NeuroImage 202 (2019) 116042

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Brain dynamics and connectivity networks under natural auditory stimulation

Po-Chih Kuo [a], Yi-Li Tseng [b], Karl Zilles [c], Summit Suen [a], Simon B. Eickhoff [d,e], Juin-Der Lee [f], Philip E. Cheng [a], Michelle Liou [a,*]

[a] Institute of Statistical Science, Academia Sinica, Taipei, Taiwan
[b] Department of Electrical Engineering, Fu Jen Catholic University, New Taipei City, Taiwan
[c] Institute of Neuroscience and Medicine (INM-1), Research Centre Jülich, Jülich, Germany
[d] Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
[e] Institute of Neuroscience and Medicine (INM-7), Research Centre Jülich, Jülich, Germany
[f] Graduate Institute of Business Administration, National Chengchi University, Taipei, Taiwan

ARTICLE INFO

Keywords:
Cytoarchitectonic maps
fMRI
Intraclass correlation
Natural auditory stimulation
Stationarity

ABSTRACT

The analysis of functional magnetic resonance imaging (fMRI) data is challenging when subjects are under exposure to natural sensory stimulation. In this study, a two-stage approach was developed to enable the identification of connectivity networks involved in the processing of information in the brain under natural sensory stimulation. In the first stage, the degree of concordance between the results of inter-subject and intra-subject correlation analyses is assessed statistically. The microstructurally (i.e., cytoarchitectonically) defined brain areas are designated either as concordant in which the results of both correlation analyses are in agreement, or as discordant in which one analysis method shows a higher proportion of supra-threshold voxels than does the other. In the second stage, connectivity networks are identified using the time courses of supra-threshold voxels in brain areas contingent upon the classifications derived in the first stage. In an empirical study, fMRI data were collected from 40 young adults (19 males, average age 22.76 ± 3.25), who underwent auditory stimulation involving sound clips of human voices and animal vocalizations under two operational conditions (i.e., eyes-closed and eyes-open). The operational conditions were designed to assess confounding effects due to auditory instructions or visual perception. The proposed two-stage analysis demonstrated that stress modulation (affective) and language networks in the limbic and cortical structures were respectively engaged during sound stimulation, and presented considerable variability among subjects. The network involved in regulating visuomotor control was sensitive to the eyes-open instruction, and presented only small variations among subjects. A high degree of concordance was observed between the two analyses in the primary auditory cortex which was highly sensitive to the pitch of sound clips. Our results have indicated that brain areas can be identified as concordant or discordant based on the two correlation analyses. This may further facilitate the search for connectivity networks involved in the processing of information under natural sensory stimulation.

1. Introduction

There is an increasing amount of literature on the analysis of functional magnetic resonance imaging (fMRI) data across experimental repetitions (Franco et al., 2009; Friedman et al., 2008; Somandepalli et al., 2015). Considerable strides have been made on temporal cross-correlations (TCCs) among the fMRI blood-oxygen-level-dependent (BOLD) time courses elicited by natural sensory stimulation. This appears to be a promising research direction for two reasons. First, this approach can be used to reveal the degree of similarity among subjects' responses when probing inter-subject synchronization. Second, it can be used to find reproducible brain activity within a subject when seeking to differentiate intrinsically driven from extrinsically driven systems in the brain (Hasson et al., 2010; Kauppi et al., 2010; Ren et al., 2017; Stephens et al., 2013). The term *extrinsic system* refers to all sensory motor areas involved in the processing of information derived from the outside environment (Golland et al., 2007). The extrinsic system presents high inter-subject and intra-subject TCCs in response to natural stimulation,

* Corresponding author. Institute of Statistical Science, Academia Sinica, Taipei 115, Taiwan.
E-mail address: mliou@stat.sinica.edu.tw (M. Liou).

https://doi.org/10.1016/j.neuroimage.2019.116042
Received 18 June 2019; Received in revised form 17 July 2019; Accepted 20 July 2019
Available online 22 July 2019

including primarily visual areas as well as somatomotor and dorsal attention networks. The *intrinsic system* normally shows task-related deactivations, and its activity is anti-correlated with that of the extrinsic system. The intrinsic system overlaps with the default mode network (DMN), which is recruited by episodic memory and social cognition, and the fronto-parietal control network, which is recruited by executive control tasks (Berkovich-Ohana and Glicksohn, 2014; Hacker et al., 2017; Salomon et al., 2009, 2014). The cortical subdivision into the two global systems has excluded some structures, including the limbic system, which are crucial for affective processes under natural sensory stimulation (Berridge and Kringelbach, 2013).

The influence of subjective perceptions, motivations, and real-life experiences on sensory stimulation can introduce considerable inter-subject variability (Bermudez et al., 2017; Newberg and d'Aquili, 1998). For example, it is known that language-related cortical responses differ as a function of linguistic experiences (Wang et al., 2003; Yang et al., 2015) even though the language network has been considered a part of the extrinsic system (Salomon et al., 2014). Alternatively, the classification of brain areas into networks can be based on the degree of concordance/discordance between the results of inter-subject and intra-subject TCC analyses. This notion permits a new classification criterion, wherein brain areas identified as concordant present spontaneous or induced responses that maintain strong agreement between the two correlation analyses under sensory stimulation. Moreover, responses in these areas are highly similar among subjects and independent of the random nature of the stimulus sequences assigned to individual subjects. In contrast, brain areas identified as discordant present spontaneous or induced responses exhibiting higher TCCs in one analysis but lower in the other. For example, responses in the visual cortex may present high correlation values in the inter-subject analysis under auditory stimulation, while also exhibiting low correlation values in the intra-subject analysis, if the experimental repetitions involve different operational conditions such as eyes-closed and eyes-open. Operational conditions are often selected to facilitate the evaluation of confounding effects associated with auditory instructions or visual perception (Wei et al., 2018; Zou et al., 2015). The language network, on the other hand, may be highly sensitive to the random nature of the stimulus sequences, and the corresponding responses present low correlation values in the inter-subject analysis. Nonetheless, these response patterns are reproducible across repetitions within individual subjects and insensitive to operational conditions. Note that the concordant/discordant classification has considered the microstructurally (i.e., cytoarchitectonically) defined brain areas as units of analysis in this study (Eickhoff et al., 2018). Our empirical results have revealed that the classification of brain areas based on this newly defined criterion may greatly facilitate the search for connectivity networks involved under natural sensory stimulation.

The TCC with a zero-time shift is commonly quantified using Pearson's correlation under the test-retest situation in the intra-subject analysis. In studies on natural stimulation, the group-level analysis is normally performed in conjunction with the intra-subject analysis, which means that correlations at the group level are estimated by averaging values across individual subjects (Golland et al., 2007; Harmelech et al., 2015), or by averaging BOLD time courses across subjects before computing correlations between responses to two presentations of the same stimulation by movie or video segments (Hasson et al., 2008). The pooling of correlations from the intra-subject analysis may not reveal the full range of variability in BOLD time courses among subjects. Pearson's correlation is a valid index for measuring TCCs when there are meaningful pairings in the test-retest scenario. The intraclass correlation (ICC) is a natural choice for detecting similarity among time courses in the case with more than two repetitions; it also allows the direct measurement of inter-subject cross-correlations (Friedman et al., 2008) without the need to concatenate time courses or to average all possible pairwise correlations between subjects (Hasson et al., 2004, 2010; Kauppi et al., 2010; Pajula et al., 2012; Ren et al., 2017). The ICC index reflects the degree of similarity among BOLD time courses or connectivity networks (Braun et al., 2012; Friedman et al., 2008; Kristo et al., 2014), and has been used to replicated statistical maps (t-maps or F-maps) in studies on memory encoding (Brandt et al., 2013), electro-acupuncture stimulation (Kong et al., 2007), and motor control (Kristo et al., 2014). It has also been applied to graph metrics in studies on reproducible brain networks under resting-states (Termenon et al., 2016) and natural viewing (Wang et al., 2017) conditions. Under natural stimulation, the ICC index is similar to Pearson's correlation in that they can both be applied directly to pre-processed BOLD time courses when assessing inter- and intra-subject TCCs.

Naturally occurring states of the brain are more likely to be induced by natural sensory stimulation than by artificially designed stimuli (Hanke et al., 2014). Our primary objective in this study was to develop a methodology based on ICC statistics to examine the connectivity networks in different brain structures, including the limbic system, and to interpret with respect to their functional roles in the processing of information under natural sensory stimulation (Jansma et al., 2000; Liou et al., 2012). The ICC index is sensitive to the number of repetitions, sample sizes, and outliers. Thus, the index must be standardized by its sampling error before statistical thresholding for scientific interpretation. As an example, Fig. 1 depicts intra- and inter-subject ICC maps with and without standardization based on sampling errors. The supra-threshold maps in the figure have indicated that empirical ICC indices can be excessively noisy, particularly in the intra-subject analysis. In summary, the proposed methodology is a two-stage approach. The first stage involves the statistical evaluation of concordance between the results of inter- and intra-subject ICC analyses and the classification of microstructurally defined brain areas as either concordant or discordant (Eickhoff et al., 2018). In the second stage, connectivity networks are identified based on BOLD time courses of supra-threshold voxels in the nodal areas contingent upon classifications derived in the first stage. The application of the methodology was demonstrated by conducting an fMRI experiment involving auditory stimulation using sound clips of different human voices and animal vocalizations under eyes-closed as well as eyes-open conditions. After constructing ICC maps based on supra-threshold voxels, the degree of concordance between the two maps was used to guide the classification of brain areas. Analysis on the sensitivity or insensitivity of connectivity networks to external stimulation was based on time courses of supra-threshold voxels in individual brain areas. The results were then interpreted with respect to the functional roles of these networks in the processing of information under natural sound stimulation.

2. Material and methods

2.1. Participants

A total of 40 right-handed young adults with an average age of 22.76 ± 3.25 years (21 females, average age: 22.24 ± 2.64 years; 19 males, average age: 23.30 ± 3.79) were recruited for the experiment. Subjects were mainly undergraduate or postgraduate students with normal hearing and without any history of neurological or mental disorders. Most of the subjects (26 out of 40) had previous experiences performing tasks in MRI scanners and 2 subjects had previously undergone MRI physical examinations. All subjects provided informed written consent before enrollment in the study. The experiment was approved by the Human Subject Research Ethics Committee/Institutional Review Board of Academia Sinica, Taiwan, in accordance with the Declaration of Helsinki.

2.2. Stimuli and tasks

The experimental task was conducted under two conditions: 4-min eyes-closed followed by 4-min eyes-open, each involving 12 types of sounds, as given in Kuo et al., 2019. During the experiment, the natural sounds were presented through high-quality MR-compatible insert

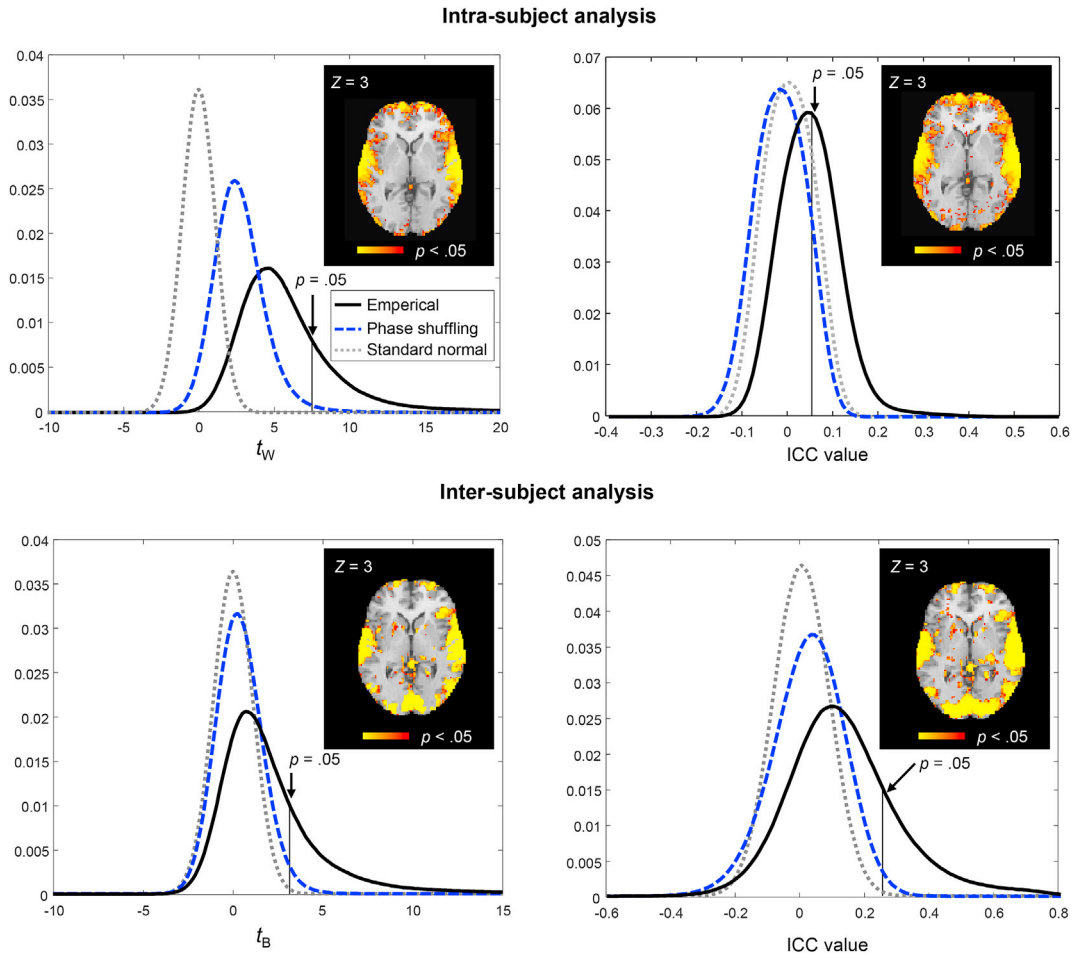

Fig. 1. Empirical distributions of t_W and t_B values estimated using the proposed method (left panel) along with the original ICC values (right panel) in intra-subject and inter-subject analyses. The ICC maps in the axial section at Z = 3 in the left panel were thresholded by comparing the empirical t_W and t_B values against phase randomized distributions with the FDR control at $\alpha = 0.05$. For the sake of comparison, the estimated ICC values were thresholded against phase randomized distributions with the FDR control at $\alpha = 0.05$ (intra-subject ICC values were averaged across all subjects).

earphones (Sensimetrics Model S14), which were worn under hearing protection ear muffs. Under the eyes-closed condition, subjects were instructed to rest quietly with eyes closed and passively listen to the natural sounds. Under the eyes-open condition, subjects were instructed to rest quietly and maintain eye-fixation on a crosshair in the central screen while listening to the sounds. The sound stimuli were edited clips from the databases at www.freesound.org and www.pacdv.com. All sounds were tuned to a similar intensity (70.81 ± 2.591 dB) and presented once at a time in a random sequence using Psychtoolbox-3 (www.psychtoolbox.org). The length of each sound clip was randomized within an interval range [16, 26]-sec to prevent synchronization between repetition time (TR) and the onset of individual sound clips (Price et al., 1999).

