NeuroImage 183 (2018) 280–290

Contents lists available at ScienceDirect

NeuroImage

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

Interpersonal synchronization of inferior frontal cortices tracks social interactive learning of a song

Yafeng Pan [a,c], Giacomo Novembre [b], Bei Song [a,d], Xianchun Li [a], Yi Hu [a,*]

[a] Shanghai Key Laboratory of Brain Functional Genomics, Key Laboratory of Brain Functional Genomics, Ministry of Education, School of Psychology and Cognitive Science, East China Normal University, Shanghai, People's Republic of China
[b] Department of Neuroscience, Physiology and Pharmacology, University College London, London, United Kingdom
[c] Neuropsychology and Functional Neuroimaging Research Unit (UR2NF), ULB Neuroscience Institute (UNI), Université Libre de Bruxelles, Bruxelles, Belgium
[d] Department of Musicology, Harbin Conservatory of Music, Heilongjiang, People's Republic of China

ARTICLE INFO

Keywords:
Social interactive learning
Interpersonal brain synchronization
fNIRS hyperscanning
Music
Turn-taking
Inferior frontal cortex
Involvement theory

ABSTRACT

Much of human learning emerges as a result of interaction with others. Yet, this interpersonal process has been poorly characterized from a neurophysiological perspective. This study investigated (*i*) whether Interpersonal Brain Synchronization (IBS) can reliably mark social interactive learning, and specifically (*ii*) during what kind of interactive behavior. We recorded brain activity from learner-instructor dyads using functional Near-Infrared Spectroscopy (fNIRS) during the acquisition of a music song. We made four fundamental observations. First, during the interactive learning task, brain activity recorded from the bilateral Inferior Frontal Cortex (IFC) synchronized across the learner and the instructor. Second, such IBS was observed in particular when the learner was observing the instructor's vocal behavior and when the learning experience entailed a turn-taking and more active mode of interaction. Third, this specific enhancement of IBS predicted learner's behavioral performance. Fourth, Granger causality analyses further disclosed that the signal recorded from the instructor's brain better predicted that recorded from the learner's brain than vice versa. Together, these results indicate that social interactive learning can be neurophysiologically characterized in terms of IBS. Furthermore, they suggest that the learner's involvement in the learning experience, alongside the instructor's modeling, are key factors driving the alignment of neural processes across learner and instructor. Such alignment impacts upon the real-time acquisition of new information and eventually upon the learning (behavioral) performance. Hence, besides providing a biological characterization of social interactive learning, our results hold relevance for clinical and pedagogical practices.

1. Introduction

Much of human learning emerges as a result of interaction with others (Marchiori and Warglien, 2008). As Greek philosophers highlighted, more than two thousand years ago, verbal "turn-taking" interactions are powerful means of education and pedagogical practice ("Socratic dialog", Kahn, 2013). Today, despite the rise of individual approaches to learning such as e-learning and multimedia learning (Clark and Mayer, 2016; Rennie and Morrison, 2013), learning through social interaction still plays a vital role in the daily lives of many people (Edwards-Groves, 2017). Examples of social interactive learning span from infancy to adulthood, including learning of new rules (Williamson et al., 2010), new words (Verga and Kotz, 2013, 2017), solution of novel problems (Williamson et al., 2008), recognition of emotional expressions (Carr et al., 2003), body movements (Liao et al., 2015), or the acquisition of fine motor skills such as music making (Belyk et al., 2016).

An influential pedagogical theory, namely the involvement theory (Astin, 1984, 1996), postulates that the quantity of interaction between learner and instructor is one of the main factors that facilitate social learning (Lundberg and Schreiner, 2004). Previous evidence supporting this hypothesis indeed showed that the amount of learner-instructor interactions (i.e. how often learners interact with an instructor) is positively related to gains in learners' skills (Bjorklund et al., 2004) and academic achievements (Anaya and Cole, 2001). Other studies found that more learner-instructor interactions probably enhance learners' expectations about their ability to succeed (Tauber, 1997), increase their engagement in academic activities (Umbach and Wawrzynski, 2005), and strengthen their critical thinking skills (Light, 2001). However, the

* Corresponding author.
E-mail address: yhu@psy.ecnu.edu.cn (Y. Hu).

https://doi.org/10.1016/j.neuroimage.2018.08.005
Received 29 November 2017; Received in revised form 19 July 2018; Accepted 3 August 2018
Available online 4 August 2018
1053-8119/

biological origins of these benefits remain largely elusive.

