OXFORD

Investigation of functional brain network reconfiguration during vocal emotional processing using graph-theoretical analysis

Shih-Yen Lin,[1,2] Chi-Chun Lee,[3] Yong-Sheng Chen,[2] and Li-Wei Kuo[1,4]

[1]Institute of Biomedical Engineering and Nanomedicine, National Health Research Institutes, 35053 Miaoli, Taiwan, [2]Department of Computer Science, National Chiao Tung University, 30010 Hsinchu, Taiwan, [3]Department of Electrical Engineering, National Tsing Hua University, 30013 Hsinchu, Taiwan, and [4]Institute of Medical Device and Imaging, National Taiwan University College of Medicine, 10051 Taipei, Taiwan

Correspondence should be addressed to Li-Wei Kuo, Institute of Biomedical Engineering and Nanomedicine, National Health Research Institutes, 35 Keyan Road, Zhunan Town, Miaoli County, 35053, Taiwan. Email: lwkuo@nhri.org.tw.

Abstract

Vocal expression is essential for conveying the emotion during social interaction. Although vocal emotion has been explored in previous studies, little is known about how perception of different vocal emotional expressions modulates the functional brain network topology. In this study, we aimed to investigate the functional brain networks under different attributes of vocal emotion by graph-theoretical network analysis. Functional magnetic resonance imaging (fMRI) experiments were performed on 36 healthy participants. We utilized the Power-264 functional brain atlas to calculate the interregional functional connectivity (FC) from fMRI data under resting state and vocal stimuli at different arousal and valence levels. The orthogonal minimal spanning trees method was used for topological filtering. The paired-sample *t*-test with Bonferroni correction across all regions and arousal–valence levels were used for statistical comparisons. Our results show that brain network exhibits significantly altered network attributes at FC, nodal and global levels, especially under high-arousal or negative-valence vocal emotional stimuli. The alterations within/between well-known large-scale functional networks were also investigated. Through the present study, we have gained more insights into how comprehending emotional speech modulates brain networks. These findings may shed light on how the human brain processes emotional speech and how it distinguishes different emotional conditions.

Key words: functional magnetic resonance imaging; vocal emotion; resting state; brain network; graph-theoretical analysis

Introduction

Emotion is one of the crucial cognitive factors that affect our daily life and social interaction. Various facial and vocal expressions convey the emotion during social interaction. Thus, comprehending these emotional expressions and their underlying neural mechanism is essential to modern society and to build new communication technologies. Several prior neuroimaging studies have aimed to elucidate the neural mechanism for emotional processing; however, most of them have studied emotion based on facial expressions and visual stimuli (Lane *et al.,* 1997; Phan *et al.,* 2002). Of late, several neuroimaging studies have focused on vocal emotion. Functional magnetic resonance imaging (fMRI) studies have shown that emotional prosody (especially emotions such as anger) consistently activates amygdala as well as numerous brain regions in the lateral temporal lobe and frontal lobe (Mitchell *et al.,* 2003; Grandjean *et al.,* 2005; Sander *et al.,*

2005; Liebenthal *et al.*, 2016). Additionally, electrophysiological [electroencephalography (EEG) and magnetoencephalography (MEG)] studies using event-related potentials have strived to delineate the neural dynamics related to the effects of vocal emotions. For example, Paulmann *et al.* (2013) suggested that valence information is decoded during early processing, while arousal effects occur at a later stage of processing.

In addition to changes in regional brain activity, emotional perception may also alter the interregional functional connectivity (FC) as well as the brain network topology. At the connectivity level, studies have employed FC analysis and investigated reorganized FC induced by emotional processing (Anticevic *et al.*, 2011; Kim *et al.*, 2011; Klapwijk *et al.*, 2013; Eckstrand *et al.*, 2018; Ewbank *et al.*, 2018). At the network level, the recent advancement in computational approaches, especially graph-theoretical analysis, has provided the means to characterize the brain network topology (Rubinov and Sporns, 2010). Several studies have used graph-theoretical analysis to investigate the alteration of brain networks when interpreting facial emotional expressions that have shown significant changes of global efficiency and clustering coefficient (CC) compared with the resting state (Di and Biswal, 2018; Zuo *et al.*, 2018). Compared with regional functional activation, connectivity-based and network-based studies may help us gain more insights into the neural mechanism of emotion and learn how emotion may modulate the cognition states. Currently, a growing body of evidence supports the affective workspace hypothesis, suggesting that either positive or negative affective state is not necessarily associated with activating a specific set of regions (Barrett and Bliss-Moreau, 2009; Lindquist *et al.*, 2015). Alternatively, it can emerge as 'brain state' at the population level. Therefore, to strengthen the validity of this hypothesis, it is essential to understand how emotion-related cortical regions interact during processing of emotional information.

Several studies have explored the alteration of network topology due to facial emotional expressions, but not much is known about the effects of vocal emotional stimuli on brain networks at connectivity and network topological levels. Therefore, to explore the vocal emotion and its underlying neural mechanism, this study has the following three objectives. First, we sought to examine the feasibility of graph-theoretical analysis on fMRI data with vocal emotional stimuli. Second, we sought to explore whether vocal emotional stimuli induce any alteration of brain networks—at both network topological and connectivity levels. We also investigated the differences within/between well-known large-scale functional networks. Third, we sought to investigate whether there is any difference in network topology between the resting state and the ones induced by vocal emotional stimuli.

Materials and methods

Participants

A total of 36 healthy volunteers (27 male and 9 female) participated in our study. Furthermore, to reduce the risk of possible confounding factors, the participants were recruited based on several criteria: being free of any brain disease or major brain injury, age ranging between 20 and 35 years and a college or higher-level education to understand the vocal emotion stimuli pronounced in English. Furthermore, we only recruited right-handed subjects to exclude any potential variability due to handedness. The Institutional Review Board at

National Health Research Institutes approved this study, and all volunteers provided informed consent.

Experimental stimuli

The vocal emotion stimuli were generated from part of the USC IEMOCAP database (Busso *et al.*, 2008). The audio data from IEMOCAP database consist of recordings of scripted or spontaneous speech during dyadic interaction between a pair of voice actors. Naive raters rated each recording with attributes including valence, arousal and dominance with the continuous rank from 1 to 5. For our study, we used the scripted dialogs by a chosen male voice actor whose recordings yielded the highest variability in the speech attributes among all voice actors. From 639 segments spoken by the selected voice actor, we selected 251 voice segments as the stimuli for our experiments.