The stimuli consisted of two types of natural sounds: human voices (baby prattling, woman crying, man guffawing, woman laughing, man sneezing, and crowd chatter), and animal vocalizations (bird chirping, cow mooing, dog barking, sheep bleating, rooster crowing, and actual sounds from a farm). Each subject was assigned a unique sequence of stimuli; that is, the same random sequence of sound clips was repeated under eyes-closed and eyes-open conditions. The stimuli included human voices followed by animal vocalizations. The acoustic features of the various sound clips were outlined in Kuo et al., 2019. The acoustic features of interest included harmonics-to-noise ratio (HNR), intensity, pitch (as computed using Praat; see www.praat.org/), zero crossing rate, energy, average frequency of events, tempo, energy above cut-off frequency, energy below cut-off frequency, centroid spread, skewness, kurtosis, flatness of spectral distribution, spectrum entropy, and silence ratio (as computed using MIRtoolbox; Lartillot and Toiviainen (2007)). HNR measured the ratio between the strength of periodic and noisy components and was indexed by $HNR = 10\log_{10}[s/(1\text{-}s)]$, where s denotes the strength of the strongest periodic component. Intensity was calculated by the mean amplitude of the signals, whereas pitch was calculated using an autocorrelation method. There was no significant difference in these acoustic features between the human voices and animal vocalizations, based on the two-sample *t*-test. The results were as follows: HNR: $p = 0.66$; intensity: $p = 0.86$; pitch: $p = 0.12$; zeros crossing rate: $p = 0.56$; energy: $p = 0.79$; average frequency of events: $p = 0.33$; tempo: $p = 0.18$; energy above cut-off frequency: $p = 0.72$; energy below cut-off frequency: $p = 0.13$; centroid spread: $p = 0.56$; skewness: $p = 0.95$; kurtosis: $p = 0.69$; flatness of spectral distribution: $p = 0.63$; spectrum entropy: $p = 0.85$; and silence ratio: $p = 0.68$.

2.3. fMRI data acquisition and preprocessing

The MRI scans were performed using a 3T MAGNETOM Skyra scanner (Siemens Healthcare, Erlangen, Germany) and a standard 20-channel head-neck coil. The echo planar imaging (EPI) scans were performed with parameters TR/TE = 2000 ms/30 ms, flip angle = 84°; 35 slices, slice thickness = 3.4 mm, FOV = 192 mm, and resolution 3 × 3 × 3.74 mm to cover the whole brain including the cerebellum. In total, 120 image volumes were collected under the eyes-closed condition and 120 image volumes under the eyes-open condition within each subject. T1-weighted

anatomical images were acquired with parameters TR/TE = 2530 ms/3.30 ms, flip angle = 7°, 192 slices, FOV = 256 mm, and resolution 1 × 1 × 1 mm. The data were preprocessed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm). For each subject, low-resolution EPI images were realigned using rigid body transformation (translation and rotation corrections did not exceed 2 mm and 2°, respectively). The T1-weighted high-resolution image volume was co-registered to the mean of the realigned EPI images and spatially normalized using a voxel size of 2 × 2 × 3 mm to the standard MNI space. The signal-to-noise ratio of the functional images was enhanced via spatial smoothing using a 4-mm FWHM (full width at half maximum). In fMRI experiments, imaging techniques commonly cause systematic non-random artifacts, such as pulse sequences, imaging parameters, and scanner performance. In addition to regular pre-processing procedures (i.e., slice timing and motion correction), the BOLD time courses were corrected for major trend effects in order to account for drifts in the magnetic field using the GLM (Fox et al., 2009; Worsley et al., 2002). Following removal of these trends, the ICC statistics were applied to the pre-processed data.

2.4. Data analysis

2.4.1. ICC statistics

With M repetitions, each voxel-wise ICC estimate and its standard deviation can be computed. In this study, M refers to the number of conditions or subjects, and $\mathbf{S}$ is denoted as the M-by-M cross-covariance matrix with a zero-time shift among preprocessed BOLD time courses within a single voxel. The ICC can be estimated using the following equation:

$$\widehat{\mathrm{ICC}} = \frac{M}{(M-1)}\left[1 - \mathrm{tr}(\mathbf{S})/(\underline{\mathbf{1}}'\mathbf{S}\underline{\mathbf{1}})\right], \tag{1}$$

where $\mathrm{tr}(\mathbf{S})$ denotes the trace of the matrix $\mathbf{S}$, and $\underline{\mathbf{1}}$ is the summing vector of order M with the transpose $\underline{\mathbf{1}}'$. This equation is equivalent to $\widehat{\mathrm{ICC}}$(3, M) (Shrout and Fleiss, 1979) and $\widehat{\mathrm{ICC}}$(C, M) (McGraw and Wong, 1996). As the number of scan volumes (denoted by n) is sufficiently large within each repetition, the sampling distribution of the elements in $\mathbf{S}$ (with bounded sample variances of order $1/n$) will be approximately M(M+1)/2-dimensional Gaussian with zero mean and the covariance matrix $2K_M(\mathbf{\Sigma}\otimes\mathbf{\Sigma})K_M'$:

$$\sqrt{n}(\mathrm{vech}\mathbf{S} - \mathrm{vech}\mathbf{\Sigma}) \rightarrow N\left(0,\ 2K_M(\mathbf{\Sigma}\otimes\mathbf{\Sigma})K_M'\right), \tag{2}$$

where $\otimes$ is the Kronecker product and $\mathbf{\Sigma}$ is the population counterpart of $\mathbf{S}$; $\mathrm{vech}\mathbf{\Sigma}$ is the half-vectorization of the matrix $\mathbf{\Sigma}$; K_M is the transformation matrix of order M(M+1)/2-by-M^2 with the identity $\mathrm{vech}\mathbf{\Sigma} = K_M(\mathrm{vec}\mathbf{\Sigma})$, where $\mathrm{vec}\mathbf{\Sigma}$ is the vectorization of the matrix $\mathbf{\Sigma}$ (Henderson and Searle, 1979; Iwashita and Siotani, 1994; van Zyl et al., 2000).

Because the ICC index is a scalar-valued function of $\mathbf{S}$ in (1) and differentiable in a neighborhood of $\mathbf{S} = \mathbf{\Sigma}$, the asymptotic variance of $\widehat{\mathrm{ICC}}$ can be estimated by

$$\mathrm{Var}(\widehat{\mathrm{ICC}}) = 2n^{-1}\eta' K_M(\mathbf{\Sigma}\otimes\mathbf{\Sigma})K_M'\eta\big|_{\mathbf{\Sigma}=\mathbf{S}} \tag{3}$$

In the equation, η' is the derivative of ICC with respect to $\mathrm{vech}'\mathbf{\Sigma} = (\mathrm{vech}\mathbf{\Sigma})'$ with $\mathbf{\Sigma}$ evaluated at $\mathbf{S}$ in application; it has the following form:

$$\eta' = \frac{M}{(M-1)(\underline{\mathbf{1}}'\mathbf{\Sigma}\underline{\mathbf{1}})^2}\left[-(\underline{\mathbf{1}}'\mathbf{\Sigma}\underline{\mathbf{1}})\mathrm{vec}'(I_M) + \mathrm{tr}(\mathbf{\Sigma})(\underline{\mathbf{1}}'\otimes\underline{\mathbf{1}}')\right]G_M\big|_{\mathbf{\Sigma}=\mathbf{S}}, \tag{4}$$

where I_M is the identity matrix of order M, and G_M is the transformation matrix of order M^2-by-M(M+1)/2 with the identity $\mathrm{vec}\mathbf{\Sigma} = G_M(\mathrm{vech}\mathbf{\Sigma})$. A brief derivation of η' in (4) is given in Appendix-A, where $\sqrt{n}$-rate asymptotic normality of (2) and $(1/n)$-rate asymptotic variance of (3) are valid because the lagged correlations between the entries in $\mathbf{S}$ satisfy the strong mixing condition of stationary time series, as evidenced by the uniformly bounded variance estimates (of order $1/n$) of $\widehat{\mathrm{ICC}}$ values given in Supplementary Figure S8. The standard error estimator in (3) was originally proposed by Browne (1974) and van Zyl et al. (2000) for values at time points assumed to be independently and identically distributed. The use of the estimator for evaluating temporal dependent BOLD time courses requires the stationarity assumption. Details pertaining to this assumption are presented in Appendix-A, including a statistical evaluation of the assumption using the Kwiatkowski-Phillips-Schmidt-Shin and Augmented Dickey–Fuller tests (Hamilton, 1995; Kwiatkowski et al., 1992). The t-value for testing the significance of inter-subject estimates is defined as follows:

$$t_B = \frac{\widehat{\mathrm{ICC}}}{\sqrt{\mathrm{Var}(\widehat{\mathrm{ICC}})}}. \tag{5}$$

Data acquisition provided $n = 240$ and M = 40 in the inter-subject analysis.

In the intra-subject analysis, the BOLD time courses were divided into two repetitions (i.e., M = 2), under either eyes-closed or eyes-open conditions. The ICC assessment was applied to each subject, and $\widehat{\mathrm{ICC}}_j$ as well as $\mathrm{Var}(\widehat{\mathrm{ICC}}_j)$ were estimated for the j-th subject. The individual estimates were synthesized as follows:

$$t_W = \frac{\sum_{j=1}^{\mathrm{N}}\widehat{\omega}_j\widehat{\mathrm{ICC}}_j}{\sqrt{\left(\sum_{j=1}^{\mathrm{N}}\frac{1}{\mathrm{Var}(\widehat{\mathrm{ICC}}_j)}\right)^{-1}}}, \tag{6}$$

where $\widehat{\omega}_j = \frac{1}{\mathrm{Var}(\widehat{\mathrm{ICC}}_j)}\left(\sum_{j=1}^{\mathrm{N}}\frac{1}{\mathrm{Var}(\widehat{\mathrm{ICC}}_j)}\right)^{-1}$ and N = 40 in this study. The denominator of t_W is the standard deviation of $\sum_{j=1}^{\mathrm{N}}\widehat{\omega}_j\widehat{\mathrm{ICC}}_j$, which is the weighted average of $\widehat{\mathrm{ICC}}_j$ for $j = 1, \ldots, 40$. The derivation of the variance estimate for $\sum_{j=1}^{\mathrm{N}}\widehat{\omega}_j\widehat{\mathrm{ICC}}_j$ can be found in Appendix-B.

2.4.2. ICC maps

The BOLD time courses were temporally and spatially correlated; therefore, surrogate data distributions for evaluating the significance of t_B and t_W were generated using the phase-randomization procedure (Lerner et al., 2011). This procedure was validated based on the stationarity tests presented in Appendix-A. To control for familywise Type-I error, the false discovery rate (FDR) control procedure was used in the thresholding of p-values (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001; Friederici, 2012; Genovese et al., 2002; Langers et al., 2007). In the FDR procedure, a sequence of ordered p-values (i.e., the probability of t statistics in the surrogate distributions equaling or exceeding the observed t_B or t_W values) were compared with the critical value $(i/\mathrm{V})[\alpha/\mathrm{C(V)}]$ for the i-th p-value in the ordered sequence, where C(V) is a predetermined constant and α is the nominal Type-I error rate ($\alpha = 0.05$). The choice of this constant depends on the joint distribution of p-values in the sequence. We selected $C(V) = \sum_{i=1}^{\mathrm{V}} i^{-1}$ as a weak FDR control (Genovese et al., 2002) in order to take into account the possibility of false negatives. The ICC maps were constructed by identifying the supra-threshold voxels using t_B or t_W statistics that exceeded the FDR critical values. The ICC maps were co-registered to the JuBrain Cytoarchitectonic Atlas provided by the SPM Anatomy toolbox (Eickhoff et al., 2005) for the topographical definitions of area-specific effects. An interested reader may refer to Amunts and Zilles (2015) for the definition and localization of brain areas in the JuBrain Atlas.

2.4.3. Concordance analysis

In a given area of the brain, a single voxel had a pair of supra-threshold or infra-threshold classification labels derived using intra-

Table 1
The 2-by-2 table for the concordance analysis.

	Inter-Subject (t_B)		
Intra-Subject (t_W)	Supra-threshold	Infra-threshold	Sum
Supra-threshold	a (P_a)	b (P_b)	a+b (P_W)
Infra-threshold	c (P_c)	d (P_d)	c+d ($1 - P_W$)
Sum	a+c (P_B)	b+d ($1 - P_B$)	a+b+c+d (1)

and inter-subject analyses, respectively (i.e., the supra-threshold label was assigned to a voxel if its t_B or t_W value was above the FDR threshold, and the infra-threshold label was assigned if otherwise). In the 2-by-2 table below, the numbers of voxels within the four classification categories are denoted as a, b, c, and d:

In Table 1, P_i (with the population value p_i) denotes the observed joint probability in the i-th classification category for $i =$ a, b, c, and d, and P_W and P_B respectively denote the marginal probabilities of supra-threshold voxels in intra- and inter-subject analyses. McNemar's test was used to examine the number of voxels with discordant statuses, b and c. If b and c were of equal size, then the numbers of supra-threshold voxels in the two analyses would be concordant; that is, the two analyses provided the same analysis results in terms of the total number of supra-threshold voxels in a given brain area. McNemar's test has been widely used in clinical studies to evaluate the effects of medication on a given group of subjects before and after treatment.

The exact probability that one can test the null hypothesis, $p_W = p_B$ with k = b+c, can be computed as follows:

$$\text{Two-sided } p\text{-value} = 2 \sum_{x=0}^{\min(b,c)} \binom{k}{x} \left(\frac{1}{2}\right)^k \tag{7}$$

(Breslow et al., 1987). Fisher's mid-p test is a modified version of the exact McNemar's test by subtracting two-sided p-value with a factor $\binom{k}{c}\left(\frac{1}{2}\right)^k$, which provides more accurate test results under the null hypothesis (Fagerland et al., 2013). If the mid-p-value is less than the critical value (e.g., $\alpha = 0.05$), then the null hypothesis cannot be retained. Otherwise, it may be claimed that the two marginal probabilities are concordant. Based on Fisher's mid-p test, each brain area was classified within one of the three categories: $p_W > p_B$, $p_W < p_B$, and $p_W = p_B$ (i.e., null hypothesis).

In the $p_W > p_B$ case, the BOLD time courses were reproducible in the two repetitions; nonetheless, there was considerable inter-subject variability, which could be attributed to the random nature of the stimulus sequences assigned to individual subjects, to subjective (internal) appraisals of stimuli, or to anatomical variations among subjects (Amunts et al., 2000). In the $p_W < p_B$ case, the time courses differed between the two repetitions (i.e., brain reactions to the eyes-open instruction were stronger than those to the eyes-closed instruction); however, the pattern of the reactions was consistent across subjects. In a comparison of time courses between the two repetitions, cases where $p_W = p_B$ tended to be insensitive to operational conditions. In the concordance analysis, the FDR threshold at 0.05 was applied to the empirical mid-p-values.

2.4.4. Connectivity network

The importance of conducting intra- and inter-subject analyses concurrently lies in the scientific interpretation of connectivity networks when subjects are under exposure to sensory stimulation. In the second stage, connectivity networks were considered in the three cases ($p_W = p_B$, $p_W > p_B$ and $p_W < p_B$) separately. The identification of regions of interest (ROIs) in a network was based on supra-threshold voxels in the three cases. The time courses with significant ICCs were first averaged across supra-threshold voxels to generate the mean time course in a given ROI within each subject. For example, if an ROI was classified to as a $p_W > p_B$ case, then the supra-threshold voxels in the intra-subject analysis were identified for the computation of mean time courses. A brain area was eliminated from analysis if it contained less than 5% supra-threshold voxels or if the numbers of supra-threshold voxels were less than 5 in both analyses (Konrad et al., 2005). Pairwise correlations were computed using the mean time courses between each pair of ROIs in a network (Betti et al., 2013; O'Neil et al., 2014; Wang et al., 2007). Fisher's Z transformation was applied to improve the normality of correlations. There were 40 Fisher Z-transformed correlations between each pair of ROIs in a network, the sample mean of which was evaluated for statistical significance using the Student's t-test (Betti et al., 2013; O'Neil et al., 2014; Wang et al., 2007). To facilitate visualization, the connectivity networks were reconstructed using the BrainNet Viewer (Xia et al., 2013). Table 2 gives a list of cytoarchitectonic areas that are referred to in the connectivity study along with their corresponding Brodmann nomenclature.