In this study, we aimed at providing a neurobiological testbed for the involvement theory. Our goal was to characterize the learner-instructor interaction from a neurophysiological perspective, and explore whether and how the quantity of social interactive learning might facilitate these neurophysiological processes. We adopted a musical paradigm entailing a music instructor teaching a music song to a group of learners – a paradigm that allowed us to study social interaction under ecological, yet controlled, experimental conditions (D'Ausilio et al., 2015). Indeed, music is an inherently social activity and has proven to be a good test case for social phenomena (D'Ausilio et al., 2015; Lindenberger et al., 2009; Osaka et al., 2015; Keller et al., 2017). Moreover, song learning involves social vocal imitation, which is a widespread learning behavior observable in both humans and other animals (Mooney, 2009; Patel, 2006; Patel et al., 2009).

We compared learning of a music song across two well-known pedagogical methods differing one another in terms of the amount of learner-instructor interactions that they imply (Klinger et al., 1998; Persellin et al., 2002; Persellin and Bateman, 2009). One of such methods, the "part learning" (PL), assumes that sensory information presented to the learner is typically organized into chunks or parts (Gobet et al., 2001; Miller, 1956). During PL, learners receive and retain knowledge as it is presented in repeated segments, rather than a complete whole (Herrold, 2005). Accordingly, PL involves many interactions between learner and instructor. The second method is the "whole learning" (WL), which provides learners with repeated exposure to the to-be-learned materials in their entirety and therefore may offer a greater sense of continuity and integrity of the materials (Klinger et al., 1998; Miller, 2005; Persellin and Bateman, 2009). This method is somehow in accordance with Gestalt's view that natural systems and their properties should be viewed as wholes, not as collections of parts (Köhler, 2015). For example, teaching songs through PL requires instructors to interact with learners on a phrase-by-phrase basis, whereas teaching through WL requires instructors to perform the entire song to learners. As a result, the PL compared with the WL implies more turn-taking interactions.

We recorded brain activity from twenty-four learner-instructor dyads using functional near-infrared spectroscopy (fNIRS). We reasoned that the information exchange between a learner and an instructor cannot be fully understood by examining the two brains in isolation. Therefore, we adopted a hyperscanning approach that permitted us to record activity from two brains simultaneously during an ecologically-valid interaction (Montague et al., 2002), and compute the degree of similarity or reciprocal influence of one brain signal over the other. Previous studies have demonstrated that social interactions are associated with enhanced interpersonal brain synchronization (IBS), as measured by EEG (e.g., Lindenberger et al., 2009; Hu et al., 2018), fNIRS (e.g., Cui et al., 2012; Cheng et al., 2015; Hu et al., 2017; Osaka et al., 2015; Pan et al., 2017), or fMRI (e.g., Koike et al., 2016; Saito et al., 2010). For example, using EEG-based hyperscanning, Lindenberger et al. (2009) reported increased interpersonal synchronization of sensorimotor regions when guitar dyads played a melody together. Recently, using fNIRS-based hyperscanning, Osaka et al. (2015) found that when two people sang or hummed face-to-face, their brain activities of inferior frontal cortex (IFC) were synchronized. Similarly, other studies have reported IBS between signals recorded from IFC during tasks entailing turn-taking interactions (Jiang et al., 2012; Liu et al., 2015). Compared to EEG, fNIRS is a more robust method for measuring brain activity during motor tasks in that it is less affected by motor artifacts. By measuring local hemodynamic effects, fNIRS is suitable for measuring IBS during social interactions in realistic situations that require active movement (Cui et al., 2015; Quaresima and Ferrari, 2016).

The goal of the current study was twofold. First, we computed IBS and compared it across the two distinct learning methods, namely PL and WL, entailing more (PL) or less (WL) interactions, respectively. To the extent that the involvement theory holds true, we would expect PL to lead to better learning performance compared to WL. Furthermore, we hypothesized that, to the extent that IBS can reliably mark the success of social interactive learning, then IBS should also be higher during PL as compared to WL. In accordance with the previous fNIRS hyperscanning literature (see above), we were particularly interested in monitoring brain activity over inferior frontal regions (Jiang et al., 2012; Liu et al., 2015; Osaka et al., 2015), and adapted our measurement patch in order to properly extract signals from these neural regions.