We categorized the stimuli into two types of emotional attributes, namely, arousal and valence, and designed three conditions for each feature. Each experiment comprised six 5 min vocal emotion stimuli and a 1 min break between any two stimuli. For the arousal attribute, the conditions were categorized into low (value ≤ 2.5), medium ($2.5 <$ value < 3.5) and high ($3.5 \leq$ value) levels. For the valence attribute, the conditions were negative (value ≤ 2.5), neutral ($2.5 <$ value < 3.5) and positive ($3.5 \leq$ value) levels. For each condition, the speech segments with the given attribute and level were shuffled to remove any contextual information and were then concatenated to form a 5 min continuous vocal emotional stimulus. The participants were asked to pay attention to the speech-based stimuli without being informed of the purpose or the details of the experiment.

Image acquisition

MR experiments were performed on a 3T MRI scanner (Prisma, Siemens, Erlangen, Germany) at National Taiwan University. Each scanning session included T1-weighted imaging (T1WI), resting-state fMRI (rs-fMRI) and task-evoked fMRI (t-fMRI) of all vocal emotional stimuli. The T1WI protocol was employed using a magnetization-prepared rapid gradient-echo sequence with repetition time (TR) of 2000 ms, echo time (TE) of 2.3 ms, inversion time (TI) of 900 ms, flip angle (α) of 8°, voxel size of $1 \times 1 \times 1$ mm^3, matrix size of 256×256 and 192 slices. Each fMRI scan with blood oxygen level-dependent (BOLD) contrast was acquired using gradient-echo echo-planar imaging sequence with TR/TE of 3000/32 ms, α of 90°, voxel size of $2.5 \times 2.5 \times 3$ mm^3, matrix size of 96×96, 40 slices and 100 repetitions.

Data pre-processing

Before network analyses, all rs-fMRI and t-fMRI data sets were pre-processed using DPARSF toolbox (Chao-Gan and Yu-Feng, 2010). The pre-processing procedures included the removal of the first 10 volumes, slice-timing correction, co-registration to T1WI, covariate regression of head motion, white matter signals and cerebrospinal fluid signals, nonlinear spatial normalization using T1WI, linear detrending and band-pass filtering (0.01–0.1 Hz). To estimate the FC over the whole brain, brain regions were parcellated using the Power-264 functional atlas (Power *et al.*, 2011), which comprises 264 putative functional regions-of-interest (ROIs) associated to 13 large-scale functional networks and a group of unlabeled regions (Table 1). We also provide the region definitions used in Automated Anatomical Labeling (AAL) atlas (Supplementary Table S2) and reported

Table 1. Abbreviations of the large-scale functional networks defined in Power-264 functional atlas

Abbreviations	Functional networks
SM.H	Sensory/somatomotor hand
SM.M	Sensory/somatomotor mouth
CO	Cingulo-opercular task control
Aud	Auditory
DMN	Default mode
MR	Memory retrieval
Vis	Visual
FP	Fronto-parietal task control
Sal	Salience
Sub	Subcortical
VA	Ventral attention
DA	Dorsal attention
CB	Cerebellum
Unlabeled	All unlabeled regions

the corresponding anatomical locations of functional ROIs using the definitions by the AAL atlas. The averaged time series of each putative functional ROI defined in Power-264 functional atlas was derived by averaging the pre-processed rs-fMRI signals within the ROI. The pairwise between ROI FC was derived by quantifying the temporal dependency between two extracted averaged time series. We computed two types of FC measures—Pearson's correlation (PC) and covariance (COV). It should be noted that the negative FCs were excluded in the following analysis, i.e. only positive FCs were used. Subsequently, we employed the orthogonal minimal spanning trees (OMSTs) method on constructed FC matrices to filter out spurious connections (Dimitriadis *et al.,* 2017a; Dimitriadis *et al.,* 2017b). Briefly, the OMSTs iteratively extract the minimal spanning trees from a given graph, and the filtered graph is the aggregate of OMSTs that maximizes the global efficiency subtracted by the wiring cost of the brain network. Compared with the conventional sparsity thresholding method based on either a given FC value or a network sparsity, the OMSTs method is parameter-free and more reproducible in group-wise or even individual-level brain network (Dimitriadis *et al.,* 2017a; Dimitriadis *et al.,* 2017b).

Graph-theoretical analysis

After applying OMSTs, the graph-theoretical analysis was employed to derive both nodal and global graph-theoretical network measures from the filtered FC matrices. The nodal network measures used in this study are degree centrality (DC), CC (Saramäki *et al.,* 2007), local efficiency (E_{loc}) (Latora and Marchiori, 2001) and PageRank centrality (PR) (Rubinov and Sporns, 2010). In addition to investigating the network attributes at nodal scale, we also examined the network attributes at the global scale—the whole brain—using a set of global graph-theoretical network measures. The global network measures in our study includes characteristic path length (L), global efficiency (E_{glob}), mean local efficiency ($\overline{E}_{loc}$), mean clustering coefficient ($\overline{CC}$), transitivity (T) (Newman, 2003), modularity (M) and assortativity coefficient (AC) (Humphries and Gurney, 2008), in addition to the network wiring cost (S). We provided the detailed definitions of network measures in the Supplementary Section S1. One can also refer to a previous review article for more details (Rubinov and Sporns, 2010). We also performed an analysis of complementarity among different network measures and provided the discussion in the Supplementary Section S2.

Statistical analysis

In this study, we sought to explore the topological reconfiguration of t-fMRI networks with vocal emotional stimuli and rs-fMRI networks. By categorizing these vocal emotional stimuli into multiple arousal and valence levels, we further investigated the relationship within and between these levels, as well as their differences with the resting-state condition. The comparisons were performed at nodal network, global network and FC levels. As for each well-known large-scale functional network, we calculated the averaged nodal measures across its member ROIs, the averaged intra-network FC and inter-network FC connecting to other functional networks. Except the above analyses, we also performed the analysis of common connections across all subjects for each specific arousal, valence or resting-state condition and discussed in the Supplementary Section S3. All statistical analyses of the FC and graph-theoretical network measures were performed using the paired-sample t-test. All significant levels were subsequently adjusted for multiple comparisons jointly across 264 ROIs and 6 pairs of conditions (either resting-state and 3 arousal levels or resting-state and 3 valence levels) using Bonferroni correction.

