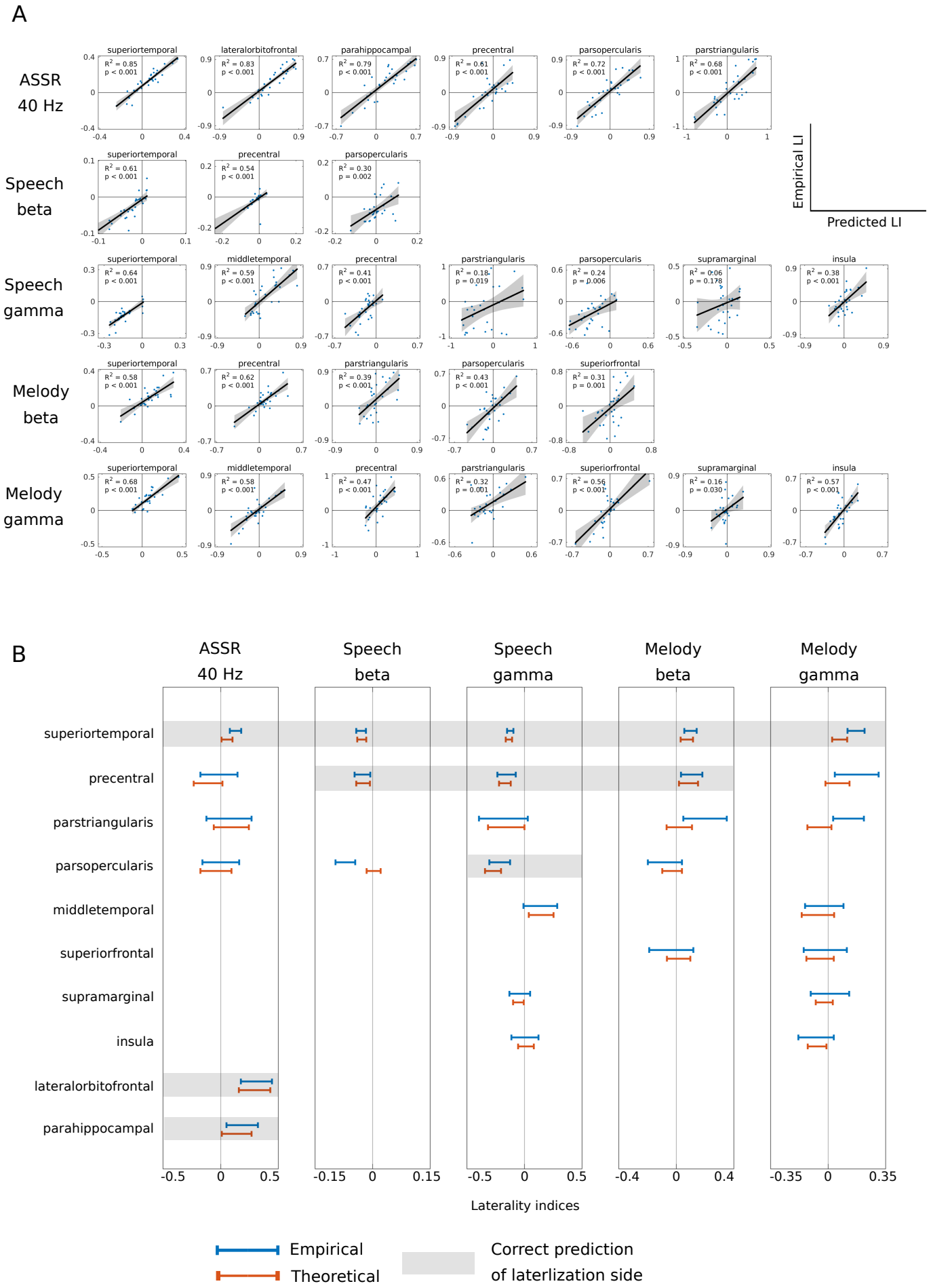
A
ASSR 40 Hz
superiortemporal R² = 0.85 p < 0.001
lateralorbitofrontal R² = 0.83 p < 0.001
parahippocampal R² = 0.79 p < 0.001
precentral R² = 0.61 p < 0.001
parsopercularis R² = 0.72 p < 0.001
parstriangularis R² = 0.68 p < 0.001
Speech beta
superiortemporal R² = 0.61 p < 0.001
precentral R² = 0.54 p < 0.001
parsopercularis R² = 0.30 p = 0.002
Empirical LI
Predicted LI
Speech gamma
superiortemporal R² = 0.64 p < 0.001
middletemporal R² = 0.59 p < 0.001
precentral R² = 0.41 p < 0.001
parstriangularis R² = 0.18 p = 0.019
parsopercularis R² = 0.24 p = 0.006
supramarginal R² = 0.06 p = 0.178
insula R² = 0.38 p < 0.001
Melody beta
superiortemporal R² = 0.58 p < 0.001
precentral R² = 0.62 p < 0.001
parstriangularis R² = 0.39 p < 0.001
parsopercularis R² = 0.43 p < 0.001
superiorfrontal R² = 0.31 p = 0.001
Melody gamma
superiortemporal R² = 0.68 p < 0.001
middletemporal R² = 0.58 p < 0.001
precentral R² = 0.47 p < 0.001
parstriangularis R² = 0.32 p = 0.001
superiorfrontal R² = 0.56 p < 0.001
supramarginal R² = 0.16 p = 0.030
insula R² = 0.57 p < 0.001
B
ASSR 40 Hz
Speech beta
Speech gamma
Melody beta
Melody gamma
superiortemporal
precentral
parstriangularis
parsopercularis
middletemporal
superiorfrontal
supramarginal
insula
lateralorbitofrontal
parahippocampal
-0.5 0 0.5
-0.15 0 0.15
-0.5 0 0.5
-0.4 0 0.4
-0.35 0 0.35
Laterality indices
Empirical
Theoretical
Correct prediction of laterlization side