Second, we explored what behavior would best enhance IBS during interactive PL. To this end, we used video coding techniques to classify IBS recordings specifically associated either with OBSERVATION (i.e. the learners attending to the instructor's performance) or IMITATION (i.e. the learners performing under the instructor's supervision). Next, having dissociated the neural processes associated with OBSERVATION and IMITATION, we tested which process induced the higher IBS contributing to PL. Complementary to this, Granger causality analysis (GCA) was also used to provide a neurobiological suggestion of which individual (the learner or the instructor) was more actively driving the other. Finally, a series of correlational analyses were adopted as an exploratory investigation of potential IBS-behavior relationships (e.g., correlation between IBS and song learning performance).

2. Methods

2.1. Participants

Twenty-four undergraduate students (aged 20.58 ± 2.15 years) and a music instructor (22 years old) participated in the study. All participants were female, healthy and right-handed, and were recruited by advertisements on East China Normal University. No learner had received any formal musical training apart from compulsory school lessons, while the instructor had 13 years music learning experience. Each learner paired up with the instructor as a learner-instructor dyad, forming 24 participant dyads in total. This arrangement was intended to make the teaching style as similar as possible across dyads (Thepsoonthorn et al., 2016). Each participant signed an informed consent prior to the experiment and was paid ￥30 for participation. The study was approved by the University Committee of Human Research Protection (HR 044–2017), East China Normal University.

2.2. Materials and apparatus

Two Chinese songs were selected: "The Moon Reflection" (Lyrics: B. Peng, Music: Z. Liu and S. Yan) and "A Tune of Homesickness" (Lyrics: C. Qu, Music: Q. Zheng). These songs were selected because (*i*) they entailed simple lyrics and melodies (i.e., they were easy to be acquired) and (*ii*) they were likely to be unfamiliar to the participants (something we determined using a pre-screening survey and that was important in order to exclude potential confounding effects of prior knowledge) (Simmons-Stern et al., 2010). These two songs have similar musical structures (e.g., quadruple rhythm, eight bars, and slow tempo) and convey similar musical emotions and semantics (i.e., nostalgia, missing home). Note that each song is composed of four phrases, in accordance with previous song learning tasks (e.g., Klinger et al., 1998). Performing one musical phrase at the correct tempo would take approximately 6 s, while performing the entire song would take around 24 s. Two songs were selected (instead of only one) in order to exclude the possibility that any detected effect was specific to a certain song and hence limited in generalizability. Within both the PL ($N = 12$) and WL ($N = 12$) groups, six dyads were assigned with "The Moon Reflection" and six with "A Tune of Homesickness". Because the results reported hereafter did not differ across performance of the two songs, these were pooled together.

A digital video camera (Sony, HDR-XR100, Sony corporation, Tokyo, Japan) was used to record the vocal and non-vocal interactions within dyads through the experiment. The camera recordings were used to classify (following the experiment) behaviors such as listening, singing, facial expressions, gestures and orofacial movements.

2.3. Tasks and procedures

The task was split into three phases (Fig. 1A). During the initial resting phase (3 min), both participants (sitting face-to-face, 0.8 m apart) were asked to relax and to remain still while keeping their eyes closed. Such 3-min resting phase served as the baseline.

The resting phase was followed by the learning phase (~9 min), whose procedures (similar to Klinger et al., 1998) are reported in Table 1. In both PL and WL, the instructor, who was blind to the purpose of the experiment, began the task by singing the entire song twice (i.e. initial learning sub-phase, ~1 min). Next, the learner and instructor engaged in the interactive learning (interactive learning sub-phase, 8 min) through either of two learning methods: (*i*) PL, namely learning the song phrase by phrase; (*ii*) WL, namely learning the song in a holistic way. Specifically, in the PL, the learner attended (OBSERVATION) and then imitated (IMITATION) each single phrase performed by the instructor. So the learner and the instructor interacted in a turn-taking fashion. In contrast, in WL, the learner attended and then imitated the entire song (Fig. 1B). These procedures allowed participants to attend and imitate the song for an equal time across the two methods (Persellin and Bateman, 2009)