Results

Investigation on nodal network measures

Table 2 shows the statistical comparisons of the nodal network measures among different t-fMRI and resting-state conditions. Note that we denoted the type of FC in superscripts for a given network metric in the following sections. For example, $DC^{(PC)}$ denotes the DC calculated using PC as definition of FC. For t-fMRI with arousal stimuli, significantly reduced $PR^{(PC)}$ of low-arousal condition was found in an ROI (in STG.R) within the auditory network by comparing with resting-state condition. However, no significant differences among those arousal and resting-state conditions were found by using all nodal network measures derived from COV. For t-fMRI with valence stimuli, significant differences of nodal network measures were only found between neutral- and negative-valence conditions. Compared to neutral-valence condition, our results show decreased $CC^{(PC)}$ in one ROI (located between MOG.L and IOG.L) within the visual network, $CC^{(COV)}$ in one ROI (between DCG and SMA.L) within the hand sensory/somatomotor network and $E_{loc}^{(COV)}$ in three ROIs (one in IOG.R, one between ITG.R and IOG.R and one between CUN.R an PCUN.R) in visual network of negative-valence condition.

Investigation on global network measures

Table 3 shows the statistical comparisons of the global network measures among different t-fMRI and resting-state conditions. For t-fMRI with arousal stimuli, the high-arousal condition showed increased $E_{glob}^{(COV)}$, $\overline{E}_{loc}^{(COV)}$, $T^{(COV)}$ and $\overline{CC}^{(COV)}$ compared to mid-arousal condition. The decreased $L^{(COV)}$ was also found in high-arousal condition compared to mid-arousal condition. Compared to low-arousal condition, increased $\overline{CC}^{(COV)}$ was found in high-arousal condition. However, no significant between-condition difference of global network measures was found by utilizing *PC* as definition of FC. For t-fMRI with valence stimuli, significant between-group differences

Table 2. The statistical comparisons of nodal network measures among different task-evoked, (a) arousal stimuli and (b) valence stimuli and resting-state conditions. All P-values were corrected for multiple comparison across the arousal–valence levels by Bonferroni correction (*$P < 0.01$; **$P < 0.001$). For each ROI, both of the corresponding network in Power-264 atlas and the corresponding regions in Automated Anatomical Labeling atlas were shown. Please see Supplementary Table S1 for abbreviations of AAL region

(a) Arousal stimuli

No. of ROI in Power-264 atlas	Network (AAL region)	PC Rest > Low
63	Auditory (STG.R)	PR

(b) Valence stimuli

#ROI in Power-264 atlas	Network (AAL regions)	PC Neutral > negative	COV Neutral > negative
168	Visual (MOG.L, IOG.L)	CC, E_{loc}	
15	Sensory/somatomotor hand (DCG.R, DCG.L, SMA.L)		CC, E_{loc}
153	Visual (IOG.R)		E_{loc}
161	Visual (ITG.R, IOG.R)		E_{loc}
163	Visual (CUN.R, PCUN.R)		E_{loc}

Table 3. The statistical comparisons of global network measures among different task-evoked, (a) arousal stimuli and (b) valence stimuli and resting-state conditions. All P-values were corrected for multiple comparison across the arousal–valence levels by Bonferroni correction (*$P < 0.01$; **$P < 0.001$)

(a) Arousal stimuli

	COV Low < high	Mid < high
E_{glob}		*
$\overline{E}_{loc}$	*	**
$\overline{CC}$		*
T		*

	COV Mid > high
L	*

(b) Valence stimuli

	PC Neutral > negative	COV Rest > negative	Neutral > negative	Positive > negative
E_{glob}	*	**	***	*
$\overline{E}_{loc}$	*	**	***	
$\overline{CC}$		*	**	
T	*	*	**	

	PC Neutral < negative	COV Rest < negative	Neutral < negative	Positive < negative
L	*	*	***	*

of global network measures were found mainly in negative-valence condition compared to other valence or resting-state conditions. For utilizing COV as FC definition, the altered global network measures includes increased $L^{(COV)}$ and decreased $E_{glob}^{(COV)}$ (compared to all other conditions), decreased $\overline{E}_{loc}^{(COV)}$, $\overline{CC}^{(COV)}$ and $T^{(COV)}$ (compared to resting-state and neutral-valence conditions). For utilizing PC as FC definition, decreased $E_{glob}^{(PC)}$, $\overline{E}_{loc}^{(PC)}$, $T^{(PC)}$ and increased $L^{(PC)}$ (compared to neutral-valence condition) were found. No significant between-condition difference of M, AC nor S was found among all comparisons.

Investigation on interregional FC

In addition to network metrics—either nodal or global—we also performed the between-condition comparisons of interregional FC (PC based and COV based). For arousal stimuli, significant between-condition differences were found by using PC-based FC, while no significant between-condition differences were found by using COV-based FC. Compared with resting-state condition, significantly reduced PC-based FC were found in either low- or high-arousal condition for Aud, SM.M, SM.H, DA, VA and Vis networks, as shown in Figure 1. For valence stimuli, significant between-condition differences were found by using both PC-based and COV-based FC. By comparing the resting-state and positive-valence conditions, the significantly different connection with PC-based FC was found between DA and DMN. Most of the significant between-condition differences were found to associated with reduced FC in negative-valence condition, including Vis-FP (resting-state > negative-

valence; neutral-valence > negative-valence), Vis-DMN (resting-state > negative-valence), SM.H-Vis (neutral-valence > negative-valence), SM.M-CO (neutral-valence > negative-valence) and DA-Vis (neutral-valence > negative-valence). Two significantly different connections with increased FC in negative-valence condition were found in intra-DMN (resting-state < negative-valence) and SM.H-CO (positive-valence < negative-valence). For COV-based FC, a significantly different connection was found between DMN and Vis (negative-valence < neutral valence; Figure 2).