3. Results

3.1. Randomized distributions and ICC maps

Fig. 1 depicts the empirical distributions of t_W and t_B values computed using empirical BOLD time courses in in-brain voxels along with the randomized distributions. For the sake of comparison, the empirical distributions of raw ICC values averaged across subjects in the intra-subject analysis and unstandardized ICC values in the inter-subject analysis are also presented in the right panel of Fig. 1 along with their randomized distributions. An FDR control was used to constrain the experiment-wise false positive rate at 0.05 for thresholding all of the ICC and t values in the figure. The number of subjects was moderate to large ($\geq$40); therefore, the inter-subject analysis was less affected by the standardization of ICC values, such that standardized and unstandardized ICC values provided similar supra-threshold ICC maps. However, the intra-subject analysis was greatly affected by the weighting scheme; that is, the unweighted ICC averages were noisier than the weighted averages. Supplementary Figures S1 and S2 depict the ICC maps after FDR thresholding in the two analyses. The experiment involved natural auditory sounds; therefore, BOLD time courses in the superior temporal cortex, prefrontal cortex, and a portion of the parietal cortex presented a large number of supra-threshold voxels in the intra-subject ICC maps. The occipital cortex had more supra-threshold voxels in the inter-subject analysis than in the intra-subject analysis. These results suggested that the occipital cortex was more sensitive to whether the sequences were performed under eyes-closed or eyes-open conditions.

3.2. Concordance between the two analyses

Fig. 2 depicts the bivariate relationships between the proportions of supra-threshold voxels in the intra-subject (horizontal axis) versus inter-subject (vertical axis) analyses for the brain areas listed in Table 2. The spatial locations of these areas in the brain can be visualized using the BrainNet Viewer as shown in Supplementary Figure S3. The blue balls in the figure denote brain areas in the $p_W = p_B$ case. Most of the blue balls fall on the diagonal line in the figure. For example, the primary auditory cortex (Areas Te 1.0, 1.1, and 1.2 after Morosan et al. (2001); Brodmann area (BA) 41 after Brodmann (1909)) was insensitive either to the randomness in stimulus sequences assigned to subjects in the inter-subject analysis or eyes-closed and -open conditions in the intra-subject analysis. Most brain areas were classified as discordant cases such that proportions of supra-threshold voxels were significantly higher in either the intra-subject analysis than in the inter-subject analysis (denoted by red balls) or vice versa (denoted by green balls). For example, Area 45 of Broca's region (Amunts et al., 1999) and Area Fp2 (Bludau et al. (2014); medial part of BA 10) in the frontal pole were sensitive to the randomness of the stimulus sequences assigned to individual subjects (or subjective appraisals of the stimulus contents), and

Table 2
Cytoarchitectonic brain areas referred to in the connectivity study.

Area names	Brain regions	Relationship with Brodmann areas	Atlases	References
hOc1	Primary visual cortex	17	JuBrain	Amunts et al. (2000)
hOc2	Secondary visual cortex	18	JuBrain	Amunts et al. (2000)
hOc3d, hOc3v, hOc4la, hOc4lp, hOc4v, hOc5/V5, hOc6/V6	Higher visual cortex	Parts of 19	JuBrain	Malikovic et al. (2016); Kujovic et al. (2013); Malikovic et al. (2007)
Te1.0, 1.1, 1.2	Primary auditory cortex	41	JuBrain	Morosan et al. (2001)
Te3	Superior temporal gyrus	Central part of 22	JuBrain	Morosan et al. (2005)
MTG	Middle temporal gyrus	21	Neuromorphometrics, Inc.	http://neuromorphometrics.com/
44, 45	Broca's region	44, 45	JuBrain	Amunts et al. (1999)
47	Part of the inferior frontal gyrus	47	MRIcro	http://www.mricro.com/
6d1, 6d2, 6d3, 6mc, 6mr	Premotor cortex	6	JuBrain	Geyer (2004)
4a, 4p	Primary motor cortex	4	JuBrain	Geyer et al. (1996)
Fp1, Fp2	Frontal pole	10	JuBrain	Bludau et al. (2014)
Fo1, Fo2, Fo3	Parts of orbito-frontal cortex	11	JuBrain	Henssen et al. (2016)
1, 2, 3a, 3b	Somatosensory cortex	1, 2, 3	JuBrain	Geyer et al. (2000); Grefkes et al. (2001)
7A, 7M, 7P, 7PC, 5L, 5M	Superior parietal cortex	5, 7	JuBrain	Scheperjans et al. (2008)
PFt	Part of the inferior parietal cortex	Part of 40	JuBrain	Caspers et al. (2006)
OP1, OP4	Parts of the secondary somatosensory cortex	Opercular part of 43	JuBrain	Eickhoff et al. (2006b)
Ch4	Nucleus basalis of Meynert	n.d.	JuBrain	Zaborszky et al. (2008)
FEF	Frontal eye field	Part of 8 at the junction of the superior frontal and precentral sulci	MRIcro	http://www.mricro.com/
PCC	Posterior cingulate cortex	23, 31	Neuromorphometrics, Inc.	http://neuromorphometrics.com/
s32, s24, 25	Subgenual anterior cingulate cortex	Subgenual parts of 24 and 32, 25	JuBrain	Palomero-Gallagher et al. (2009)
CM, SF, MF, VTM	Centro-medial, superficial, medial fiber bundles, and ventro-medial nuclei of the amygdala	n.d.	JuBrain	Amunts et al. (2005)
EC	Entorhinal cortex	28	JuBrain	Amunts et al. (2005)
NA	Nucleus accumbens	n.d.	Neuromorphometrics, Inc.	http://neuromorphometrics.com/

Note: n.d. Denotes not determined.

were relatively insensitive to eyes-closed and -open conditions. In contrast, Area 7M (Scheperjans et al. (2008); medial part of BA 7) in the superior parietal cortex and the primary visual area hOc1 (Amunts et al. (2000); BA 17) in the visual cortex were sensitive to eyes-closed and -open conditions, and relatively insensitive to the random sequences.

The classification of cytoarchitectonically defined brain areas according to the degree of concordance/discordance can enable a more precise assessment of the functional dynamics engaged during natural sensory stimulation. For illustration purposes, Fig. 3(A)–(E) depict the supra-threshold ICC maps in Areas s24 ($p_W > p_B$) in the subgenual anterior cingulate (Palomero-Gallagher et al., 2015), 45 ($p_W > p_B$) in the Broca's region (Amunts et al., 1999), Te1.0 ($p_W = p_B$) in the superior temporal cortex (Morosan et al., 2001), PFt ($p_W < p_B$) in the inferior parietal cortex (Caspers et al., 2006), and hOc1 ($p_W < p_B$) in the occipital cortex (Amunts et al., 2000). In Area s24, for example, 62.7% supra-threshold voxels (31.1% + 31.6%) were observed in the intra-subject analysis, and 45.1% voxels (31.1% + 14.0%) were observed in the inter-subject analysis. Although the BOLD time courses in supra-threshold voxels exhibited similar patterns in the two analyses (also see the mean time courses in Fig. 4(A)), the inter-subject analysis had under-estimated reproducible responses in this area. The ICC maps in Area 45 demonstrated that the supra-threshold voxels in the intra-subject analysis were distributed almost symmetrically in the two hemispheres, but these supra-threshold voxels common in both intra- and inter-subject analyses were more noticeable in the left hemisphere. If connectivity was computed based on supra-threshold voxels in the inter-subject analysis, the results would be biased in favor of the left hemisphere. Area Te1.0 presented considerable overlap between the supra-threshold voxels in the two analyses. The BOLD time courses indicated that this area was insensitive to the mode of experimental instructions (i.e., eyes-closed or eyes-open; also see Fig. 4(B)). Area PFt displayed decreased responses following the eyes-open instruction and those responses were mainly distributed in the right hemisphere (also see Fig. 5(A)). Supra-threshold voxels in the intra-subject analysis, however, were distributed in the left hemisphere. If the computation of connectivity was based on supra-threshold voxels in the intra-subject analysis, then the results would be biased against convergent responses (across subjects) in this area. In Area hOc1, 73.2% supra-threshold voxels were observed in the inter-subject analysis, and 6.2% voxels were observed in the intra-subject analysis. Empirical BOLD time courses revealed that response amplitudes in these areas were higher after the onset of the eyes-open instruction than after the onset of the eyes-closed instruction (also see Fig. 5(B)).

In summary, brain areas classified as the $p_W < p_B$ case were more sensitive to operational conditions (e.g., Areas PFt and hOc1). Those classified as the $p_W > p_B$ case were insensitive to operational conditions (e.g., Areas s24 and 45); however, they were sensitive to the randomness of stimulus sequences, subjective appraisals of the stimulus contents, and anatomical variations among subjects. Brain areas in the $p_W = p_B$ case showed similar response patterns under the two operational conditions (e.g., Area Te1.0), and were insensitive to random sequences; some of these areas were sensitive to the pitch of auditory stimulation as detailed below.

3.3. Acoustic features

The mean time course in each brain area, contingent upon the classification in Fig. 2, was correlated with auditory features within each subject. The mean of the Z-transformed correlations was evaluated using the Student's *t*-test based on a sample size of 40. Among the acoustic features, the pitch was found to have significant correlations with the empirical BOLD time courses in Areas OP4 (Eickhoff et al. (2006b); parietal ventral cortex), BA 45 in the left hemisphere (part of Broca's

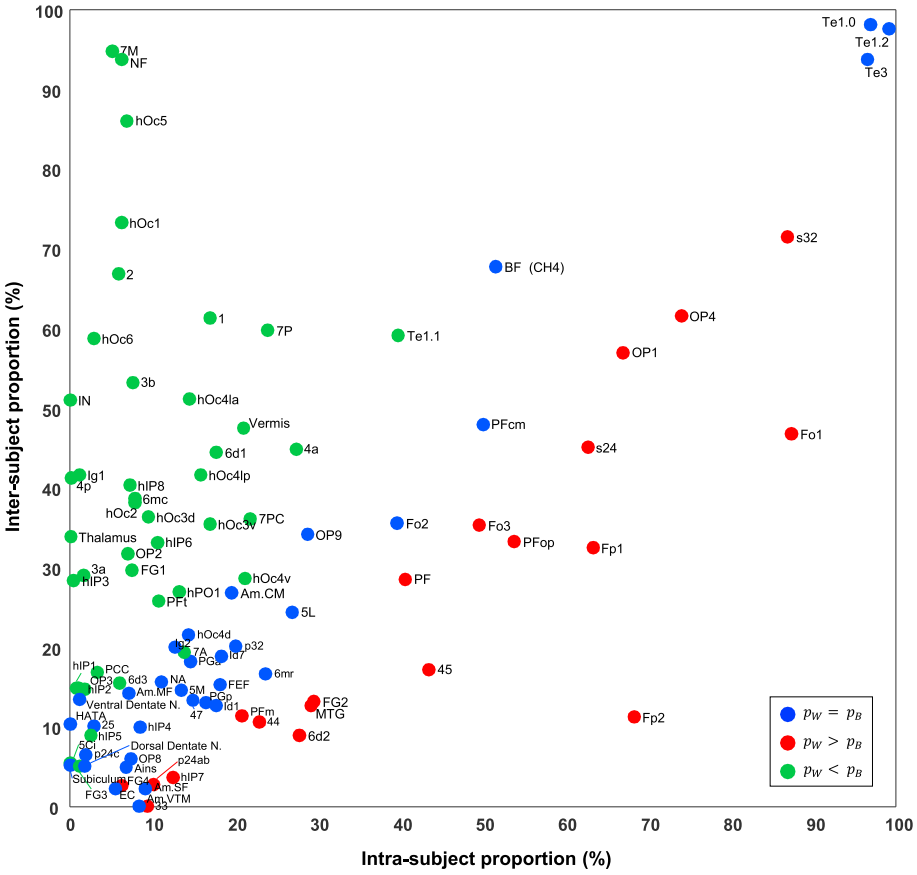

Fig. 2. Bivariate plots showing the proportions of supra-threshold voxels in brain areas in the intra-subject (horizontal axis) and inter-subject (vertical axis) analyses. Labeled brain areas can be visualized using the BrainNet Viewer, as shown in Supplementary Figure S3. The blue balls indicate the corresponding areas in concordant cases (i.e., the temporal cortex, insular cortex, limbic structures, and cerebellum). The red balls (i.e., the ventral medial prefrontal cortex, Broca's region, and parts of the parietal cortex) and green balls (i.e., the premotor, motor, parietal and occipital cortices) indicate areas in discordant cases. Brain areas were defined mainly using the JuBrain Atlas provided by the SPM Anatomy toolbox (Eickhoff et al., 2005, 2006a). Some brain areas were defined using the Neuromorphometrics Inc. Atlas (http://neuromorphometrics.com/) and MRIcro software (http://www.mricro.com/) in support of the connectivity network analysis. The areas and their corresponding brain regions can be found in Table 2. Abbreviations in the figure denote: NA ≡ Nucleus accumbens, Am ≡ Amygdala, EC ≡ Entorhinal cortex, NF ≡ Nucleus fastigii, FEF ≡ Frontal eye field, IN ≡ Interposed nuclei; MTG ≡ Middle temporal gyrus; PCC ≡ Posterior cingulate cortex; Ains ≡ Anterior insula; [L] ≡ Left; [R] ≡ Right. Brain areas eliminated from the analysis due to small proportions of supra-threshold voxels (<5%) included the following: hippocampus (CA1-3, and DG), and amygdala (IF and LB).

region), Te1.0, Te1.2, and Te3 (posterior BA 22). Supplementary Table S1 lists those features that had correlations greater than 0.2. Among these areas, the highest correlation value was found in the left Te1.0 ($r = -0.47$, $p < 0.01$). These pitch sensitive areas fall along the diagonal line in Fig. 2, which are consistent with recent findings on pitch-selective regions, including the superior temporal gyrus and Broca's region (Bendor and Wang, 2005; Hall and Plack, 2009; Leaver and Rauschecker, 2010; Muller et al., 2001; Tierney et al., 2013). The parietal operculum (OP) was reported to be selective to specific pitches (Humphries et al., 2010), and the inferior parietal cortex was found to be activated in pitch discrimination tasks (Royal et al., 2016). Such pitch-selective property was also revealed by the correlation between the kurtosis/flatness of spectral distributions and responses in Te3 ($p < 0.01$). In addition, we found that the energy had correlations with the activation in the left Te1.0 and OP4 ($p < 0.01$), which supported previous findings showing that these areas exhibited higher activation when the sound level decreased because of the foreground-background decomposition (Brechmann et al., 2002). Moreover, the correlation between tempo and time courses in Area Te1.2 in the right hemisphere suggested that this area was related to the perception of rhythm, of which a key feature is a regular tempo (or beat) (Chen et al., 2006; Royal et al., 2016).