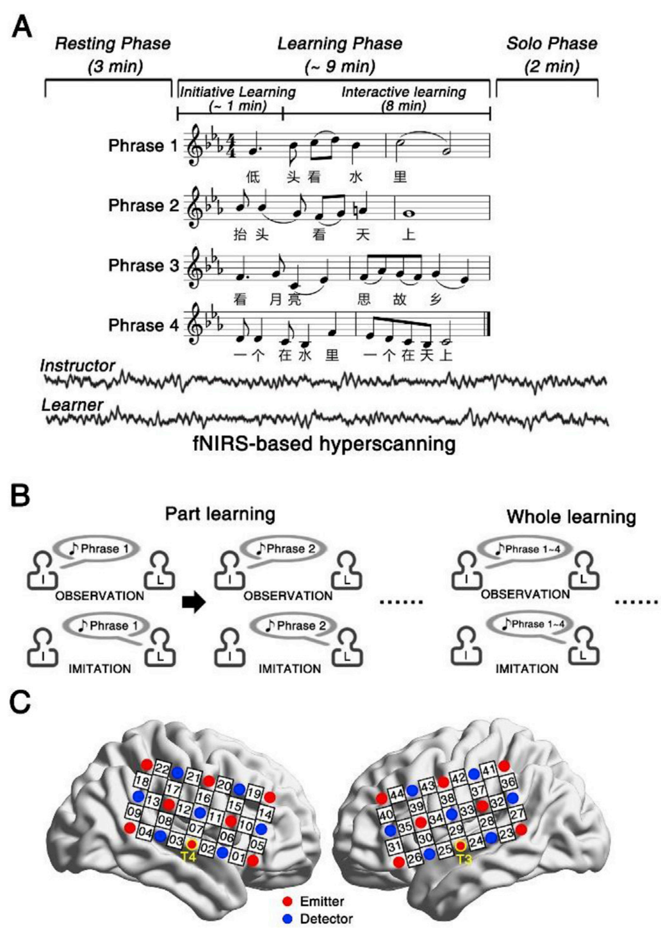

Fig. 1. Experimental task and procedures. (**A**) Dyads composed of a learner and an instructor start the task by resting with eyes closed (resting phase/baseline), then the instructor teaches a song to the learner (learning phase), and finally the learner sings the newly acquired song (solo phase). During the whole procedure, brain activities from the instructor and the learner are acquired simultaneously using fNIRS. **(B)** Two learning methods are compared. In the "part learning" (PL) method, the learner attends and imitates the song in a phrase-by-phrase fashion. In the "whole leaning" (WL) method, the learner attends all phrases at once and then imitates the whole song (I = instructor, L = learner). **(C)** Optode probe set. The set was placed over the bilateral fronto-temporo-parietal cortices.

Table 1
Procedures associated with the part learning (PL) and whole learning (WL) methods.

The part learning (PL)
- *Initial learning sub-phase*
The instructor sings the entire song to the learner twice.
- *Interactive learning sub-phase*
The learner observes while the instructor sings phrase 1 (OBSERVATION).
The learner imitates phrase 1 under the instructor's supervision (IMITATION).
The learner observes while the instructor sings phrase 2 (OBSERVATION).
The learner imitates phrase 2 under the instructor's supervision (IMITATION).
The learner observes while the instructor sings phrase 3 (OBSERVATION).
The learner imitates the phrase 3 under the instructor's supervision (IMITATION).
The learner observes while the instructor sings phrase 4 (OBSERVATION).
The learner imitates phrase 4 under the instructor's supervision (IMITATION).
(OBSERVATION alternates with IMITATION until the end of this phase)
The whole learning (WL)
- *Initial learning sub-phase*
The instructor sings the entire song to the learner twice.
- *Interactive learning sub-phase*
The learner observes while the instructor sings all four phrases consecutively (OBSERVATION).
The learner imitates the four phrases under the instructor's supervision (IMITATION).
(OBSERVATION alternates with IMITATION until the end of this phase)

while being continuously supervised by the instructor. On average, the songs were repeated 7.83 ± 1.03 ($M \pm SD$) times in the PL and 6.75 ± 2.01 times in the WL (and these values were comparable across the two groups; $t_{22} = 1.66, p = 0.11$). All participants were allowed to use non-vocal communication (e.g., facial expressions, gestures) to facilitate the acquisition of the song. Learners were randomly allocated to one of the two learning methods, half for the PL (n = 12) and half for the WL (n = 12).