Figure 5

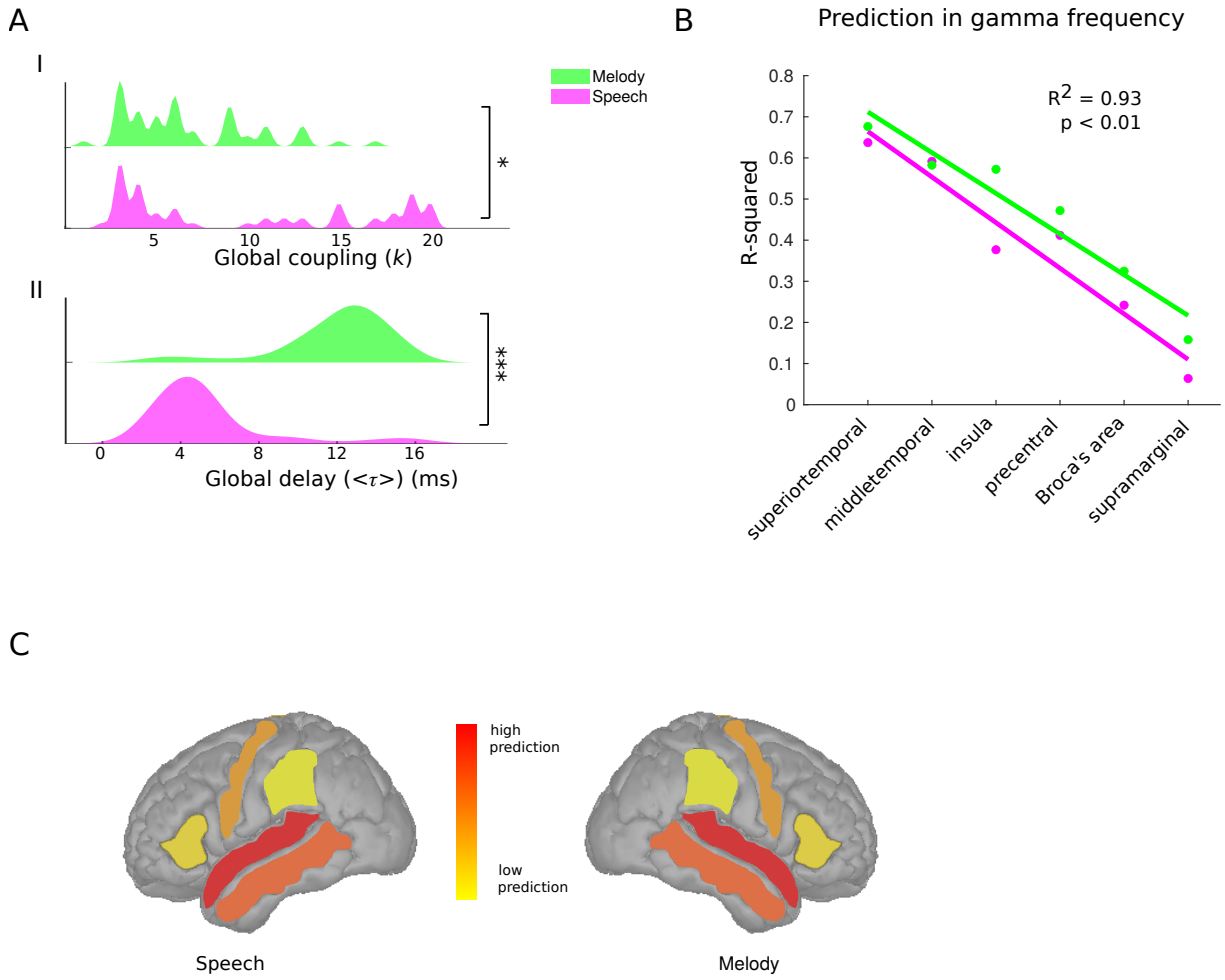
A
I
Melody
Speech
*
5
10
15
20
Global coupling (k)
II
***
0
4
8
12
16
Global delay (<τ>) (ms)
B
Prediction in gamma frequency
R-squared
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
R2 = 0.93
p < 0.01
superiortemporal
middletemporal
insula
precentral
Broca's area
supramarginal
C
high prediction
low prediction
Speech
Melody

Table 1: Frequncy-specific target regions of bilateral primary auditory cortices during different auditory conditions.

Condition	Frequency Band	Regions
Speech	Delta	Left Pars Opercularis, Left Precentral, Left Superior temporal
	Theta	Left Middle temporal lobe, Left Pars Opercularis, Left Pars Triangularis, Left Precentral, Left Superior temporal, Left Supramarginal, Left Insula, Right Middle temporal lobe, Right Pars Opercularis, Right Pars Triangularis, Right Precentral, Right Superior temporal
	Beta	Left Pars Opercularis, Left Precentral, Left Superior temporal, Right Precentral, Right Superior temporal
	Gamma	Left Middle temporal lobe, Left Pars Opercularis, Left Pars Triangularis, Left Precentral, Left Superior temporal, Left Supramarginal, Right Middle temporal lobe, Right Precentral, Right Superior temporal, Right Supramarginal, Right Insula
Melody	Delta	Left Superior temporal, Right Pars Opercularis, Right Pars Triangularis, Right Precentral, Right Superior frontal, Right Superior temporal
	Theta	Left Middle temporal lobe, Left Pars Triangularis, Left Precentral, Left Superior frontal, Left Superior temporal, Right Middle temporal lobe, Right Pars Triangularis, Right precentral, Right Superior frontal, Right Superior temporal, Right Supramarginal, Right Insula
	Beta	Left Pars Opercularis, Left Pars Triangularis, Left Precentral, Left Superior frontal, Left Superior temporal, Right Pars Opercularis, Right Pars Triangularis, Right Precentral, Right Superior temporal
	Gamma	Left Middle temporal lobe, Left Pars Triangularis, Left Precentral, Left Superior frontal, Left Superior temporal, Left Supramarginal, Left Insula, Right Middle temporal lobe, Right Pars Triangularis, Right Precentral, Right Superior frontal, Right Superior temporal
ASSR	40 Hz	Left Pars Opercularis, Left Precentral, Right Lateral orbitofrontal cortex, Right Parahippocampal, Right Pars Opercularis, Right Precentral, Right Superior temporal