3.4. Connectivity networks

As mentioned, the standardized ICC statistics were sensitive to brain dynamics in the limbic system and subcortical structures. For illustration purposes, two connectivity networks were chosen from brain areas in the $p_W > p_B$ case: the stress modulation network for regulating affective processes distributed primarily in the limbic structures (Berridge and Kringelbach, 2013; Cerqueira et al., 2008; Pannese et al., 2015; Pruessner et al., 2008), and the auditory language network distributed over various cortical structures (Friederici, 2012; Friederici and Alter, 2004; Parker et al., 2005). Fig. 4(A) depicts the stress modulation network, in which the nodal ROIs include the orbital frontal cortex (Henssen et al. (2016); Areas Fo1, Fo2, and Fo3), subgenual anterior cingulate (Palomero-Gallagher et al. (2009); Areas s32, s24 and BA 25), nucleus accumbens, nucleus basalis Meynert of the basal forebrain (Zaborszky et al. (2008); Ch4), amygdala (Amunts et al. (2005); centro-medial, superficial, region of medial fiber bundles, and ventro-medial nuclei), and entorhinal cortex.[1] See Cerqueira et al. (2005) and Berridge and Kringelbach (2013) for the definition and functions of the network. Note that several ROIs in this network (e.g., Areas Fo2 and 25) were parts of the $p_W = p_B$ case, which showed similar results in intra- and inter-subject analyses as illustrated in Fig. 3(C). The inclusion of nodal areas in the $p_W = p_B$ case would not affect the precision in evaluating the network. As an example, Supplementary Figure S5 depicts empirical BOLD time courses of supra-threshold voxels in Area Fo2 in the intra-subject analysis. This area is one of the nodal areas in Fig. 4(A) and showed concordant results between intra- and inter-subject analyses (i.e., time courses of supra-threshold voxels in this area were almost the same in the two analyses). According to Figure S5, activity in Area Fo2 was consistent with that in other nodal areas in the network. The diameter of a ball in the figure indicates the proportion of supra-threshold voxels in a given area. All lines connecting two areas were significantly correlated, as determined by the Student's *t*-test.

Fig. 4(B) depicts the auditory language network, including the extended Broca region (BA 44, BA 45, and BA 47), premotor cortex (6d2

[1] Note: The insular cortex has been considered one of the key regions in processing affective components in stimuli. However, the time courses of supra-threshold voxels in this cortex were similar to those in the language network, especially in the anterior insula. The time courses of the anterior insula in the intra-subject analysis are depicted in Supplementary Figure S4.

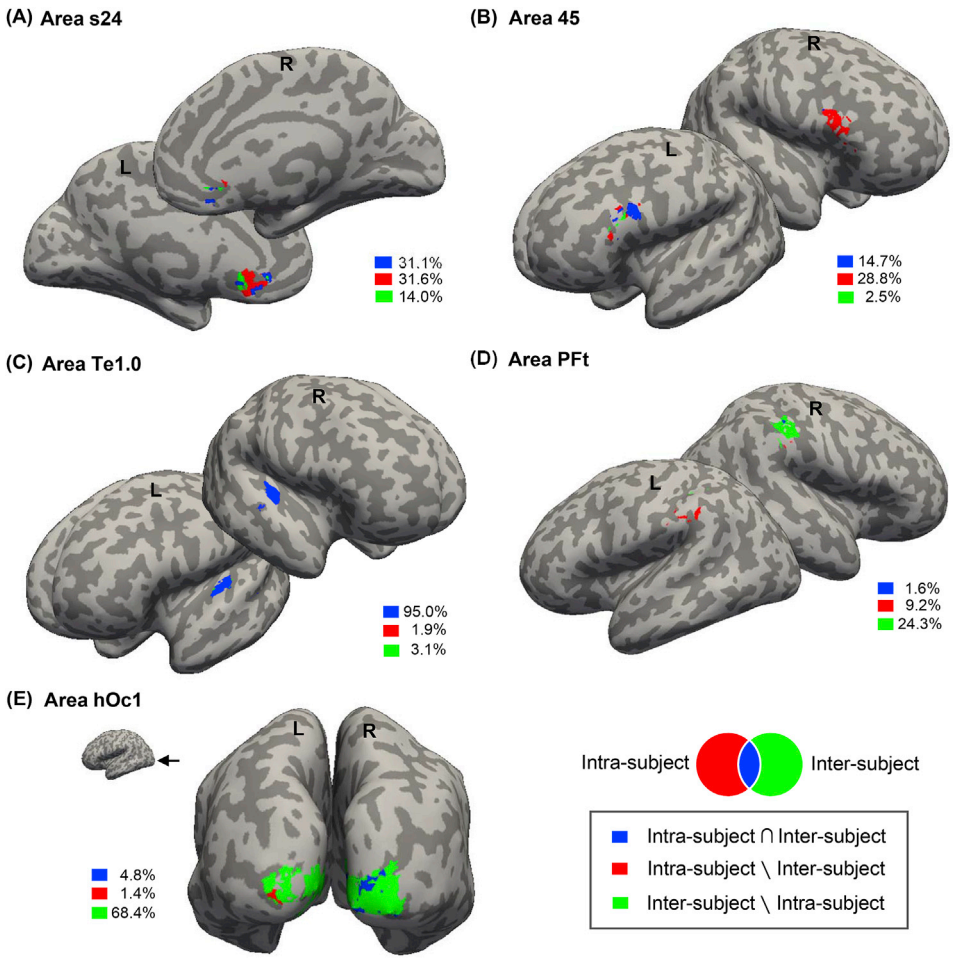

Fig. 3. Supra-threshold voxels in Areas (A) s24 (subgenual BA24), (B) 45 (Broca's region), (C) Te1.0 (primary auditory cortex), (D) PFt (part of BA40), and (E) hOc1 (BA17). The supra-threshold voxels are plotted in blue, red, or green as shown by FreeSurfer v. 5.0.0 (http://surfer.nmr.mgh.harvard.edu). Blue areas correspond to supra-threshold voxels common in the two analyses (Intra∩Inter); red areas correspond to supra-threshold voxels uniquely identified in the intra-subject analysis (Intra\Inter); green areas correspond to supra-threshold voxels uniquely identified in the inter-subject analysis (Inter\Intra).

and 6mr), primary auditory cortex (Te1.0–1.2), posterior part of BA22 (Te3; part of Wernicke's area in the left hemisphere), and middle temporal gyrus (MTG). See Friederici (2012) and Parker et al. (2005) for the definition of the auditory language network. The chart under each connectivity network depicts the mean time courses averaged across supra-threshold voxels and brain areas within a given network for individual subjects (i.e., there are 40 mean time courses in each chart), and the average across the 40 mean time courses is marked in bold. The 40 mean time courses showed considerable variability among subjects in both networks. On average, responses in the stress modulation network displayed prolonged decreases under the eyes-closed condition as well as the eyes-open condition. Overall, human voices generated significantly higher amplitudes than did animal vocalizations in the language network (two-sample t-test, $p < 0.05$). Nodal areas in the networks unavailable in JuBrain were parcellated using other atlases as specified in Table 2. These additional brain areas can also be found in Fig. 2 along with JuBrain areas. As shown in Supplementary Figure S6, several nodal areas had insignificant connectivity values with other areas in the two networks when connectivity was computed based on supra-threshold voxels in the inter-subject analysis.

Most of the brain areas in the occipital and parietal cortices were classified as $p_W < p_B$ cases, and could be delineated into two groups depending on an increase or decrease of activity following the eyes-open instruction. One group, connecting the primary motor cortex (Geyer et al. (1996); Areas 4a and 4p), somatosensory cortex (Areas 1, 2, 3a and 3b), higher visual cortex (Malikovic et al. (2007); hOc4la, and hOc5/V5), parts of the premotor cortex (Areas 6d1, and 6mc), part of the inferior parietal cortex (Area PFt), and the superior parietal cortex (Scheperjans et al. (2008); 5L, 7A, and 7 PC), exhibited a short-term decrease in activity immediately following the eyes-open instruction. Following the same instruction, an increase in activity occurred in the thalamus, cerebellar vermis, nucleus fastigii, interposed nuclei, posterior and superior parietal cortex (5M, 7A, 7P, and 7M), and occipital cortex (hOc1, hOc2, hOc3d, hOc3v, hOc4lp, hOc4v, and hOc6/V6). In the literature, the parietal and occipital cortices have been acknowledged in the visuomotor control, including saccadic eye movements (Culham et al., 2006). To interpret the functional roles of brain areas in the $p_W < p_B$ case, the frontal eye field (FEF) and posterior cingulate cortex (PCC) were also parcellated. The FEF has been associated with tasks involving voluntary control of saccadic eye movements (Kleiser et al., 2017), whereas the PCC has been referred to in reflexive saccades (Carrasco, 2011). The empirical BOLD time courses suggest that the FEF exhibited a decrease in activity following the eyes-open instruction, whereas the PCC showed an increase in activity. Fig. 5 depicts the connectivity of these two groups of brain areas (Note: Thalamic nuclei as well as lobules in the cerebellar vermis exhibited similar patterns of activity). As a comparison, Supplementary Figure S7 depicts the same connectivity networks based on supra-threshold voxels in the intra-subject analysis. Several nodal areas in Fig. 5 had insignificant connectivity values with other areas when connectivity was computed based on supra-threshold voxels in the intra-subject analysis.

In summary, the inter-subject analysis suggested that visual and parietal cortices exhibited less functional variability among subjects while processing natural sounds; however, these regions were activated differently following the eyes-open instruction. The intra-subject analysis, on the other hand, indicated that functional dynamics in structures including the limbic system, frontal, and temporal cortices showed considerable variability among subjects, but were insensitive to the two

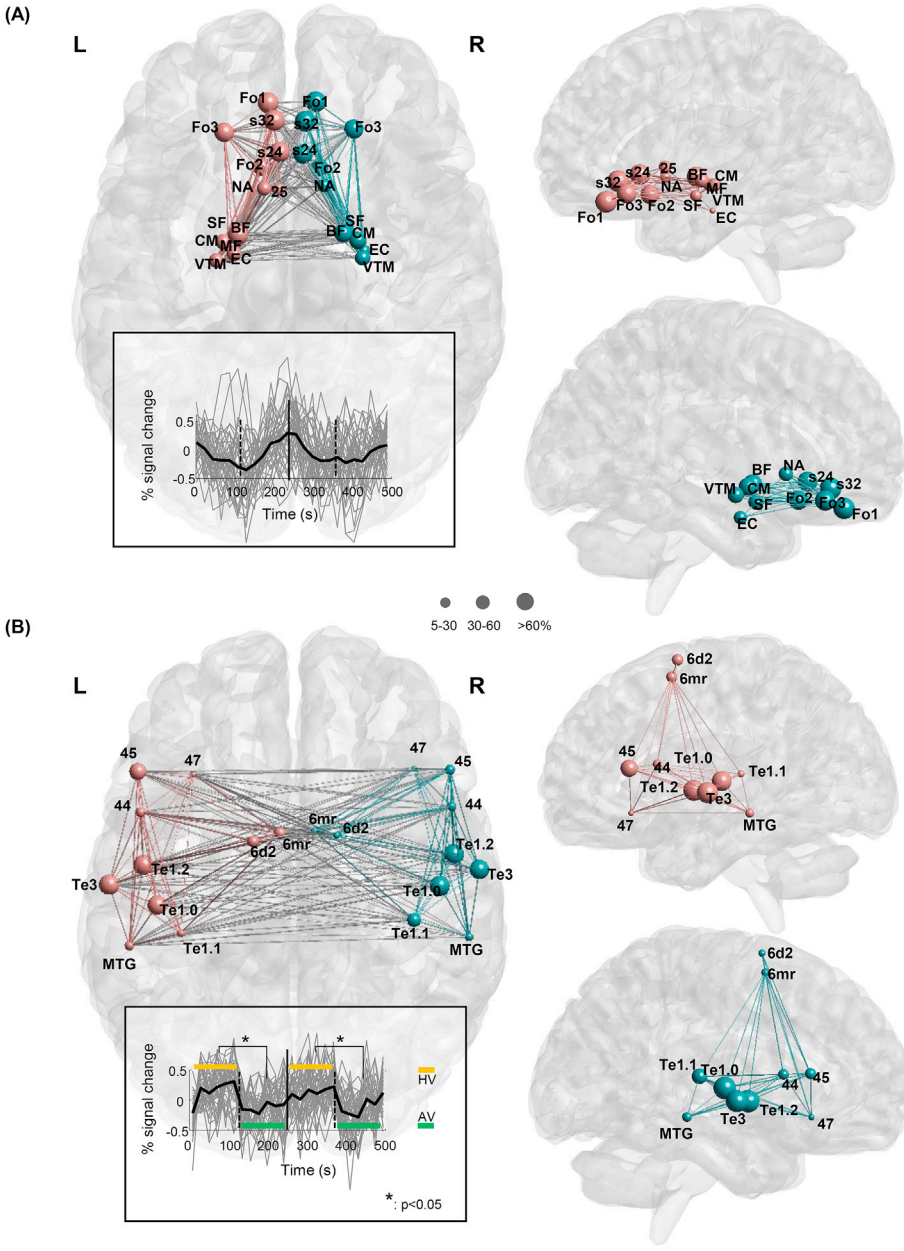

Fig. 4. Connectivity networks in the intra-subject analysis. The stress modulation network in (A) includes the fronto-orbital (Fo1, Fo2, and Fo3), subgenual anterior cingulate (s32, s24 and 25), nucleus accumbens (NA), basal forebrain (BF), amygdala (CM, MF, SF, and VTM), and entorhinal cortex (EC). The average correlations between ROIs in this network were 0.31 ± 0.14 (left hemisphere), 0.33 ± 0.15 (right hemisphere), and 0.33 ± 0.14 (inter-hemisphere) under the eyes-closed condition, and slightly lower under the eyes-open condition with 0.31 ± 0.12 (left hemisphere), 0.30 ± 0.12 (right hemisphere), and 0.31 ± 0.12 (inter-hemisphere). The language network in (B) includes the pars opercularis (BA 44), pars triangularis (BA 45), pars orbitalis (BA 47), premotor cortex (6d2 and 6mr), superior temporal cortex (Te1.0-1.2, Te3), and middle temporal gyrus (MTG). The average correlations in this network were 0.42 ± 0.13 (left hemisphere), 0.40 ± 0.12 (right hemisphere), and 0.40 ± 0.14 (inter-hemisphere) under the eyes-closed condition, and slightly lower under the eyes-open condition with 0.37 ± 0.15 (left hemisphere), 0.36 ± 0.12 (right hemisphere), and 0.36 ± 0.17 (inter-hemisphere). The chart under each network depicts the averaged time courses in gray across supra-threshold in related ROIs for individual subjects; the grand average is shown in bold. The solid vertical line in each chart indicates the onset time of the eyes-open instruction. (Note: HV ≡ Human voice; AV ≡ Animal vocalization).

operational conditions (eyes-open or eyes-closed conditions). High concordance between the two analyses was observed in the primary auditory cortex, which was highly sensitive to the pitch and other acoustic features in sound clips.