After the learning phase, learners were instructed to sing the entire song as best as they could (solo phase). They sang the song repeatedly until the end of this phase (2 min). This practice was recorded and later used to assess how well the learners had acquired the song (see below).

2.4. Pre- and post-experiment assessments

Before the experiment, learners completed an online test assessing their musical skills (Mandell et al., 2007; http://jakemandell.com/tonedeaf/; range of score: 0–100%; higher score indicates better pitch discrimination and musical memory abilities).

After the experiment, the two participants forming each dyad completed a battery of subjective ratings that assessed their impressions of the experiment and the interacting partner. Both participants rated the nervousness, awkwardness, difficulty, empathy, satisfaction, and likability (see Table S1 for details). The ratings were on a 7-point Likert scale, which ranged from 1 ("not very much") to 7 ("very much"). No discussion was allowed during the rating task.

2.5. Video-recording data processing

Four graduate students were recruited to independently code activities in the interactive learning phase. Similar to previous studies (Jiang et al., 2012, 2015), three types of activities were categorized: (*i*) vocal interactions (VI), such as the learners' observation of the instructors' modeled singing (VI during OBSERVATION), or learners' imitation of the singing under the supervision of the instructor (VI during IMITATION), (*ii*) non-vocal interactions (NVI), such as body language (including facial expressions, sign gestures); and (*iii*) no interaction (NI). Each second (s) from the 8 minutes entailing the interactive learning phase was coded either as VI, NVI, or NI. If an interactive activity consisted of both VI and NVI for a given second (s), the dominant behavior (i.e., the behavior exerting greater impact on the interaction) was coded. For all coding activities, inter-coder reliability was calculated by the intra-class correlation (Werts et al., 1974). Inter-coder reliability was 0.84 for the VI (vs. NVI) and 0.87 for the OBSERVATION (vs. IMITATION). If there was an inconsistency, the four coders discussed it and made an agreement on it. Based on the coded activities, two indices were calculated. One was the number of interactive learning activities, which was calculated as the cumulative number of sequential vocal or non-vocal interactions (for example, within a 10-s activity coded as "VI VI VI VI NVI VI VI NVI NVI VI", the cumulative number of VI and NVI would be 3 and 2, respectively). The other was the duration ratio of interactive learning activities, which was calculated as the proportions of time (out of 8 min) when vocal or non-vocal interactions occurred (see Jiang et al., 2015).

Two music experts, who were blind to the group assignment, rated the singing performance of learners in the solo phase. The experts had at least 15-year experience in music teaching. Six aspects of music performance were evaluated through 7-point scales (i.e., melody, rhythm, lyric, pitch, emotion, and tune; see Table S2 for details). For each learner, inter-coder reliability was calculated by the intra-class reliability on six aspects (ranging from 0.765 to 0.912). The rating scores provided by the two experts were averaged. The sum of the judgements made on all six aspects (for a given learner) was considered as the index of overall learning performance (maximum score: 7 points $\times$ 6 aspects = 42).

2.6. Image acquisition

Signals were acquired using ETG-7100 optical topography system (Hitachi Medical Corporation, Japan), measuring the absorption of near infrared light (two wavelengths: 695 and 830 nm). Signals were sampled at 10 Hz. The oxyhemoglobin (HbO) and deoxyhemoglobin (HbR) were obtained through the modified Beer-Lambert law. In this study, we only focused on the HbO concentration, which is frequently used in fNIRS-based hyperscanning studies (Cheng et al., 2015; Jiang et al., 2015; Pan et al., 2017; Tang et al., 2015). Furthermore, previous evidence indicated that HbO is the most sensitive indicator of changes in the cerebral blood flow in fNIRS measurements (Hoshi, 2007).