Investigation on large-scale functional networks

Table 4 shows the statistical comparisons of averaged nodal network measures within well-known large-scale functional networks among different conditions. No significant difference was found in the arousal condition. In contrast, significant differences of averaged nodal network measures were found to be mostly associated with negative-valence condition. Compared with the neutral-valence condition, decreased averaged nodal network measures were found in negative-valence condition, including $CC^{(COV)}$ and $E_{loc}^{(COV)}$ in SM.H, $E_{loc}^{(COV)}$ in DMN and DC, CC and E_{loc} in Vis for both FC definitions. Additionally, increased $PR^{(COV)}$ in MR and decreased $E_{loc}^{(COV)}$ in Vis were found in negative-valence condition by comparing with the neutral-valence and resting-state conditions, respectively. Figure 3 shows the inter-network and intra-network comparisons of FCs. The alterations of FCs were only found in t-fMRI with
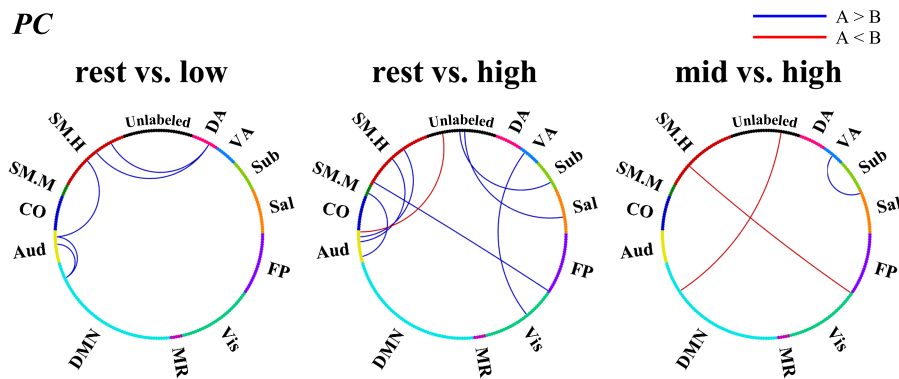
Fig. 1. Significant changes of FC associated with arousal stimuli and resting state. Blue and red lines signify decrease and increase of connectivity in the latter condition compared with the former condition, respectively. For having a better visualization, the ROIs are reordered and colored according to their correspondence to the large-scale functional networks. Please see Table 1 for the abbreviations of the functional networks. Note that we excluded CB in the illustration. All FCs are corrected for multiple comparisons across arousal levels using Bonferroni correction.
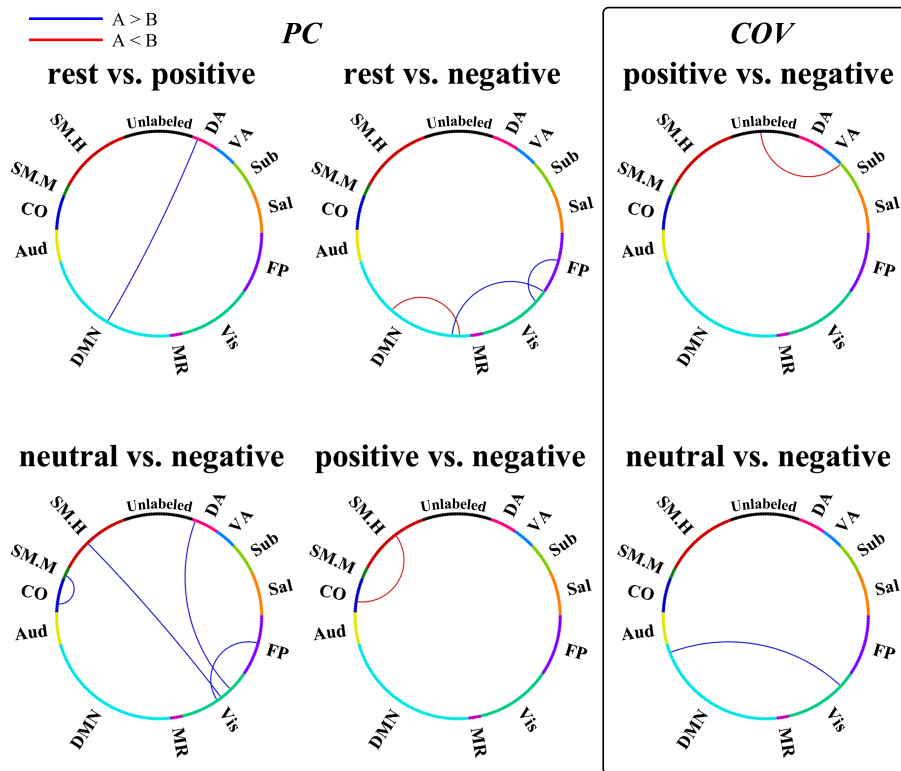

Fig. 2. Significant changes of FC associated with valence stimuli and resting state. Blue and red lines signify decrease and increase of connectivity in the latter condition compared with the former condition, respectively. For having a better visualization, the ROIs are reordered and colored according to their correspondence to the large-scale functional networks. Please see Table 1 for the abbreviations of the functional networks. Note that we excluded CB in the illustration. All FCs are corrected for multiple comparisons across valence levels using Bonferroni correction.

valence stimuli, and most of them were associated with the negative-valence condition. A total of five inter-network alternations of PC-based FC were found, including VA-Sub (resting-state < positive-valence), CO-DA (positive-valence < negative-valence), Sub-FP (positive-valence < negative-valence), Aud-SM.M (neutral-valence > negative-valence) and Vis-DA (neutral-valence > negative-valence). In contrast, the only intra-network alteration was found in Vis (neutral-valence > negative-valence) by using COV-based FC.

Discussion

Our study demonstrates that perception of emotional speech could modulate brain network topology in several cortical regions associated with emotion processing. We also found the altered global network topology among different task-based and resting-state conditions. Beyond regional level, we further investigated the alterations of network metrics and FCs within/between large-scale functional networks and reported
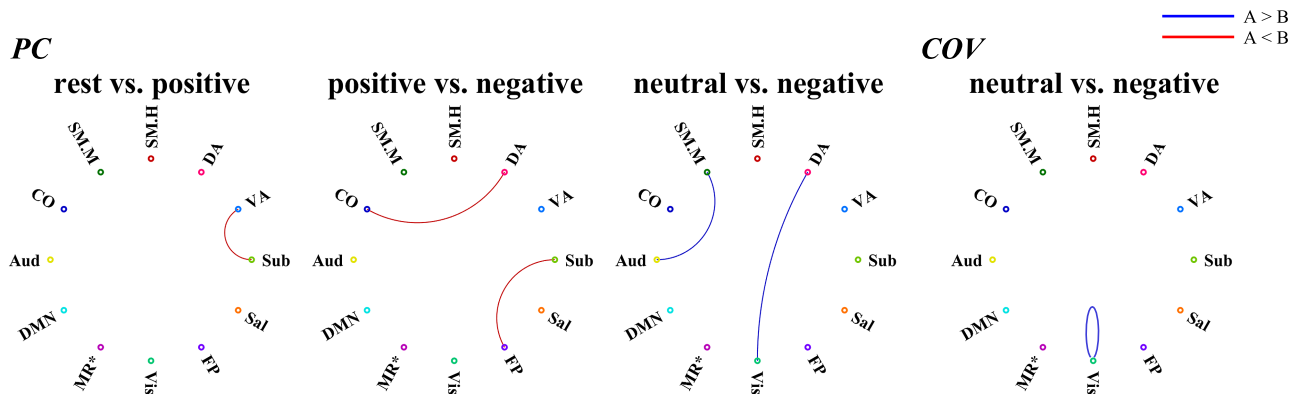
Fig. 3. Significant changes of averaged intra-network or inter-network FCs with respect to the 12 large-scale functional networks (without CB and unlabeled) associated with the valence stimuli and resting state. Both PC- and COV-based FCs were investigated. Blue and red lines signify decrease and increase of connectivity in the latter condition compared with the former condition, respectively. For having a better visualization, the ROIs are reordered and colored according to their correspondence to the large-scale functional networks. Please see Table 1 for the abbreviations of the functional networks. All FCs are corrected for multiple comparisons across valence levels using Bonferroni correction.