4. Discussion

The auditory stimulation experiment closely resembled the natural sensory stimulation methods proposed in the literature (Hasson et al., 2004, 2010), wherein each subject received a unique random sequence of sound clips, which was presented twice under either eyes-closed or eyes-open conditions. As illustrated in Fig. 3, the intra-subject ICC analysis helped to disentangle brain areas sensitive to operational conditions (eyes-closed and eyes-open) from areas that were insensitive to these conditions. In contrast, assigning random sequences of sound clips to individual subjects made it possible to differentiate brain areas that were sensitive to random sequences from those that were independent. Nonetheless, there remained the possibility that random sequences under-predicted subjective (internal) appraisals of sound clips (e.g., emotional versus non-emotional sounds, and male versus female sounds). This issue would necessitate a detailed analysis of stress modulation and auditory language networks to enable scientific interpretation. Unlike the conventional practice of using the seed voxel for research into functional connectivity, the proposed two-stage methodology allows the use of mean time courses averaged across supra-threshold voxels in each nodal ROI. The averaged time course should be more reliable than a single time course in the seed voxel.

4.1. Stress modulation network

The limbic system and its associated structures have frequently been acknowledged as the major loci for investigating affective processes in patients and healthy controls (Adler et al., 2007; Berridge and Kringelbach, 2013; Frazier et al., 2005). Most natural sounds carry emotional meaning that may consciously or unconsciously induce affective processes in subjects (Fruhholz et al., 2014). Several studies have shown

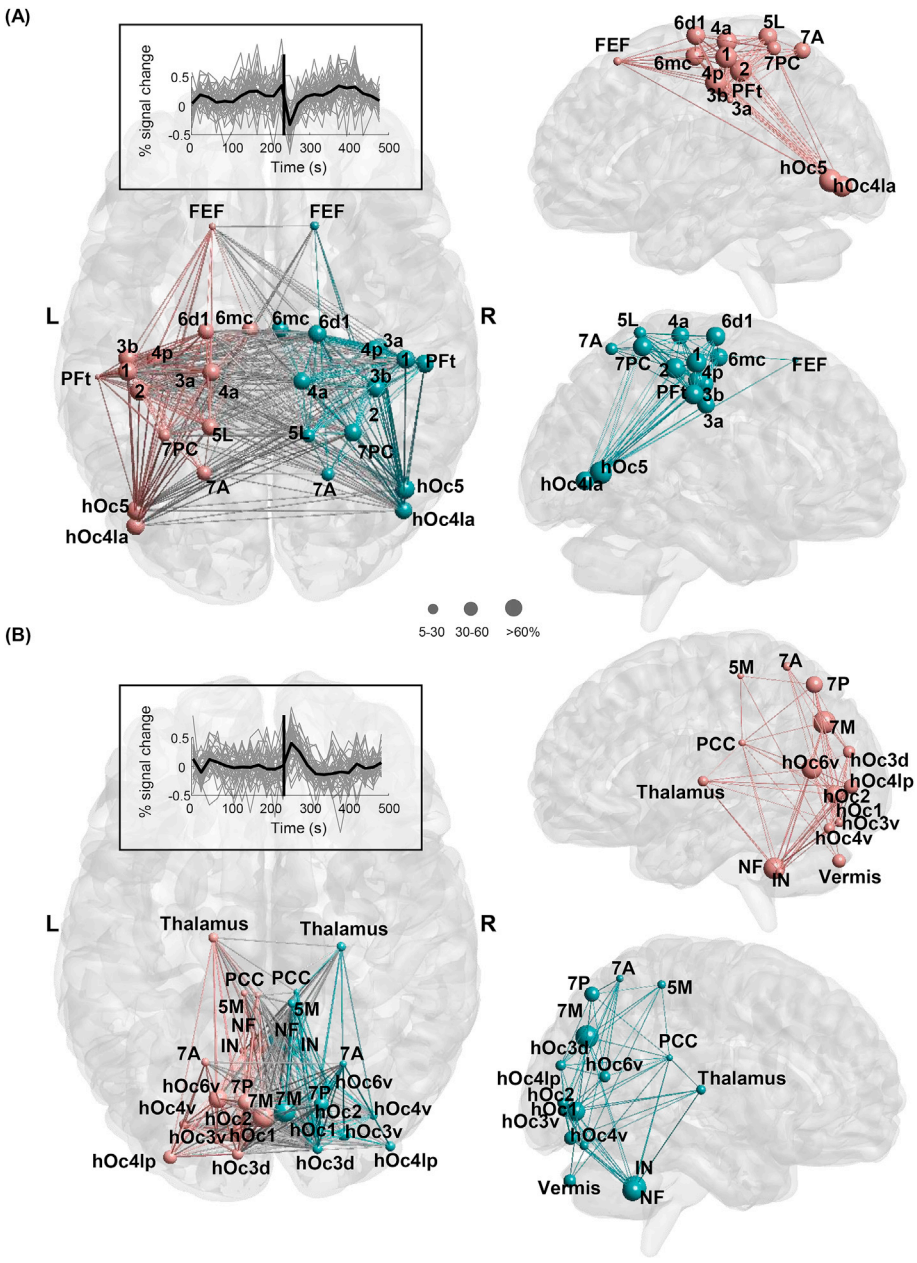

Fig. 5. The visuomotor control network in the inter-subject analysis. Areas in the sub-network in (A) include the primary motor cortex (4a and 4p), somatosensory cortex (1, 2, 3a, and 3b), higher visual cortex (hOc4la and hOc5), superior parietal cortex (5L, 7A, and 7 PC), inferior parietal cortex (PFt), areas of premotor cortex (6d1 and 6mc), and frontal eye field (FEF). These areas showed short-term decreases in activity following the eyes-open instruction. The average correlations between ROIs in this sub-network were 0.36 ± 0.12 (left hemisphere), 0.37 ± 0.12 (right hemisphere), and 0.35 ± 0.13 (inter-hemisphere) under the eyes-closed condition, and 0.36 ± 0.11 (left hemisphere), 0.33 ± 0.11 (right hemisphere), and 0.35 ± 0.12 (inter-hemisphere) under the eyes-open condition. Areas in the sub-network in (B) exhibited short-term increases in activity following the eyes-open instruction, including the thalamus, cerebellar vermis, nucleus fastigii (NF), interposed nuclei (IN), posterior and superior parietal cortex (5M, 7A, 7P, and 7M), occipital cortex (hOc1, hOc2, hOc3d, hOc3v, hOc4lp, hOc4v, and hOc6), and the posterior cingulate cortex (PCC). The average correlations in this sub-network were 0.39 ± 0.08 (left hemisphere), 0.33 ± 0.09 (right hemisphere), and 0.41 ± 0.07 (inter-hemisphere) under the eyes-closed condition, and 0.40 ± 0.10 (left hemisphere), 0.34 ± 0.11 (right hemisphere), and 0.42 ± 0.10 (inter-hemisphere) under the eyes-open condition. The averaged time courses across supra-threshold voxels are shown in gray for individual subjects; the grand average is shown in bold.

connectivity networks consistently distributed in the limbic system during the processing of various sounds (Escoffier et al., 2013; Fruhholz et al., 2014, 2016); however, fluctuations in networks caused by individual reactions to a given group of sounds can greatly compromise inter-subject ICCs. The inter-subject variability could result from individual reactions to extrinsic features in sound clips (random sequences) and subjective motivations. As shown in Fig. 4(A), the overall mean time course in the network exhibited prolonged decreases under the eyes-closed as well as the eyes-open conditions. Note that induced responses in 22 out of 40 subjects clearly elicited prolonged decreases in brain responses, and 18 out of those 22 subjects had previous experiences with fMRI. Thus, brain activity showing prolonged decreases would also be associated with the internal state of subjects during sensory stimulation, and be less affected by operational conditions.

Limbic deactivations have been reported in studies on stress provoking tasks (Pruessner et al., 2008), vagus nerve stimulation using audible 'OM' chanting (Kalyani et al., 2011), and bladder filling even with small volumes in patients with urge incontinence (Griffiths and Tadic, 2008). The limbic deactivations have been interpreted either as suppression of unwanted emotional reactions aroused by tasks or conversely as full relaxation responses. The stress modulation network defines tightly interconnected structures, which receive sensory inputs from cortical structures with updated outputs for regulating motor, autonomic, and hormonal functions (Lopes da Silva et al., 1990). During natural sound stimulation, subjects were required to passively listen to a variety of sounds without any behavioral responses. Other than the suppression of emotional reactions to sounds, it would also be reasonable to posit that the network had played a role in monitoring the external world through the use of sensory information (e.g., sounds and temperature), while subjects were relaxed and passively receiving sound stimulation. When sensory information became familiar (natural sounds) to the subjects, a prolonged decrease could reflect reduced connectivity within the network as well as reduced connectivity between the network and cortical structures (Barron et al., 2016). In Fig. 4(A), the network presented a decrease in correlations between ROIs (on average) when subjects were exposed to the same group of sounds a second time (i.e.,

eyes-open). In the study on OM chanting, all of the subjects were trained by an experienced yoga teacher prior to the fMRI experiment, and were fully familiar with the audible vowel (O) and consonant (M) parts of the OM. Constant monitoring of the external world is associated with motivational processes, which appeared to be strengthened by "natural sound stimulation" when subjects were relaxed and not performing any voluntary control tasks. This consideration could partially explain reactions involving prolonged decreases in the 22 subjects under the two conditions. Other subjects showed stronger reactions to specific sound clips, such as baby prattling, compared with the other sounds. On average, the group of subjects showed prolonged decreases under the two operational conditions.

4.2. Auditory language network

Although there was no explicit linguistic task involved in the sound stimulation, the auditory language network was identified with nodal ROIs lying below the diagonal in Fig. 2, including the inferior frontal, superior temporal and middle temporal gyri (Friederici, 2012). It has been known for some time that the language network can be induced in non-linguistic tasks. For example, self-referential tasks often induce lateralized activation in the dorsolateral prefrontal and temporo-parietal systems as a result of abstract language-based concepts of the self (Binder et al., 2009; Davey et al., 2016). Moreover, inner speech and overt speech share a common network (Aleman et al., 2005; Martin et al., 2014). It has also been observed that the language network is automatically induced when subjects perform visual object recognition tasks (Binkofski et al., 1999). Notably, the left Te3 (part of Wernicke's area) and left 45 (Broca's region) presented higher proportions of supra-threshold voxels than did the right Te3 and 45. The overall mean time course in the network exhibited far higher response amplitudes in response to human voices than to animal vocalizations, provided that there were insignificant differences in intensity, harmonics-to-noisy ratio, pitch, and other auditory features between the two types of sounds. The variability among subjects was also strong, particularly under the eyes-open condition.

Emotional vocalizations can be used to communicate across cultures (Sauter et al., 2010) and are important acoustic cues in human communication and social interaction (Scott et al., 2010). Emotional cues could be more salient in the sound clips of human voices selected for this experiment, such as the sounds of woman crying and woman laughing. However, further comparisons of response amplitudes between individual sound clips revealed that, on average, the sounds of woman crying and woman laughing did not induce higher activity than did the other human voices. Voice-selective areas such as the superior temporal sulcus are known to be more strongly activated when subjects listen passively to speech or non-speech human voices than when they listen to environmental sounds (Belin et al., 2000, 2004). In addition to the superior temporal areas, the inferior frontal gyrus is another important nodal area in the language network; it has also been reported to have significant activation during speech and non-speech stimulations (Fadiga et al., 2009). Unlike the stress modulation network, human voices, on average, exhibited higher response amplitudes than did the animal vocalizations in nearly every subject (i.e., 36 out of 40 subjects). One possible explanation of the functional selectivity for human voices may be mechanisms related to action, understanding, and imitation (Caspers et al., 2006; Gazzola et al., 2006; Molenberghs et al., 2012). Studies have reported mirror activation in the MTG and premotor cortex when viewing actions or hearing human sounds (Galati et al., 2008; Gazzola et al., 2006). These regions respond to action-execution as well as action-sounds and play a major role in understanding the actions of others (Thioux et al., 2008). The averaged time courses during human voice stimulation appeared to be non-specific to operational conditions. In other words, activation was more pronounced in response to human voices than to animal vocalizations under both conditions. Furthermore, activation tended to be specific to the differences in content between human voices and animal vocalizations based on the real-life experiences of subjects. The language network was similar to the stress modulation network in terms of decreased connectivity values when subjects were exposed to a given group of sounds a second time. As mentioned, the anterior insula demonstrated time courses similar to those in the language network. Recent studies have suggested that the anterior insula is activated bilaterally during overt speech and singing (Ackermann and Riecker, 2004) and plays an integrative role in speech processing (Oh et al., 2014; Rodriguez et al., 2018). The anterior insula contains the anatomical substrate for the evolved capacity of humans to be aware of themselves, others, and the environment (Craig and Craig, 2009); it could participate in the auditory language network supporting the processing of information related to natural sounds.

4.3. Visuomotor control network

The proposed method allows the inclusion of operational conditions in different repetitions of the sensory stimulation. Under the eyes-open condition, subjects were required to maintain eye-fixation on a cross-hair in the central screen while listening to sound stimuli. Among the brain areas lying above the diagonal in Fig. 2, two distinct groups of brain areas were sensitive to the two conditions. Those areas could be further delineated into those presenting a short-term increase in activity or a decrease in activity following the eyes-open instruction. One group could be associated with the voluntary control of saccadic eye movements including the FEF, primary motor cortex, somatosensory cortex, part of premotor cortex, part of inferior parietal cortex, higher visual cortex and superior parietal cortex. The other group could be associated with reflexive saccades, including the PCC (Pierrot-Deseilligny et al., 2005; Walker et al., 2000), thalamus, cerebellar vermis, nucleus fastigii and interposed nuclei, posterior and superior parietal cortex, and occipital cortex. As shown in Fig. 2, the inter-subject ICCs presented by these two sub-networks were higher than the intra-subject ICCs, indicating the effects of operational conditions. Sound stimulation did not require any behavioral responses; therefore, the areas associated with voluntary control over saccadic eye movements could be inhibited through reduced activity. The PCC is known to be active during reflexive saccades but not during voluntary saccades (Mort et al., 2003). Reflexive saccades involve visual perception, which would automatically engage the thalamus, cerebellum, and a major part of the occipital and parietal areas.

It is interesting to note that Area 7A exhibited increased (67.1%) as well as decreased (32.9%) activity following the eyes-open instruction in the inter-subject analysis as shown in Fig. 5(A) and (B). There were other areas showing distinct responses in the inter- or intra-subject analysis. Although these areas could engage at least two processes following the eyes-open instruction, it was also possible that the distinct temporal responses were the result of anatomical variability among subjects. The two-stage methodology further highlighted the importance of evaluating individual areas by comparing the results of inter- and intra-subject analyses. It is possible that brain areas associated with processing voluntary and reflexive saccades were independent of each other (Carrasco, 2011; Pinto et al., 2013), despite the high correlations between the two groups of brain areas shown in Fig. 5 (i.e., between-network correlations ranged from −0.37 to 0.73). The connectivity values between ROIs in visuomotor control networks were higher than those in other networks when the subjects were exposed to the same group of sounds a second time.