Two optode probe sets were used to cover each participant's left and right fronto-temporo-parietal regions (Fig. 1C), which were associated with the production of words and melody during singing (Osaka et al., 2015), and also with social interactions (Decety et al., 2004; Pan et al., 2017). Specifically, one 3 $\times$ 5 optode probe set (eight emitters and seven detectors forming 22 measurement points with 3 cm optode separation) was placed over the right frontal, temporal, and neighboring parietal cortices. The middle optode of the lowest probe row of the patch was placed at T4 (Fig. 1C), following the international 10–20 system (Okamoto et al., 2004). Another probe set (of identical size) was placed symmetrically over the left hemisphere (T3 corresponds to T4, Fig. 1C). The middle probe set columns were placed along the sagittal reference curve. Channels (CHs) 1–22 and 23–44 represent 22 point-of-interest in the right and left hemispheres, respectively. The correspondence between the NIRS CHs and the measured points on the cerebral cortex was determined using the virtual registration approach (Singh et al., 2005; Tsuzuki et al., 2007).

2.7. Imaging-data analyses

2.7.1. Interpersonal brain synchronization (IBS)

Data collected during the resting phase (3 min, served as the baseline) and interactive learning phase (8 min, served as the task) were entered into the interpersonal brain synchronization (IBS) analysis.

During preprocessing, a principal component spatial filter algorithm was used to remove the global components in the fNIRS data (for more details, see Zhang et al., 2016). To remove head motion artifacts, we used a "Correlation Based Signal Improvement" (CBSI) method (see more details in Cui et al., 2010), which is based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. To remove slow drifts, a set of discrete cosine basic functions with a cutoff period of 128 s (i.e., 7.81×10^{-3} Hz) was also used to perform the high-pass filtering in the fNIRS signals (Ikeda et al., 2017).

We then employed wavelet transform coherence (WTC) analysis to estimate IBS, here entailing the relationship between HbO time series within each dyad (Grinsted et al., 2004). The WTC ranges between 0 and 1, and can be conceptualized as a localized *correlation coefficient* in time and frequency space (Chang and Glover, 2010; Grinsted et al., 2004). For the wavelet function in WTC analysis, we used the Morlet wavelet, in accordance with previous studies (Grinsted et al., 2004; Nozawa et al., 2016). As a first step, we estimated whether the interactive learning task enhanced IBS (estimated by WTC) with respect to baseline. In order to do so, IBS (averaged across all channels in each dyad) was compared between the resting phase (i.e. baseline session) and the interactive learning phase (i.e. task session) using a series of paired sample *t*-tests, one for each frequency band (frequency range: 0.01–1 Hz; period range: 1–100 s) (Ikeda et al., 2017; Nozawa et al., 2016; Xue et al., 2018). Note that this range covers nearly all frequencies investigated in previous fNIRS hyperscanning studies (e.g., Cui et al., 2012; Nozawa et al., 2016; Jiang et al., 2012, 2015; Pan et al., 2017). This analysis yielded a series of *p*-values that were FDR corrected (threshold for significance was 0.05, Benjamini and Hochberg, 1995). The results indicated that IBS associated with the interactive learning phase was significantly higher than that associated with the resting phase for frequencies ranging between 0.07 and 0.10 Hz (i.e., period 10.50–14.01 s) and 0.11–0.15 Hz (i.e., period 6.61–9.35 s) (see Fig. S1). These two ranges (separated by only one period) were clustered and chosen as our frequency of interest (FOI): 0.07–0.15 Hz (corresponding to period between 6.61 and 14.01 s). Note that this range nicely encompasses the temporal structure of the task performed by the PL group, especially considering that performing one musical phrase took approximately 7 s on average (6.64 $\pm$ 1.56 s in this study). Also note that this frequency band of interest enabled the removal of high- and low-frequency undesired signals, such as those related to cardiac pulsation ($\sim$1 Hz) and respiration ($\sim$0.2–0.3 Hz).