Table 4. The statistical comparisons of averaged nodal network measures within large-scale functional networks among different task-evoked and resting-state conditions. All P-values were corrected for multiple comparison across the arousal–valence levels by Bonferroni correction (*$P < 0.01$; **$P < 0.001$).

	(a) Arousal experiment (None) (b) Valence experiment			
	PC		COV	
Network	Neutral > negative	Rest > negative	Neutral > negative	Neutral < negative
Sensory/somato-motor mouth			CC, E_{loc}	
Default mode network			E_{loc}	
Memory retrieval				PR
Visual	DC, CC, E_{loc}	E_{loc}	DC, CC, E_{loc}	

our findings. To our knowledge, this is the first study that investigates the effects of vocal emotional stimuli on brain network topology using graph-theoretical analysis of fMRI data. These findings may shed light on how the human brain processes emotional speech and how it distinguishes different emotions. In the following sections, the results from our analysis and their interpretations are elaborated. Also, the limitations of the experimental design and data interpretation are discussed.

Task-related alterations in nodal network measures

Results for valence stimuli revealed a tendency that the network topology was significantly altered under the negative-valence condition compared with that of neutral valence or resting state, suggesting that the negative-valence stimuli may modulate or reorganize the brain network. In addition, we should note that alterations of averaged network measures in the large-scale functional networks are highly consistent with that by investigating individual ROIs, further supporting our findings. One interesting finding from the experiments with valence stimuli was that the reductions of functional segregation ($CC^{(COV)}$ and $E_{loc}^{(COV)}$) were observed in the visual network. These alterations

were observed in several individual ROIs in visual network and from investigating the averaged measures in visual network. A meta-analytic review by Kober et al. (2008) has reported that a group of visual sub-regions would be activated under visual emotional stimuli. Furthermore, we hypothesized that these visual sub-regions could be stimulated by not only visual stimuli but also other modalities. A similar hypothesis has been introduced in an fMRI study by Sander et al. (2005) in which activation in CUN was observed under attended anger prosody compared with neutral or unattended anger prosody. Therefore, we speculated that the alteration of network topology may be resulted from the complex cross-modal interactions during emotional processing. One possible explanation about cross-modal interactions in our case is the visual mental imagery triggered by the speech stimuli. A previous fMRI study showed that the mental imagery evokes greater emotional response than verbal representation (Holmes and Mathews, 2005). Another fMRI study by Just et al. (2004) also revealed that the visual imagery is crucial for sentence comprehension. Essentially, the theory of multi-modal mental imagery has been supported by a growing body of evidence. For instance, an fMRI study by Pekkola et al. (2005) showed that using silent visual speech stimulus (facial videos during speech overlaid with written pronunciation) could activate primary auditory cortex. Other than visual mental imagery, a few studies have also reported different kinds of cross-modal interactions during emotional processing. Brosch et al. (2008, 2009) have reported that visual attention could be modulated by anger prosody. Another EEG study by Jessen et al. (2013) also showed that the cross-modal prediction of emotion exists in the multimodal processing of audiovisual emotion. Based on these previous studies, we could suggest that a similar cross-modal interaction mechanism to alter the network topology might also be revealed in visual sub-regions. However, a more sophisticated experimental design in further study would be needed to verify our speculation.

We also observed significantly reduced nodal functional segregation ($CC^{(COV)}$ and $E_{loc}^{(COV)}$) in the sensorimotor network by comparing negative-valence and neutral-valence conditions. These alterations were found in one ROI in the hand sensorimotor network and by investigating the averaged network measure of the mouth sensorimotor network. Consistently, previous studies have also reported the association of sensorimotor network with speech, language and emotional processing

(Oliveri *et al.,* 2003; Nummenmaa *et al.,* 2014b; Hertrich *et al.,* 2016). A study using transcranial magnetic stimulation suggested the role of supplementary motor area in movement control triggered by emotional stimuli (Oliveri *et al.,* 2003). An fMRI study showed that the vocal emotion was associated with the BOLD responses in emotion, attention and sensorimotor circuits, in addition to the inter-subject synchronization within somatosensory and supplementary motor cortices (Nummenmaa *et al.,* 2014b). Intense emotion can trigger corresponding physiological and bodily response through sensorimotor and visceral nervous systems (Vrana and Lang, 1990; Costa *et al.,* 2010; Nummenmaa *et al.,* 2014a). Therefore, it is reasonable to speculate that the alterations in sensorimotor network were likely due to the increased demand for physiological and bodily emotion response.

Task-related alterations in global network measures

Our results showed that vocal emotional stimuli altered not only nodal network measures but also global network measures. For arousal stimuli, significant increases of functional integration (increased E_{glob} and decreased L) and segregation (increased $\bar{E}_{loc}$, CC, and T) were found in high-arousal condition compared with low- and mid-arousal conditions. For valence stimuli, significantly reduced functional integration and segregation were found in negative-valence condition compared with all the other conditions (neutral valence, positive valence and resting state). We hypothesized that the brain network for processing emotional speech with high-arousal condition might intrinsically exhibit distinct level of functional integration and segregation as compared with other conditions or resting state. In this case, the brain network under high-arousal condition may show higher degree of integration and segregation, while the task-negative resting-state network is being suppressed. However, the brain network under low- or mid-arousal condition may be presented as a mixed pattern of task-positive and resting-state networks. Having different combinations of task-positive and resting-state networks may contribute to our speculation about the altered global network topology between high-arousal and the other two arousal levels.