4.4. Limitations and challenges

At least three connectivity networks (stress modulation, auditory language, and visuomotor control) were engaged during auditory stimulation involving natural human voices and animal vocalizations that reflected real-life experiences. The superior temporal areas were engaged in processing low-level acoustic features (Bendor and Wang, 2006; Leaver and Rauschecker, 2010), whereas category-selective areas were engaged in sound discrimination (Doehrmann et al., 2008; Engel et al.,

2009; Leaver and Rauschecker, 2010; Lewis et al., 2005; Zhang et al., 2015). The three groups of ROIs were also engaged in processing other aspects of information under auditory stimulation. Although Pearson's correlations were used for deciding the connectivity maps in the second stage, other connectivity indices, such as mutual information and cross-covariance structures (Chai et al., 2009; Rykhlevskaia et al., 2006), could also be used to examine the causal effects among ROIs. The proposed two-stage approach might encounter the same methodological problems as other competing methods (e.g., independent component analysis) under long-term stimulation (i.e., >8 min), wherein a failure to correct for motion artifacts could compromise the results, particularly in the inter-subject analysis. It is recommended that, in long-term stimulation, the inter-subject analysis be applied to small groups of subjects and efforts be made to minimize motion artifacts. The resulting ICC maps may be used as a standard for processing data from the entire group of subjects. Note that the proposed methodology was able to distinguish between two network types in the $p_W > p_B$ case, that is, one network displaying moderate inter-subject variability (stress modulation) and the other displaying small inter-subject variability (auditory language). The mean time courses of 22 out of 40 subjects clearly exhibited prolonged decreases in the stress modulation network; however, the mean time courses of 36 out of 40 subjects showed stronger responses to human voices than to animal vocalizations in the auditory language network. These two types of processes were insensitive to the operational conditions, but they were sensitive to subjective motivations (e.g., long-term decreased activity in 22 subjects), or real-life experiences (i.e., stronger reactions to human voices than to animal vocalizations in 36 subjects).

Most fMRI studies on connectivity networks require well-defined ROIs. In the two-stage methodology, the classification of brain areas into concordant/discordant cases also depends on well-defined nodal ROIs. The JuBrain Cytoarchitectonic Atlas (Eickhoff et al., 2005) relies on microscopically defined cytoarchitectonic features in brain structures, in contrast to other publicly available atlases based on macroscopic characteristics like sulci and gyri such as the AAL (Tzourio-Mazoyer et al., 2002), Neuromorphometrics, Inc., and ICBM (Mazziotta et al., 1995) Atlases. The definition of regional boundaries in a cytoarchitectonic atlas is observer-independent and statistically determined (Amunts and Zilles, 2015), the use of which requires a high degree of rigor in preprocessing fMRI data, as noted in previous studies (Clos et al., 2014; Xu et al., 2016). Brain areas defined according to their microstructures (regarded as functional modules) have been widely used in studies on functional connectivity (Cole et al., 2012; Eickhoff et al., 2011; Uddin et al., 2010; Umarova et al., 2009). Benefiting from histological parcellations in the JuBrain Atlas, the two-stage procedure differentiated two sub-networks in the occipital and parietal cortices according to whether they exhibited increased or decreased activity following the eyes-open instruction. One group connected the superior parietal cortex (5M, 7A, 7P, and 7M) and occipital cortex (hOc1, hOc2, hOc3d, hOc3v, hOc4lp, hOc4v, and hOc6v). The other group connected the higher visual cortex (hOc4la and hOc5/V5) and a part of the superior parietal cortex (5L, 7A, and 7 PC). In contrast, when using macroscopic atlases, most brain areas were clustered in the bottom left corner in the bivariate plots showing the proportions of supra-threshold voxels in intra- and inter-subject analyses. This made the connectivity networks more difficult to identify (Note: The bivariate plots based on other atlases are available from the first and corresponding authors upon request). In the example in Fig. 2, the subgenual anterior cingulate included 52.2% supra-threshold voxels in the inter-subject analysis and 65.2% in the intra-subject analysis. This important region has previously been linked to fear, negative emotions, and predicting errors in social interactions (Hornboll et al., 2018; Lockwood and Wittmann, 2018; Palomero-Gallagher et al., 2015). During real-life stimulation, Area s32 presented higher proportions of supra-threshold voxels (71.5% in inter-subject and 86.9% in intra-subject analyses) than did Areas s24, and 25. This type of information should provide a target substrate for the comparison of connectivity changes in patients and healthy controls. Therefore, we recommend the use of the JuBrain Atlas in conjunction with the two-stage approach for the connectivity network study.

5. Conclusions

In this study, we developed a two-stage methodology for assessing functional networks in subcortical and cortical brain structures during natural sensory stimulation. In the first stage, the concordance analysis is used to assess the degree of agreement between the results obtained using intra- and inter-subject ICC analyses on the same dataset. In the second stage, connectivity networks are constructed based on supra-threshold time courses provided in the first stage. Under natural auditory stimulation using human voices and animal vocalizations, our results demonstrated that the highest concordance between intra- and inter-subject analyses was observed in the primary auditory cortex, which was particularly sensitive to the pitch of sound clips. Three connectivity networks were identified among the cases of discordance: stress modulation (affective), auditory language, and visuomotor control. These networks were generally sensitive to subjective motivations, real-life experiences, and the eyes-open instruction. Based on the findings in this study, it appears that designating brain areas as concordant/discordant may enhance the precision in assessing connectivity networks involved in the processing of internal/external information related to natural sensory stimulation.

Declaration of interest

None declared.

Acknowledgement

We are indebted to three reviewers and Professors Stephen M. Smith and Michael Breakspear for helpful comments on the earlier draft of this manuscript, to Professor SH Annabel Chen for recommending the fMRI scanning sequences suitable for real-life experiments, and to the Research Center for Mind, Brain & Learning at the National Chengchi University for the great assistance during fMRI data acquisition. This research was supported by grants MOST-103-2410-H-001-058-MY2 and MOST-106-2410-H-001-026 from the Ministry of Science and Technology, Taiwan, and a grant from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 785907 (HBP SGA2).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.116042.

APPENDIX

A. Variance of $\widehat{\mathrm{ICC}}$ (Proof of Equation (2))

Assume that $\mathbf{\Sigma}$ is positive definite, and the variance of $\widehat{\mathrm{ICC}}$ can be estimated by

$$\mathrm{Var}(\widehat{\mathrm{ICC}}) = 2n^{-1}\eta' \mathrm{K}_M(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma})\mathrm{K}'_M\eta|_{\boldsymbol{\Sigma}=\mathbf{S}} \tag{A.1}$$

The derivative of ICC with respect to $\mathrm{vech}'\boldsymbol{\Sigma}$ in the delta method includes the following factors for use:

$$\partial \mathrm{tr}(\boldsymbol{\Sigma})/\partial \mathrm{vech}'\boldsymbol{\Sigma} = \partial \mathrm{vech}'\mathrm{I}_M\mathrm{G}_M\mathrm{vech}\boldsymbol{\Sigma}/\partial \mathrm{vech}'\boldsymbol{\Sigma} = \mathrm{vech}'\mathrm{I}_M\mathrm{G}_M \tag{A.2}$$

$$\partial(\underline{\mathbf{1}}'\boldsymbol{\Sigma}\underline{\mathbf{1}})/\partial \mathrm{vech}'\boldsymbol{\Sigma} = \partial(\underline{\mathbf{1}}'\otimes\underline{\mathbf{1}}')\mathrm{G}_M\mathrm{vech}\boldsymbol{\Sigma}/\partial \mathrm{vech}'\boldsymbol{\Sigma} = (\underline{\mathbf{1}}'\otimes\underline{\mathbf{1}}')\mathrm{G}_M \tag{A.3}$$

where $\underline{\mathbf{1}}$ in (A.3) is a summing vector of order M. By (A.2) and (A.3), η' for estimating the variance of $\widehat{\mathrm{ICC}}$ can be derived as follows:

$$\begin{aligned}
\eta' &= \partial \mathrm{ICC}/\partial \mathrm{vech}'\boldsymbol{\Sigma} \\
&= \frac{-M}{M-1}\left\{ (\underline{\mathbf{1}}'\boldsymbol{\Sigma}\underline{\mathbf{1}})^{-1}\left[\partial \mathrm{tr}(\boldsymbol{\Sigma})/\partial \mathrm{vech}'\boldsymbol{\Sigma}\right] + \mathrm{tr}(\boldsymbol{\Sigma})\left[\partial(\underline{\mathbf{1}}'\boldsymbol{\Sigma}\underline{\mathbf{1}})^{-1}/\partial \mathrm{vech}'\boldsymbol{\Sigma}\right] \right\} \\
&= \frac{-M}{M-1}\left\{ (\underline{\mathbf{1}}'\boldsymbol{\Sigma}\underline{\mathbf{1}})^{-1}[\mathrm{vech}'\mathrm{I}_M G_M] - \mathrm{tr}(\boldsymbol{\Sigma})(\underline{\mathbf{1}}'\boldsymbol{\Sigma}\underline{\mathbf{1}})^{-2}[(\underline{\mathbf{1}}'\otimes\underline{\mathbf{1}}')G_M] \right\} \\
&= \frac{M}{M-1}(\underline{\mathbf{1}}'\boldsymbol{\Sigma}\underline{\mathbf{1}})^{-2}\left\{ -(\underline{\mathbf{1}}'\boldsymbol{\Sigma}\underline{\mathbf{1}})\mathrm{vech}'\mathrm{I}_M + \mathrm{tr}(\boldsymbol{\Sigma})(\underline{\mathbf{1}}'\otimes\underline{\mathbf{1}}') \right\} G_M
\end{aligned} \tag{A.4}$$

where I_M is the identity matrix of order M. The variance of $\widehat{\mathrm{ICC}}$ can be estimated by substituting $\mathbf{S}$ for $\boldsymbol{\Sigma}$ in (A.1) and (A.4), and plugging η' into (A.1).

The asymptotic formula (A.1) is valid under the conditions that the dependent vector responses across the n image scans in each session have bounded fourth moments and satisfy two basic conditions: (i) strictly stationary along the sequence of image scans, and (ii) ρ-mixing, that is, the maximal modulus of lagged correlations between $s_{i,j}^{(q)}$ and $s_{k,l}^{(q+v)}$ converge toward 0 for $1 \le i, j, k, l \le \mathrm{M}$, $q \ge 1$, $1 \le v < n$, and $\mathbf{S} = (s_{ij})$, or $\max_{i,j,k,l}^{(q)}\left|\rho\left(s_{i,j}^{(q)}, s_{k,l}^{(q+v)}\right)\right| \to 0$, as $v \to \infty$. It follows from Theorem 1.3 in Bradley (2012) that the stationary elements $s_{i,j}$ satisfy asymptotic normality of Equation (2) in Section 2.4.1. To validate the application of (A.1), the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test was applied to individual time courses for all in-brain voxels to evaluate the null hypothesis that time courses are trend-stationary against the alternative of a unit root (Kwiatkowski et al., 1992). In addition, the Augmented Dickey–Fuller test (ADF) was also applied to evaluate the null hypothesis of non-stationarity against the alternative hypothesis of stationarity. A time course was assumed to be stationary if the KPSS test did not reject the null hypothesis, and the ADF test rejected the null hypothesis. Time courses of all in-brain voxels in every subject were rejected by the ADF test. This result suggested that BOLD time courses were stationary. Without applying the FDR control, the time courses of 0.1% voxels were consistently rejected by the KPSS test in at least 70% of the subjects (0.6% voxels in at least 60% of the subjects, 1.98% voxels in at least 50% of the subjects). These voxels were distributed mainly in Areas Fo1, Fo2, Fo3, Fp1, Fp2, and s32. However, under the weak FDR control at $\alpha = 0.05$, none of these time courses exceeded the FDR threshold. From this, it can be concluded that the assumption of stationarity was not violated in most in-brain voxels.

B. Variance of the weighted average of $\widehat{\mathrm{ICC}}_j$ (Proof of Equation (6))

Denote $\widehat{\mathrm{ICC}}_j$ and $\mathrm{Var}(\widehat{\mathrm{ICC}}_j)$ to be the estimates in the j-th subject. The weighted average across subjects is

$$\widehat{\mathrm{ICC}}_W = \sum_{j=1}^{\mathrm{N}} \widehat{\omega}_j \widehat{\mathrm{ICC}}_j, \quad \text{and}$$

$$\widehat{\omega}_j = \frac{1}{\mathrm{Var}\left(\widehat{\mathrm{ICC}}_j\right)}\left(\sum_{j=1}^{\mathrm{N}} \frac{1}{\mathrm{Var}\left(\widehat{\mathrm{ICC}}_j\right)}\right)^{-1}$$

with $\sum_{j=1}^{\mathrm{N}} \widehat{\omega}_j = 1$. When n (number of entries in the BOLD time course) is large, it can be shown that $\widehat{\omega}_j = \omega_j + \mathrm{o}(1/\sqrt{n})$ where $\mathrm{o}(1/\sqrt{n})$ indicates that $\widehat{\omega}_j$ converges to ω_j in mean-square with negligible error of order $\mathrm{o}(1/n)$ by finite fourth moments and stationarity; that is,

$$\mathrm{Var}\left(\widehat{\mathrm{ICC}}_W\right) = \sum_{j=1}^{\mathrm{N}}\left(\widehat{\omega}_j\right)^2 \mathrm{Var}\left(\widehat{\mathrm{ICC}}_j\right) + \mathrm{o}\left(\frac{1}{n}\right) \tag{A.5}$$

$$\approx \sum_{j=1}^{\mathrm{N}}\left[\left(\frac{1}{\mathrm{Var}\left(\widehat{\mathrm{ICC}}_j\right)}\right)^2\left(\sum_{j=1}^{\mathrm{N}}\frac{1}{\mathrm{Var}\left(\widehat{\mathrm{ICC}}_j\right)}\right)^{-2}\mathrm{Var}\left(\widehat{\mathrm{ICC}}_j\right)\right]$$

$$\approx \sum_{j=1}^{\mathrm{N}}\left[\left(\frac{1}{\mathrm{Var}\left(\widehat{\mathrm{ICC}}_j\right)}\right)\left(\sum_{j=1}^{\mathrm{N}}\frac{1}{\mathrm{Var}\left(\widehat{\mathrm{ICC}}_j\right)}\right)^{-2}\right]$$

$$\approx \left(\sum_{j=1}^{N} \frac{1}{\mathrm{Var}(\widehat{\mathrm{ICC}}_j)}\right)\left(\sum_{j=1}^{N} \frac{1}{\mathrm{Var}(\widehat{\mathrm{ICC}}_j)}\right)^{-2}$$

$$\approx \left(\sum_{j=1}^{N} \frac{1}{\mathrm{Var}(\widehat{\mathrm{ICC}}_j)}\right)^{-1}.$$

Supplementary Figure S8(A) depicts the bivariate plots of $\sqrt{\mathrm{Var}(\widehat{\mathrm{ICC}})}$ (vertical axis) versus $\widehat{\mathrm{ICC}}$ values (horizontal axis) for all in-brain voxels in the inter-subject analysis. Supplementary Figure S8(B) depicts the bivariate plots of $\sqrt{\mathrm{Var}(\widehat{\mathrm{ICC}}_W)}$ versus $\widehat{\mathrm{ICC}}_W$ in the intra-subject analysis. It is remarkable that the standard error estimates generally decrease as the $\widehat{\mathrm{ICC}}$ or $\widehat{\mathrm{ICC}}_W$ values increase toward 1.0, and values in Figure S8 have confirmed that the bounded ICC variance estimates are of order $O(1/n)$, which echoes the validity of the ρ-mixing condition. In other words, the assumptions of stationarity and ρ-mixing sequences were not violated in the dataset.

References

Ackermann, H., Riecker, A., 2004. The contribution of the insula to motor aspects of speech production: a review and a hypothesis. Brain Lang. 89, 320–328.