Table 2: 95% Confidence Intervals (CI) of Laterality Indices during Speech, Melody, and ASSR Conditions.

	ASSR			**Speech beta**			**Speech gamma**			**Melody beta**			**Melody gamma**		
	Lo CI	Up CI	Mean	Lo CI	Up CI	Mean	Lo CI	Up CI	Mean	Lo CI	Up CI	Mean	Lo CI	Up CI	Mean
Superior temporal	0.08	0.18	0.13	-0.04	-0.02	-0.03	-0.15	-0.10	-0.13	0.06	0.14	0.10	0.12	0.22	0.17
Precentral	-0.18	0.14	-0.02	-0.05	-0.01	-0.03	-0.24	-0.08	-0.16	0.03	0.18	0.11	0.04	0.31	0.17
Pars triangularis	-0.13	0.27	0.07	_	_	_	-0.40	0.03	-0.19	0.05	0.35	0.20	0.03	0.22	0.12
Pars opercularis	-0.16	0.16	0.00	-0.10	-0.05	-0.07	-0.31	-0.13	-0.22	-0.20	0.04	-0.08	_	_	_
Middle temporal	_	_	_	_	_	_	-0.01	0.29	0.14	_	_	_	-0.14	0.09	-0.02
Superior frontal	_	_	_	_	_	_	_	_	_	-0.19	0.12	-0.04	-0.15	0.11	-0.02
Supramarginal	_	_	_	_	_	_	-0.13	0.05	-0.04	_	_	_	-0.11	0.13	0.01
Insula	_	_	_	_	_	_	-0.12	0.12	0.00	_	_	_	-0.18	0.03	-0.07
Lateral orbitofrontal	0.17	0.44	0.31	_	_	_	_	_	_	_	_	_	_	_	_
Parahippocampal	0.05	0.32	0.19	_	_	_	_	_	_	_	_	_	_	_	_

Table 3: Prediction specificity and normality of free parameters.

	95 % Confidence Interval						**Normality test**			
	Frequency		**k**		**tau**		**k**		**tau**	
	Lo CI	Up CI	Lo CI	Up CI	Lo CI	Up CI	kstat	*p* value	kstat	*p* value
Control ASSR	21.6	31.1	10.6	15.4	1.8	2.6	0.16	0.042	0.13	0.195
ASSR	39.4	40.9	2.5	3.5	1.5	2.2	0.24	> 0.001	0.16	0.043
Speech delta	17.6	29.8	3.2	7.7	1.7	2.4	0.32	> 0.001	0.15	0.091
Melody delta	15.7	26.9	3.8	8.4	1.8	2.6	0.26	> 0.001	0.15	0.093
Speech theta	19.2	30.8	3.8	8.4	2.4	3.1	0.27	> 0.001	0.12	0.339
Melody theta	16.0	28.1	4.6	8.9	1.7	2.4	0.19	> 0.001	0.12	0.339
Speech beta	17.2	23.7	14.8	17.5	1.1	1.8	0.19	0.006	0.28	> 0.001
Melody beta	17.6	22.8	4.7	6.3	2.6	3.1	0.18	0.015	0.18	0.010
Speech gamma	31.0	41.7	6.2	9.4	2.7	3.3	0.24	> 0.001	0.29	> 0.001
Melody gamma	31.1	37.4	3.5	4.3	1.0	1.5	0.18	0.014	0.37	> 0.001