Similarly, the brain network to process the negative-valence vocal emotion stimuli may be characterized by reduction in network integration and segregation. Our results generally showed reduced network integration and segregation in negative-valence conditions compared with the resting state. Previous studies have attempted to understand the underlying mechanism and investigate the relationship between task-specific and resting-state networks further. Di *et al.* (2013) compared the global network measures of a task-general and meta-analytic coactivation network to a group-averaged resting-state network and reported reduced clustering, reduced modularity and increased efficiency. Recently, Zuo *et al.* (2018) used binarized PC matrix for studying the change of brain network topology under seven different kinds of functional tasks, which showed significant increases in global efficiency in all functional tasks compared with resting state. Another specific study by Wang *et al.* (2012) investigated the network topology during the semantic matching task and resting state using binarized correlation matrices. Their results showed reduced global efficiency, reduced normalized global efficiency, increased E_{loc} and increased nodal centrality. Taya *et al.* (2014) used alphabet recognition tasks and discovered reduced normalized CC compared with that of resting state. Although

the experimental designs and targeted network measures of these previous studies do not converge in details, a general tendency that we could summarize from these studies is that most task-related networks would exhibit increased efficiency—in contrast to our findings. This controversy may arise from experimental designs, pre-processing of graph theoretical analysis, computation of network measures and statistical comparison approaches. A further study is needed to clarify these effects of data processing.

Task-related alterations in FC

Our results also showed altered interregional FC in several connections. For arousal stimuli, reduced FC was mostly found in those connections associated with the auditory network, mostly involving superior temporal gyrus (STG). This finding may suggest that the reduced FC centered to these regions could be a result of configuration switching between resting-state and task-positive networks. Several sub-regions within STG, e.g. primary auditory cortex and Wernicke's area, are known to be responsible for processing auditory and language information. Functionally, STG is responsible for language processing, which may also contribute to the altered FC under task-related conditions (Damasio and Geschwind, 1984; Poeppel *et al.,* 2008). In our current results, the altered FC related to STG may reflect that the patterns of network topology are different between vocal emotion modulation and resting state. However, our results cannot fully explain the association between STG and vocal emotional processing. Other than STG, we also observed that the CC and E_{loc} in TPOmid.R under high-arousal condition were significantly lower than those under mid-arousal condition. Interestingly, previous studies have suggested the temporal pole was associated with the social and emotional processing, including face recognition and theory of mind (Beauregard *et al.,* 2001; Olson *et al.,* 2007). Although the association between the temporal pole and vocal emotional processing is not clear yet, we could speculate that the arousal levels of vocal emotion stimuli may alter the network topology and result in altered nodal network characteristics in temporal pole. For valence stimuli, significantly reduced COV-based interregional FC and mean intra-network FC was found in several connections within the visual network by comparing negative-valence to neutral-valence conditions—consistent with our findings in the nodal network topology, further supporting our hypothesis of cross-modal mental imagery altering the network topology in visual-associated regions. We should note that the alterations of averaged network measures in the well-known large-scale functional networks and mean inter-network/intra-network FCs were highly consistent with those observed by investigating individual ROIs, especially in the visual and sensorimotor networks. The observations among large-scale functional networks further solidified our findings in FCs, nodal and global network metrics.

Investigation on complementarity of network measures

Since we incorporated a series of nodal and global network metrics that may be used to quantify similar network topological characteristics in theory, it was of great interest to explore the complementarity between these network metrics. How they complement each other to form a more concrete delineation of the overall brain network topology would be beneficial for our current study. Thus, we also analyzed the similarity between different nodal network metrics using correlation analysis and also

compared these between-metric similarities from two different pre-processing procedures, i.e. OMSTs and sparsity thresholding (see Supplementary Section S2 for details). Our results generally showed that the network metrics used to characterize the same topological attribute could be highly correlated even if they have different theoretical definitions. For example, E_{loc} was calculated based on the shortest path length and CC was based on triangles; however, these two metrics were highly correlated in our case. Although these metrics were highly correlated, E_{loc} revealed more between-condition differences than other metrics that also measured the functional segregation in this study. We could summarize that, in our study, the network metrics for characterizing the same topological attribute could still provide complementary information, e.g. sensitivity to differentiate subtle alterations between conditions, even if they were highly similar in their quantities. It would be beneficial if all metrics were calculated and included in providing more insights into the complex brain network architectures.

A comparison between PC- and COV-based FC

We also investigated the influence of two kinds of FCs—PC and COV—on brain network topology. Interestingly, it was revealed that the analyses using these two FCs could provide non-redundant information for depicting the brain networks under different task-related and resting-state conditions. For nodal network measures, PC and COV reflected the influences of arousal and valence stimuli on brain network topology, respectively. For global network measures, the influences of both arousal and valence stimuli were only revealed by COV. For FC, PC revealed most of the alterations induced by arousal stimuli, whereas COV revealed most of the alterations induced by valence stimuli. By definition, assuming two independent variables X and Y, PC(X,Y) is equivalent to COV(X,Y) divided by the products of the variances of those two variables ($\delta_X \delta_Y$). In other words, the calculation of COV considers both the signal amplitudes and variations, while PC is a dimensionless measure that decouples the effect of the signal variations. Therefore, PC and COV could reflect different aspects of functional dependency in principle and then result in non-redundant observations.

Here, we give two scenarios where the observations of PC and COV may not converge. In some cases where the alteration of signal variances is irrelevant to the stimulation, COV may show a lower significance level than PC due to the inclusion of signal variances. In other cases where the alteration of signal variances is highly relevant to the stimulation, COV may show a higher significance level than PC. Considering the nature of definitions, we speculate that the network modulation under arousal stimuli is less relevant to the alteration of signal variations, while network modulation under valence stimuli is mostly contributed by the alteration of the amplitude of signal fluctuations. To date, choosing optimal FC measures for graph-theoretical analysis remains challenging. Our study demonstrates that the use of multiple FC measures may be a better approach to address the complex network and could provide complementary perspectives on the task-related reconfiguration of a network.