Adler, C.M., DelBello, M.P., Jarvis, K., Levine, A., Adams, J., Strakowski, S.M., 2007. Voxel-based study of structural changes in first-episode patients with bipolar disorder. Biol. Psychiatry 61, 776–781.

Aleman, A., Formisano, E., Koppenhagen, H., Hagoort, P., de Haan, E.H., Kahn, R.S., 2005. The functional neuroanatomy of metrical stress evaluation of perceived and imagined spoken words. Cerebr. Cortex 15, 221–228.

Amunts, K., Kedo, O., Kindler, M., Pieperhoff, P., Mohlberg, H., Shah, N., Habel, U., Schneider, F., Zilles, K., 2005. Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. Anat. Embryol. 210, 343–352.

Amunts, K., Malikovic, A., Mohlberg, H., Schormann, T., Zilles, K., 2000. Brodmann's areas 17 and 18 brought into stereotaxic space—where and how variable? Neuroimage 11, 66–84.

Amunts, K., Schleicher, A., Burgel, U., Mohlberg, H., Uylings, H.B., Zilles, K., 1999. Broca's region revisited: cytoarchitecture and intersubject variability. J. Comp. Neurol. 412, 319–341.

Amunts, K., Zilles, K., 2015. Architectonic mapping of the human brain beyond Brodmann. Neuron 88, 1086–1107.

Barron, H.C., Garvert, M.M., Behrens, T.E., 2016. Repetition suppression: a means to index neural representations using BOLD? Philos. Trans. R. Soc. Biol. Sci. 371, 20150355.

Belin, P., Fecteau, S., Bedard, C., 2004. Thinking the voice: neural correlates of voice perception. Trends Cogn. Sci. 8, 129–135.

Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. Nature 403, 309–312.

Bendor, D., Wang, X., 2005. The neuronal representation of pitch in primate auditory cortex. Nature 436, 1161–1165.

Bendor, D., Wang, X., 2006. Cortical representations of pitch in monkeys and humans. Curr. Opin. Neurobiol. 16, 391–399.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Ser. Soc. B Stat. Methodol. 57, 289–300.

Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in multiple testing under dependency. Ann. Stat. 29, 1165–1188.

Berkovich-Ohana, A., Glicksohn, J., 2014. The consciousness state space (CSS)—a unifying model for consciousness and self. Front. Psychol. 5, 341.

Bermudez, J., Krizaj, D., Lipschitz, D.L., Bueler, C.E., Rogowska, J., Yurgelun-Todd, D., Nakamura, Y., 2017. Externally-induced meditative states: an exploratory fMRI study of architects' responses to contemplative architecture. Front. Archit. Res. 6, 123–136.

Berridge, K.C., Kringelbach, M.L., 2013. Neuroscience of affect: brain mechanisms of pleasure and displeasure. Curr. Opin. Neurobiol. 23, 294–303.

Betti, V., Della Penna, S., de Pasquale, F., Mantini, D., Marzetti, L., Romani, G.L., Corbetta, M., 2013. Natural scenes viewing alters the dynamics of functional connectivity in the human brain. Neuron 79, 782–797.

Binder, J.R., Desai, R.H., Graves, W.W., Conant, L.L., 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cerebr. Cortex 19, 2767–2796.

Binkofski, F., Buccino, G., Posse, S., Seitz, R.J., Rizzolatti, G., Freund, H., 1999. A fronto-parietal circuit for object manipulation in man: evidence from an fMRI-study. Eur. J. Neurosci. 11, 3276–3286.

Bludau, S., Eickhoff, S.B., Mohlberg, H., Caspers, S., Laird, A.R., Fox, P.T., Schleicher, A., Zilles, K., Amunts, K., 2014. Cytoarchitecture, probability maps and functions of the human frontal pole. Neuroimage 93 Pt 2, 260–275.

Bradley, R.C., 2012. On the behavior of the covariance matrices in a multivariate central limit theorem under some mixing conditions. Ill. J. Math. 56, 677–704.

Brandt, D.J., Sommer, J., Krach, S., Bedenbender, J., Kircher, T., Paulus, F.M., Jansen, A., 2013. Test-retest reliability of fMRI brain activity during memory encoding. Front. Psychiatry 4, 163.

Braun, U., Plichta, M.M., Esslinger, C., Sauer, C., Haddad, L., Grimm, O., Mier, D., Mohnke, S., Heinz, A., Erk, S., Walter, H., Seiferth, N., Kirsch, P., Meyer-Lindenberg, A., 2012. Test-retest reliability of resting-state connectivity network characteristics using fMRI and graph theoretical measures. Neuroimage 59, 1404–1412.

Brechmann, A., Baumgart, F., Scheich, H., 2002. Sound-level-dependent representation of frequency modulations in human auditory cortex: a low-noise fMRI study. J. Neurophysiol. 87, 423–433.

Breslow, N.E., Day, N.E., Davis, W., 1987. Statistical Methods in Cancer Research. International Agency for Research on Cancer Lyon.

Brodmann, K., 1909. Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues. Barth.

Browne, M.W., 1974. Generalized least squares estimators in the analysis of covariance structures. S. Afr. Stat. J. 8, 1–24.

Carrasco, M., 2011. Visual attention: the past 25 years. Vis. Res. 51, 1484–1525.

Caspers, S., Geyer, S., Schleicher, A., Mohlberg, H., Amunts, K., Zilles, K., 2006. The human inferior parietal cortex: cytoarchitectonic parcellation and interindividual variability. Neuroimage 33, 430–448.

Cerqueira, J.J., Almeida, O.F., Sousa, N., 2008. The stressed prefrontal cortex. Left? Right! Brain Behav. Immun. 22, 630–638.

Cerqueira, J.J., Pêgo, J.M., Taipa, R., Bessa, J.M., Almeida, O.F., Sousa, N., 2005. Morphological correlates of corticosteroid-induced changes in prefrontal cortex-dependent behaviors. J. Neurosci. 25, 7792–7800.

Chai, B., Walther, D., Beck, D., Fei-Fei, L., 2009. Exploring functional connectivities of the human brain using multivariate information analysis. In: Proc Adv Neural Inf Process Syst, pp. 270–278.

Chen, J.L., Zatorre, R.J., Penhune, V.B., 2006. Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. Neuroimage 32, 1771–1781.

Clos, M., Diederen, K.M., Meijering, A.L., Sommer, I.E., Eickhoff, S.B., 2014. Aberrant connectivity of areas for decoding degraded speech in patients with auditory verbal hallucinations. Brain Struct. Funct. 219, 581–594.

Cole, M.W., Yarkoni, T., Repovs, G., Anticevic, A., Braver, T.S., 2012. Global connectivity of prefrontal cortex predicts cognitive control and intelligence. J. Neurosci. 32, 8988–8999.

Craig, A.D., Craig, A., 2009. How do you feel–now? The anterior insula and human awareness. Nat. Rev. Neurosci. 10, 59–70.

Culham, J.C., Cavina-Pratesi, C., Singhal, A., 2006. The role of parietal cortex in visuomotor control: what have we learned from neuroimaging? Neuropsychologia. 44 (13), 2668–2684.

Davey, C.G., Pujol, J., Harrison, B.J., 2016. Mapping the self in the brain's default mode network. Neuroimage 132, 390–397.

Doehrmann, O., Naumer, M.J., Volz, S., Kaiser, J., Altmann, C.F., 2008. Probing category selectivity for environmental sounds in the human auditory brain. Neuropsychologia 46, 2776–2786.

Eickhoff, S.B., Bzdok, D., Laird, A.R., Roski, C., Caspers, S., Zilles, K., Fox, P.T., 2011. Co-activation patterns distinguish cortical modules, their connectivity and functional differentiation. Neuroimage 57, 938–949.

Eickhoff, S.B., Constable, R.T., Yeo, B.T., 2018. Topographic organization of the cerebral cortex and brain cartography. Neuroimage 170, 332–347.

Eickhoff, S.B., Heim, S., Zilles, K., Amunts, K., 2006a. Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. Neuroimage 32, 570–582.

Eickhoff, S.B., Schleicher, A., Zilles, K., Amunts, K., 2006b. The human parietal operculum. I. Cytoarchitectonic mapping of subdivisions. Cerebr. Cortex 16, 254–267.

Eickhoff, S.B., Stephan, K.E., Mohlberg, H., Grefkes, C., Fink, G.R., Amunts, K., Zilles, K., 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25, 1325–1335.

Engel, L.R., Frum, C., Puce, A., Walker, N.A., Lewis, J.W., 2009. Different categories of living and non-living sound-sources activate distinct cortical networks. Neuroimage 47, 1778–1791.

Escoffier, N., Zhong, J., Schirmer, A., Qiu, A., 2013. Emotional expressions in voice and music: same code, same effect? Hum. Brain Mapp. 34, 1796–1810.

Fadiga, L., Craighero, L., D'Ausilio, A., 2009. Broca's area in language, action, and music. Ann. N. Y. Acad. Sci. 1169, 448–458.

Fagerland, M.W., Lydersen, S., Laake, P., 2013. The McNemar test for binary matched-pairs data: mid-p and asymptotic are better than exact conditional. BMC Med. Res. Methodol. 13, 91.
Fox, M.D., Zhang, D., Snyder, A.Z., Raichle, M.E., 2009. The global signal and observed anticorrelated resting state brain networks. J. Neurophysiol. 101, 3270–3283.
Franco, A.R., Pritchard, A., Calhoun, V.D., Mayer, A.R., 2009. Interrater and intermethod reliability of default mode network selection. Hum. Brain Mapp. 30, 2293–2303.
Frazier, J.A., Chiu, S., Breeze, J.L., Makris, N., Lange, N., Kennedy, D.N., Herbert, M.R., Bent, E.K., Koneru, V.K., Dieterich, M.E., 2005. Structural brain magnetic resonance imaging of limbic and thalamic volumes in pediatric bipolar disorder. Am. J. Psychiatry 162, 1256–1265.
Friederici, A.D., 2012. The cortical language circuit: from auditory perception to sentence comprehension. Trends Cogn. Sci. 16, 262–268.
Friederici, A.D., Alter, K., 2004. Lateralization of auditory language functions: a dynamic dual pathway model. Brain Lang. 89, 267–276.
Friedman, L., Stern, H., Brown, G.G., Mathalon, D.H., Turner, J., Glover, G.H., Gollub, R.L., Lauriello, J., Lim, K.O., Cannon, T., Greve, D.N., Bockholt, H.J., Belger, A., Mueller, B., Doty, M.J., He, J., Wells, W., Smyth, P., Pieper, S., Kim, S., Kubicki, M., Vangel, M., Potkin, S.G., 2008. Test-retest and between-site reliability in a multicenter fMRI study. Hum. Brain Mapp. 29, 958–972.
Fruhholz, S., Trost, W., Grandjean, D., 2014. The role of the medial temporal limbic system in processing emotions in voice and music. Prog. Neurobiol. 123, 1–17.
Fruhholz, S., Trost, W., Kotz, S.A., 2016. The sound of emotions-Towards a unifying neural network perspective of affective sound processing. Neurosci. Biobehav. Rev. 68, 96–110.
Galati, G., Committeri, G., Spitoni, G., Aprile, T., Di Russo, F., Pitzalis, S., Pizzamiglio, L., 2008. A selective representation of the meaning of actions in the auditory mirror system. Neuroimage 40, 1274–1286.
Gazzola, V., Aziz-Zadeh, L., Keysers, C., 2006. Empathy and the somatotopic auditory mirror system in humans. Curr. Biol. 16, 1824–1829.
Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage 15, 870–878.
Geyer, S., 2004. The microstructural border between the motor and the cognitive domain in the human cerebral cortex. Adv. Anat. Embryol. Cell Biol. 174 (I-VIII), 1–89.
Geyer, S., Ledberg, A., Schleicher, A., Kinomura, S., Schormann, T., Burgel, U., Klingberg, T., Larsson, J., Zilles, K., Roland, P.E., 1996. Two different areas within the primary motor cortex of man. Nature 382, 805–807.
Geyer, S., Schormann, T., Mohlberg, H., Zilles, K., 2000. Areas 3a, 3b, and 1 of human primary somatosensory cortex. Part 2. Spatial normalization to standard anatomical space. Neuroimage 11, 684–696.
Golland, Y., Bentin, S., Gelbard, H., Benjamini, Y., Heller, R., Nir, Y., Hasson, U., Malach, R., 2007. Extrinsic and intrinsic systems in the posterior cortex of the human brain revealed during natural sensory stimulation. Cerebr. Cortex 17, 766–777.
Grefkes, C., Geyer, S., Schormann, T., Roland, P., Zilles, K., 2001. Human somatosensory area 2: observer-independent cytoarchitectonic mapping, interindividual variability, and population map. Neuroimage 14, 617–631.
Griffiths, D., Tadic, S.D., 2008. Bladder control, urgency, and urge incontinence: evidence from functional brain imaging. Neurourol. Urodyn. 27, 466–474.
Hacker, C.D., Snyder, A.Z., Pahwa, M., Corbetta, M., Leuthardt, E.C., 2017. Frequency-specific electrophysiologic correlates of resting state fMRI networks. Neuroimage 149, 446–457.
Hall, D.A., Plack, C.J., 2009. Pitch processing sites in the human auditory brain. Cerebr. Cortex 19, 576–585.
Hamilton, J.D., 1995. Time Series Analysis. Economic Theory. II. Princeton University Press, USA, pp. 625–630.
Hanke, M., Baumgartner, F.J., Ibe, P., Kaule, F.R., Pollmann, S., Speck, O., Zinke, W., Stadler, J., 2014. A high-resolution 7-Tesla fMRI dataset from complex natural stimulation with an audio movie. Sci. Data 1, 140003.
Harmelech, T., Friedman, D., Malach, R., 2015. Differential magnetic resonance neurofeedback modulations across extrinsic (visual) and intrinsic (default-mode) nodes of the human cortex. J. Neurosci. 35, 2588–2595.
Hasson, U., Malach, R., Heeger, D.J., 2010. Reliability of cortical activity during natural stimulation. Trends Cogn. Sci. 14, 40–48.
Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Intersubject synchronization of cortical activity during natural vision. Science 303, 1634–1640.
Hasson, U., Yang, E., Vallines, I., Heeger, D.J., Rubin, N., 2008. A hierarchy of temporal receptive windows in human cortex. J. Neurosci. 28, 2539–2550.
Henderson, H.V., Searle, S., 1979. Vec and vech operators for matrices, with some uses in Jacobians and multivariate statistics. Can. J. Stat. 7, 65–81.
Henssen, A., Zilles, K., Palomero-Gallagher, N., Schleicher, A., Mohlberg, H., Gerboga, F., Eickhoff, S.B., Bludau, S., Amunts, K., 2016. Cytoarchitecture and probability maps of the human medial orbitofrontal cortex. Cortex 75, 87–112.
Hornboll, B., Macoveanu, J., Nejad, A., Rowe, J., Elliott, R., Knudsen, G.M., Siebner, H.R., Paulson, O.B., 2018. Neuroticism predicts the impact of serotonin challenges on fear processing in subgenual anterior cingulate cortex. Sci. Rep. 8, 17889.
Humphries, C., Liebenthal, E., Binder, J.R., 2010. Tonotopic organization of human auditory cortex. Neuroimage 50, 1202–1211.
Iwashita, T., Siotani, M., 1994. Asymptotic distributions of functions of a sample covariance matrix under the elliptical distribution. Can. J. Stat. 22, 273–283.
Jansma, J.M., Ramsey, N.F., Coppola, R., Kahn, R.S., 2000. Specific versus nonspecific brain activity in a parametric N-back task. Neuroimage 12, 688–697.
Kalyani, B.G., Venkatasubramanian, G., Arasappa, R., Rao, N.P., Kalmady, S.V., Behere, R.V., Rao, H., Vasudev, M.K., Gangadhar, B.N., 2011. Neurohemodynamic correlates of 'OM' chanting: a pilot functional magnetic resonance imaging study. Int. J. Yoga 4, 3.
Kauppi, J.P., Jaaskelainen, I.P., Sams, M., Tohka, J., 2010. Inter-subject correlation of brain hemodynamic responses during watching a movie: localization in space and frequency. Front. Neuroinf. 4, 5.
Kleiser, R., Stadler, C., Wimmer, S., Matyas, T., Seitz, R.J., 2017. An fMRI study of training voluntary smooth circular eye movements. Exp. Brain Res. 235, 819–831.
Kong, J., Gollub, R.L., Webb, J.M., Kong, J.-T., Vangel, M.G., Kwong, K., 2007. Test–retest study of fMRI signal change evoked by electroacupuncture stimulation. Neuroimage 34, 1171–1181.
Konrad, K., Neufang, S., Thiel, C.M., Specht, K., Hanisch, C., Fan, J., Herpertz-Dahlmann, B., Fink, G.R., 2005. Development of attentional networks: an fMRI study with children and adults. Neuroimage 28, 429–439.
Kristo, G., Rutten, G.J., Raemaekers, M., de Gelder, B., Rombouts, S.A., Ramsey, N.F., 2014. Task and task-free FMRI reproducibility comparison for motor network identification. Hum. Brain Mapp. 35, 340–352.
Kujovic, M., Zilles, K., Malikovic, A., Schleicher, A., Mohlberg, H., Rottschy, C., Eickhoff, S.B., Amunts, K., 2013. Cytoarchitectonic mapping of the human dorsal extrastriate cortex. Brain Struct. Funct. 218, 157–172.
Kuo, P.-C., Tseng, Y.-L., Zilles, K., Suen, S., Eickhoff, S.B., Lee, J.-D., Cheng, P.E., Liou, M., 2019. Listening to Real-World Sounds: fMRI Data for Analyzing Connectivity Networks. Data in Brief (In press).
Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. J. Econom. 54, 159–178.
Langers, D.R., Jansen, J.F., Backes, W.H., 2007. Enhanced signal detection in neuroimaging by means of regional control of the global false discovery rate. Neuroimage 38, 43–56.
Lartillot, O., Toiviainen, P., 2007. A Matlab toolbox for musical feature extraction from audio. In: Proc Int Conf Digital Audio Effects, pp. 237–244. Bordeaux.
Leaver, A.M., Rauschecker, J.P., 2010. Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. J. Neurosci. 30, 7604–7612.
Lerner, Y., Honey, C.J., Silbert, L.J., Hasson, U., 2011. Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. J. Neurosci. 31, 2906–2915.
Lewis, J.W., Brefczynski, J.A., Phinney, R.E., Janik, J.J., DeYoe, E.A., 2005. Distinct cortical pathways for processing tool versus animal sounds. J. Neurosci. 25, 5148–5158.
Liou, M., Savostyanov, A.N., Simak, A.A., Wu, W.-C., Huang, C.-T., Cheng, P.E., 2012. An information system in the brain: evidence from fMRI BOLD responses. Chin. J. Psychol. 54, 1–26.
Lockwood, P.L., Wittmann, M.K., 2018. Ventral anterior cingulate cortex in social decision-making. Neurosci. Biobehav. Rev. 92, 187–191.
Lopes da Silva, F., Witter, M.P., Boeijinga, P.H., Lohman, A., 1990. Anatomic organization and physiology of the limbic cortex. Physiol. Rev. 70, 453–511.
Malikovic, A., Amunts, K., Schleicher, A., Mohlberg, H., Eickhoff, S.B., Wilms, M., Palomero-Gallagher, N., Armstrong, E., Zilles, K., 2007. Cytoarchitectonic analysis of the human extrastriate cortex in the region of V5/MT+: a probabilistic, stereotaxic map of area hOc5. Cerebr. Cortex 17, 562–574.
Malikovic, A., Amunts, K., Schleicher, A., Mohlberg, H., Kujovic, M., Palomero-Gallagher, N., Eickhoff, S.B., Zilles, K., 2016. Cytoarchitecture of the human lateral occipital cortex: mapping of two extrastriate areas hOc4la and hOc4lp. Brain Struct. Funct. 221, 1877–1897.
Martin, S., Brunner, P., Holdgraf, C., Heinze, H.J., Crone, N.E., Rieger, J., Schalk, G., Knight, R.T., Pasley, B.N., 2014. Decoding spectrotemporal features of overt and covert speech from the human cortex. Front. Neuroeng. 7, 14.
Mazziotta, J.C., Toga, A.W., Evans, A., Fox, P., Lancaster, J., 1995. A probabilistic atlas of the human brain: theory and rationale for its development: the international consortium for brain mapping (ICBM). Neuroimage 2, 89–101.
McGraw, K.O., Wong, S.P., 1996. Forming inferences about some intraclass correlation coefficients. Psychol. Methods 1, 30–46.
Molenberghs, P., Cunnington, R., Mattingley, J.B., 2012. Brain regions with mirror properties: a meta-analysis of 125 human fMRI studies. Neurosci. Biobehav. Rev. 36, 341–349.
Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., Zilles, K., 2001. Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. Neuroimage 13, 684–701.
Morosan, P., Schleicher, A., Amunts, K., Zilles, K., 2005. Multimodal architectonic mapping of human superior temporal gyrus. Anat. Embryol. 210, 401–406.
Mort, D.J., Perry, R.J., Mannan, S.K., Hodgson, T.L., Anderson, E., Quest, R., McRobbie, D., McBride, A., Husain, M., Kennard, C., 2003. Differential cortical activation during voluntary and reflexive saccades in man. Neuroimage 18, 231–246.
Muller, R.A., Kleinhans, N., Courchesne, E., 2001. Broca's area and the discrimination of frequency transitions: a functional MRI study. Brain Lang. 76, 70–76.
Newberg, A.B., d'Aquili, E.G., 1998. The neuropsychology of spiritual experience. J. Conscious. Stud. 7, 251–266.
O'Neil, E.B., Hutchison, R.M., McLean, D.A., Köhler, S., 2014. Resting-state fMRI reveals functional connectivity between face-selective perirhinal cortex and the fusiform face area related to face inversion. Neuroimage 92, 349–355.
Oh, A., Duerden, E.G., Pang, E.W., 2014. The role of the insula in speech and language processing. Brain Lang. 135, 96–103.
Pajula, J., Kauppi, J.P., Tohka, J., 2012. Inter-subject correlation in fMRI: method validation against stimulus-model based analysis. PLoS One 7, e41196.
Palomero-Gallagher, N., Eickhoff, S.B., Hoffstaedter, F., Schleicher, A., Mohlberg, H., Vogt, B.A., Amunts, K., Zilles, K., 2015. Functional organization of human subgenual cortical areas: relationship between architectonical segregation and connectional heterogeneity. Neuroimage 115, 177–190.

Palomero-Gallagher, N., Vogt, B.A., Schleicher, A., Mayberg, H.S., Zilles, K., 2009. Receptor architecture of human cingulate cortex: evaluation of the four-region neurobiological model. Hum. Brain Mapp. 30, 2336–2355.
Pannese, A., Grandjean, D., Frühholz, S., 2015. Subcortical processing in auditory communication. Hear. Res. 328, 67–77.
Parker, G.J., Luzzi, S., Alexander, D.C., Wheeler-Kingshott, C.A., Ciccarelli, O., Ralph, M.A.L., 2005. Lateralization of ventral and dorsal auditory-language pathways in the human brain. Neuroimage 24, 656–666.
Pierrot-Deseilligny, C., Muri, R.M., Nyffeler, T., Milea, D., 2005. The role of the human dorsolateral prefrontal cortex in ocular motor behavior. Ann. N. Y. Acad. Sci. 1039, 239–251.
Pinto, Y., van der Leij, A.R., Sligte, I.G., Lamme, V.A., Scholte, H.S., 2013. Bottom-up and top-down attention are independent. J. Vis. 13, 16.
Price, C., Veltman, D., Ashburner, J., Josephs, O., Friston, K., 1999. The critical relationship between the timing of stimulus presentation and data acquisition in blocked designs with fMRI. Neuroimage 10, 36–44.
Pruessner, J.C., Dedovic, K., Khalili-Mahani, N., Engert, V., Pruessner, M., Buss, C., Renwick, R., Dagher, A., Meaney, M.J., Lupien, S., 2008. Deactivation of the limbic system during acute psychosocial stress: evidence from positron emission tomography and functional magnetic resonance imaging studies. Biol. Psychiatry 63, 234–240.
Ren, Y., Nguyen, V.T., Guo, L., Guo, C.C., 2017. Inter-subject functional correlation reveal a hierarchical organization of extrinsic and intrinsic systems in the brain. Sci. Rep. 7, 10876.
Rodriguez, S.M., Archila-Suerte, P., Vaughn, K.A., Chiarello, C., Hernandez, A.E., 2018. Anterior insular thickness predicts speech sound learning ability in bilinguals. Neuroimage 165, 278–284.
Royal, I., Vuvan, D.T., Zendel, B.R., Robitaille, N., Schonwiesner, M., Peretz, I., 2016. Activation in the right inferior parietal lobule reflects the representation of musical structure beyond simple pitch discrimination. PLoS One 11, e0155291.
Rykhlevskaia, E., Fabiani, M., Gratton, G., 2006. Lagged covariance structure models for studying functional connectivity in the brain. Neuroimage 30, 1203–1218.
Salomon, R., Levy, D.R., Malach, R., 2014. Deconstructing the default: cortical subdivision of the default mode/intrinsic system during self-related processing. Hum. Brain Mapp. 35, 1491–1502.
Salomon, R., Malach, R., Lamy, D., 2009. Involvement of the intrinsic/default system in movement-related self recognition. PLoS One 4, e7527.
Sauter, D.A., Eisner, F., Ekman, P., Scott, S.K., 2010. Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. Proc. Natl. Acad. Sci. U. S. A. 107, 2408–2412.
Scheperjans, F., Eickhoff, S.B., Homke, L., Mohlberg, H., Hermann, K., Amunts, K., Zilles, K., 2008. Probabilistic maps, morphometry, and variability of cytoarchitectonic areas in the human superior parietal cortex. Cerebr. Cortex 18, 2141–2157.
Scott, S.K., Sauter, D., McGettigan, C., 2010. Chapter 5.5 - brain mechanisms for processing perceived emotional vocalizations in humans. In: Brudzynski, S.M. (Ed.), Handbook of Behavioral Neuroscience. Elsevier, pp. 187–197.
Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: uses in assessing rater reliability. Psychol. Bull. 86, 420–428.
Somandepalli, K., Kelly, C., Reiss, P.T., Zuo, X.-N., Craddock, R.C., Yan, C.-G., Petkova, E., Castellanos, F.X., Milham, M.P., Di Martino, A., 2015. Short-term test–retest reliability of resting state fMRI metrics in children with and without attention-deficit/hyperactivity disorder. Dev. Cogn. Neurosci. 15, 83–93.
Stephens, G.J., Honey, C.J., Hasson, U., 2013. A place for time: the spatiotemporal structure of neural dynamics during natural audition. J. Neurophysiol. 110, 2019–2026.
Termenon, M., Jaillard, A., Delon-Martin, C., Achard, S., 2016. Reliability of graph analysis of resting state fMRI using test-retest dataset from the human connectome project. Neuroimage 142, 172–187.
Thioux, M., Gazzola, V., Keysers, C., 2008. Action understanding: how, what and why. Curr. Biol. 18, R431–R434.
Tierney, A., Dick, F., Deutsch, D., Sereno, M., 2013. Speech versus song: multiple pitch-sensitive areas revealed by a naturally occurring musical illusion. Cerebr. Cortex 23, 249–254.
Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273–289.
Uddin, L.Q., Supekar, K., Amin, H., Rykhlevskaia, E., Nguyen, D.A., Greicius, M.D., Menon, V., 2010. Dissociable connectivity within human angular gyrus and intraparietal sulcus: evidence from functional and structural connectivity. Cerebr. Cortex 20, 2636–2646.
Umarova, R.M., Saur, D., Schnell, S., Kaller, C.P., Vry, M.-S., Glauche, V., Rijntjes, M., Hennig, J., Kiselev, V., Weiller, C., 2009. Structural connectivity for visuospatial attention: significance of ventral pathways. Cerebr. Cortex 20, 121–129.
van Zyl, J.M., Neudecker, H., Nel, D.G., 2000. On the distribution of the maximum likelihood estimator of Cronbach's alpha. Psychometrika 65, 271–280.
Walker, R., Walker, D.G., Husain, M., Kennard, C., 2000. Control of voluntary and reflexive saccades. Exp. Brain Res. 130, 540–544.
Wang, J., Ren, Y., Hu, X., Nguyen, V.T., Guo, L., Han, J., Guo, C.C., 2017. Test-retest reliability of functional connectivity networks during naturalistic fMRI paradigms. Hum. Brain Mapp. 38, 2226–2241.
Wang, K., Liang, M., Wang, L., Tian, L., Zhang, X., Li, K., Jiang, T., 2007. Altered functional connectivity in early Alzheimer's disease: a resting-state fMRI study. Hum. Brain Mapp. 28, 967–978.
Wang, Y., Sereno, J.A., Jongman, A., Hirsch, J., 2003. fMRI evidence for cortical modification during learning of Mandarin lexical tone. J. Cogn. Neurosci. 15, 1019–1027.
Wei, J., Chen, T., Li, C., Liu, G., Qiu, J., Wei, D., 2018. Eyes-open and eyes-closed resting states with opposite brain activity in sensorimotor and occipital regions: multidimensional evidences from machine learning perspective. Front. Hum. Neurosci. 12.
Worsley, K.J., Liao, C.H., Aston, J., Petre, V., Duncan, G.H., Morales, F., Evans, A.C., 2002. A general statistical analysis for fMRI data. Neuroimage 15, 1–15.
Xia, M., Wang, J., He, Y., 2013. BrainNet Viewer: a network visualization tool for human brain connectomics. PLoS One 8, e68910.
Xu, Y., Lin, Q., Han, Z., He, Y., Bi, Y., 2016. Intrinsic functional network architecture of human semantic processing: modules and hubs. Neuroimage 132, 542–555.
Yang, J., Gates, K.M., Molenaar, P., Li, P., 2015. Neural changes underlying successful second language word learning: an fMRI study. J. Neurolinguistics 33, 29–49.
Zaborszky, L., Hoemke, L., Mohlberg, H., Schleicher, A., Amunts, K., Zilles, K., 2008. Stereotaxic probabilistic maps of the magnocellular cell groups in human basal forebrain. Neuroimage 42, 1127–1141.
Zhang, F., Wang, J.P., Kim, J., Parrish, T., Wong, P.C., 2015. Decoding multiple sound categories in the human temporal cortex using high resolution fMRI. PLoS One 10, e0117303.
Zou, Q., Miao, X., Liu, D., Wang, D.J., Zhuo, Y., Gao, J.-H., 2015. Reliability comparison of spontaneous brain activities between BOLD and CBF contrasts in eyes-open and eyes-closed resting states. Neuroimage 121, 91–105.