Limitations

In this investigative study, we have shown that the different levels of emotional speech stimuli may alter or modulate the brain networks on either a nodal or global scale. Although we

suggest that brain network analysis could have the potential to resolve the vocal-emotion-induced topological changes, several limitations must be carefully discussed. The first limitation may come from the cultural difference between the volunteers who rated the emotional scales in the IEMOCAP data set and the participants involved in this study. The cultural difference arising from native languages and environmental factors may play a major role in comprehending the emotional speech, which might be the major confounding factor in this exploratory study. The second limitation is the design of vocal emotional stimuli, in which we tried to mimic the real-world scenarios. However, this experimental design might be too complicated to rule out some other mental confounding factors. Considering both limitations, one should use the speech database with the same native language as that of the participants involved in the experiments to investigate better the effects of vocal emotional stimuli on brain network topology. Furthermore, the scenarios of the functional stimuli should be divided into several simplified sections so one could investigate each phenomenon separately. Additionally, it is also worth noting that the method used in this study assumes a static topology under a given type of stimuli. However, it is highly likely that such an assumption does not hold—for brains are dynamic systems. Notably, several studies have also investigated how functional networks change and evolve with time using dynamic FC (Glerean *et al.*, 2012; Hutchison *et al.*, 2013). Nummenmaa *et al.* (2014b) also performed whole-brain dynamic connectivity analysis for studying the effects of emotional speech on dynamic changes of brain networks. It is highly likely that considering the dynamic nature of the brain network would provide a more valid analysis and allow for studying dynamic changes in brain states. However, some careful analysis design is required to apply high-level network analysis to a dynamic network. Furthermore, the emotional stimuli used in our study were attributed to a simple two-dimensional model (i.e. arousal and valence). However, it is also possible to extract emotion-related features directly from the stimuli (El Ayadi *et al.*, 2011). In fact, it has been shown that emotion recognition using EEG signals can be facilitated by incorporating features extracted from the stimuli (Zhu *et al.*, 2014; Gao and Wang, 2015; Wang *et al.*, 2015). Therefore, we postulated that by incorporating sound features extracted from the speech stimuli, we could achieve a more comprehensive analysis of various aspects of emotions during the speech.

Conclusions

In this study, we investigated the modulation of brain networks under emotional speech perception using high-level graph-theoretical network measures. With the use of OMSTs approach and Power-264 functional atlas, we discovered that brain network exhibits significantly altered network attributes at global, nodal and connectivity levels, especially under emotional speech with high arousal or negative valence. We also investigated the alterations of network metrics and FCs within/between large-scale functional networks and found that most of alterations were associated with negative valence. To the best of our knowledge, this is the first study employing a graph-theoretical analysis of emotional speech perception. Although this is predominantly an investigative study, we have gained crucial insights into how comprehending emotional speech modulates brain networks. Additionally, this study provides directions for high-level network analysis on emotional speech comprehension or possibly other types of brain functions.

Supplementary data

Supplementary data are available at *SCAN* online.

References

Anticevic, A., Repovs, G., Barch, D.M. (2011). Emotion effects on attention, amygdala activation, and functional connectivity in schizophrenia. *Schizophrenia Bulletin*, **38**(5), 967–80.

Barrett, L.F., Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology*, **41**, 167–218.

Beauregard, M., Lévesque, J., Bourgouin, P. (2001). Neural correlates of conscious self-regulation of emotion. *The Journal of Neuroscience*, **21**(18), RC165.

Brosch, T., Grandjean, D., Sander, D., Scherer, K.R. (2008). Behold the voice of wrath: Cross-modal modulation of visual attention by anger prosody. *Cognition*, **106**(3), 1497–503.

Brosch, T., Grandjean, D., Sander, D., Scherer, K.R. (2009). Cross-modal emotional attention: emotional voices modulate early stages of visual processing. *Journal of Cognitive Neuroscience*, **21**(9), 1670–9.

Busso, C., Bulut, M., Lee, C.-C., *et al.* (2008). IEMOCAP: interactive emotional dyadic motion capture database. *Language Resources and Evaluation*, **42**(4), 335.

Chao-Gan, Y., Yu-Feng, Z. (2010). DPARSF: a MATLAB toolbox for 'pipeline' data analysis of resting-state fMRI. *Frontiers in Systems Neuroscience*, **4**, 13.

Costa, V.D., Lang, P.J., Sabatinelli, D., Versace, F., Bradley, M.M.J.H.b.m. (2010). Emotional imagery: assessing pleasure and arousal in the brain's reward circuitry. *Human Brain Mapping*, **31**(9), 1446–57.

Damasio, A.R., Geschwind, N. (1984). The neural basis of language. *Annual Review of Neuroscience*, **7**(1), 127–47.

Di, X., Biswal, B.B. (2018). Toward task Connectomics: examining whole-brain task modulated connectivity in different task domains. *Cerebral Cortex*, **29**(4), 1572–83.

Di, X., Gohel, S., Kim, E.H., Biswal, B.B. (2013). Task vs. rest-different network configurations between the coactivation and the resting-state brain networks. *Frontiers in Human Neuroscience*, **7**, 493.

Dimitriadis, S.I., Antonakakis, M., Simos, P., Fletcher, J.M., Papanicolaou, A.C. (2017a). Data-driven topological filtering based on orthogonal minimal spanning trees: application to multi-group magnetoencephalography resting-state connectivity. *Brain Connectivity*, **7**(10), 661–70.

Dimitriadis, S.I., Salis, C., Tarnanas, I., Linden, D.E. (2017b). Topological filtering of dynamic functional brain networks unfolds informative chronnectomics: a novel data-driven thresholding scheme based on orthogonal minimal spanning trees (OMSTs). *Frontiers in Neuroinformatics*, **11**, 28.

Eckstrand, K.L., Hanford, L.C., Bertocci, M.A., *et al.* (2018). Trauma-associated anterior cingulate connectivity during reward learning predicts affective and anxiety states in young adults. *Psychological Medicine*, 1–10.

El Ayadi, M., Kamel, M.S., Karray, F. (2011). Survey on speech emotion recognition: features, classification schemes, and databases. *Pattern Recognition*, **44**(3), 572–87.

Ewbank, M.P., Passamonti, L., Hagan, C.C., Goodyer, I.M., Calder, A.J., Fairchild, G. (2018). Psychopathic traits influence amygdala–anterior cingulate cortex connectivity during facial emotion processing. *Social Cognitive and Affective Neuroscience*, **13**(5), 525–34.

Gao, Z., Wang, S. (2015). Emotion recognition from EEG signals using hierarchical Bayesian network with privileged information. In: *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, Shanghai, China: ACM.

Glerean, E., Salmi, J., Lahnakoski, J.M., Jääskeläinen, I.P., Sams, M. (2012). Functional magnetic resonance imaging phase synchronization as a measure of dynamic functional connectivity. *Brain Connectivity*, **2**(2), 91–101.

Grandjean, D., Sander, D., Pourtois, G., *et al.* (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature Neuroscience*, **8**(2), 145–6.

Hertrich, I., Dietrich, S., Ackermann, H. (2016). The role of the supplementary motor area for speech and language processing. *Neuroscience and Biobehavioral Reviews*, **68**, 602–10.

Holmes, E.A., Mathews, A. (2005). Mental imagery and emotion: a special relationship? *Emotion*, **5**(4), 489.

Humphries, M.D., Gurney, K. (2008). Network 'small-world-ness': a quantitative method for determining canonical network equivalence. *PLoS One*, **3**(4), e0002051.

Hutchison, R.M., Womelsdorf, T., Allen, E.A., *et al.* (2013). Dynamic functional connectivity: promise, issues, and interpretations. *NeuroImage*, **80**, 360–78.

Jessen, S., Kotz, S. (2008). On the role of crossmodal prediction in audiovisual emotion perception. *Frontiers in Human Neuroscience*, **7**(369).

Just, M.A., Newman, S.D., Keller, T.A., McEleney, A., Carpenter, P.A. (2004). Imagery in sentence comprehension: an fMRI study. *Neuroimage*, **21**(1), 112–24.

Kim, M.J., Loucks, R.A., Palmer, A.L., *et al.* (2011). The structural and functional connectivity of the amygdala: from normal emotion to pathological anxiety. *Behavioural Brain Research*, **223**(2), 403–10.

Klapwijk, E.T., Goddings, A.-L., Burnett Heyes, S., Bird, G., Viner, R.M., Blakemore, S.-J. (2013). Increased functional connectivity with puberty in the mentalising network involved in social emotion processing. *Hormones and Behavior*, **64**(2), 314–22.

Kober, E., Barrett, L.F., Joseph, J., Bliss-Moreau, E., Lindquist, K., Wager, T.D. (2008). Functional grouping and cortical-subcortical interactions in emotion: A meta-analysis of neuroimaging studies. *Neuroimage*, **42**(2), 998–1031.

Lane, R.D., Reiman, E.M., Bradley, M.M., *et al.* (1997). Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia*, **35**(11), 1437–44.

Latora, V., Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical Review Letters*, **87**(19), 198701.

Liebenthal, E., Silbersweig, D.A., Stern, E. (2016). The language, tone and prosody of emotions: neural substrates and dynamics of spoken-word emotion perception. *Frontiers in Neuroscience*, **10**, 506.

Lindquist, K.A., Satpute, A.B., Wager, T.D., Weber, J., Barrett, L.F. (2015). The brain basis of positive and negative affect: evidence from a meta-analysis of the human neuroimaging literature. *Cerebral Cortex*, **26**(5), 1910–22.

Mitchell, R.L., Elliott, R., Barry, M., Cruttenden, A., Woodruff, P.W. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia*, **41**(10), 1410–21.

Newman, M.E. (2003). The structure and function of complex networks. *SIAM Review*, **45**(2), 167–256.

Nummenmaa, L., Glerean, E., Hari, R., Hietanen, J.K. (2014a). Bodily maps of emotions. *Proceedings of the National Academy of Sciences of the United States of America*, **111**(2), 646–51.

Nummenmaa, L., Saarimäki, H., Glerean, E., *et al.* (2014b). Emotional speech synchronizes brains across listeners and engages large-scale dynamic brain networks. *NeuroImage*, **102**, 498–509.

Oliveri, M., Babiloni, C., Filippi, M.M., *et al.* (2003). Influence of the supplementary motor area on primary motor cortex excitability during movements triggered by neutral or emotionally unpleasant visual cues. *Experimental Brain Research*, **149**(2), 214–21.

Olson, I.R., Plotzker, A., Ezzyat, Y. (2007). The enigmatic temporal pole: a review of findings on social and emotional processing. *Brain*, **130**(7), 1718–31.

Paulmann, S., Bleichner, M., Kotz, S.A. (2013). Valence, arousal, and task effects in emotional prosody processing. *Frontiers in Psychology*, **4**, 345.

Pekkola, J., Ojanen, V., Autti, T., *et al.* (2005). Primary auditory cortex activation by visual speech: an fMRI study at 3 T. *Neuroreport*, **16**(2), 125–8.

Phan, K.L., Wager, T., Taylor, S.F., Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage*, **16**(2), 331–48.

Poeppel, D., Idsardi, W.J., van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**(1493), 1071–86.

Power, J.D., Cohen, A.L., Nelson, S.M., *et al.* (2011). Functional network organization of the human brain. *Neuron*, **72**(4), 665–78.

Rubinov, M., Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *NeuroImage*, **52**(3), 1059–69.

Sander, D., Grandjean, D., Pourtois, G., *et al.* (2005). Emotion and attention interactions in social cognition: brain regions involved in processing anger prosody. *NeuroImage*, **28**(4), 848–58.

Saramäki, J., Kivelä, M., Onnela, J.-P., Kaski, K., Kertesz, J. (2007). Generalizations of the clustering coefficient to weighted complex networks. *Physical Review E*, **75**(2), 027105.

Taya, F., Sun, Y., Thakor, N., Bezerianos, A. (2014). Information transfer efficiency during rest and task a functional connectome approach. In: *2014 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, Lausanne, Switzerland: IEEE.

Vrana, S.R., Lang, P.J. (1990). Fear imagery and the startle-probe reflex. *Journal of Abnormal Psychology*, **99**(2), 189.

Wang, Z., Liu, J., Zhong, N., Qin, Y., Zhou, H., Li, K. (2012). Changes in the brain intrinsic organization in both on-task state and post-task resting state. *NeuroImage*, **62**(1), 394–407.

Wang, S., Zhu, Y., Yue, L., Ji, Q. (2015). Emotion recognition with the help of privileged information. *IEEE Transactions on Autonomous Mental Development*, **7**(3), 189–200.

Zhu, Y., Wang, S., Ji, Q. (2014) Emotion recognition from users' EEG signals with the help of stimulus videos. In: *2014 IEEE International Conference on Multimedia and Expo (ICME)*, Bandung, Indonesia: IEEE.

Zuo, N., Yang, Z., Liu, Y., Li, J., Jiang, T. (2018). Both activated and less-activated regions identified by functional MRI reconfigure to support task executions. *Brain and Behavior*, **8**(1), e00893.