Next, we averaged IBS within this FOI, and compared our two groups of participants. To do so, we computed an index of the interactive-learning-related IBS by averaging IBS within the frequency of interest (i.e., 0.07–0.15 Hz) and computing the difference between the interactive learning phase and the resting phase. The resulting values were then entered into two complementary analyses. The first one, in line with what done previously, contrasted these values vs. zero for each channel using one-sample *t*-tests and aimed at identifying group specific enhancements of IBS. The second one contrasted the values across groups for each channel. For both analyses, the resulting *p* values were FDR corrected. The results yielded three *t* maps, two of them being group-specific and one resulting from the comparison across the two groups. These *t* maps reflected the interactive-learning-related IBS, and they were generated using a spatial interpolation linear method, separately for the right and left hemispheres. The MNI coordinates and *t* values of *t* maps were first converted into *.img files using xjView (nirs2img.m, http://www.alivelearn.net/xjview). The converted data were then rendered over the 3D brain model using BrainNet Viewer (Xia et al., 2013).

To determine whether such IBS was specific to two participants who constituted a dyad, we conducted a control analysis. The real learner-instructor pairs from the PL group were re-paired in a pseudo-random way so that 12 new pseudo pairs were created (e.g. time series from the instructor in dyad 1 were paired with time series from the learner in dyad 3) (Fig. S2A). Then, IBS was estimated again from the pseudo dyads.

Finally, we explored neural-behavioral relations. To this aim, the relationship between interactive-learning-related IBS and learning (behavioral) performance was evaluated using Pearson correlation analyses.

2.7.2. Segments of interactive-learning-related IBS

We selected the CHs that showed increased IBS during the interactive learning phase relative to the resting phase. The time course of IBS in the selected CHs was downsampled to 1 Hz, so that point-to-frame correspondence between the time series and video recordings was obtained

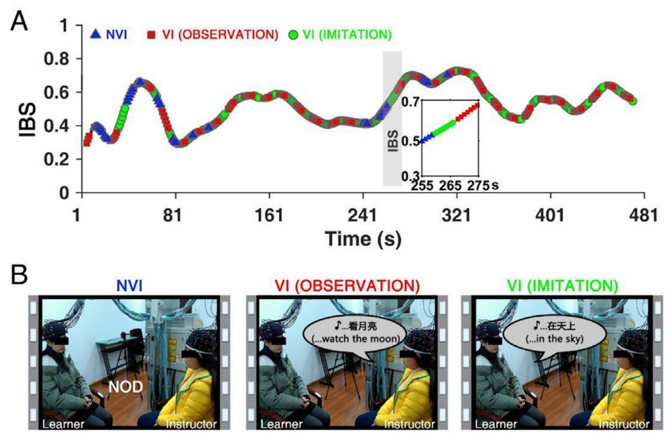

Fig. 2. Interpersonal brain synchronization (IBS) is analyzed as a function of the dyad behavior. (**A**) Time course of IBS for one randomly selected dyad from the part learning (PL) group. The grey box indicates a representative selected chunk of the time course. Blue points indicate non-vocal interaction (NVI); red points indicate vocal interaction (VI) during OBSERVATION; green points indicate VI during IMITATION. (**B**) The corresponding interactive learning behaviors coded from video frames.

(Fig. 2). Based on the results of coding procedures aforementioned, we categorized the segments of IBS in the learning phase as VI-IBS, NVI-IBS, and NI-IBS, for IBS corresponding vocal, non-vocal, no interactions, respectively. VI-IBS was further divided into VI-IBS during OBSERVATION (learners' observation of instructor's modeling) and during IMITATION (learners' imitation with instructor's supervision). The interactive-learning-related IBS data were adjusted for the delay-to-peak effect (~6 s) in the fNIRS signal (Cui et al., 2009; Jiang et al., 2015). Finally, these segments of IBS were compared between the PL and WL groups using a series of independent sample t tests. The Bonferroni correction was used to account for multiple comparisons.

2.7.3. Coupling directionality

Granger causality analysis (GCA) was used to provide a neurobiological suggestion of coupling directionality, i.e. which individual (the learner or the instructor) was more actively driving the other. GCA using vector autoregressive (AR) models has been successfully applied to measure the causal relationship between fNIRS time series data (e.g., Jiang et al., 2015; Pan et al., 2017). Here we used a linear AR model, which can achieve greater accuracy in detecting network connectivity than the widely used pair-wise granger causality model (Zhou et al., 2011). Our GCA protocol consisted of five steps. First, the raw data were preprocessed. Note that the preprocessing procedure (e.g., filtering and global effect removing, see above) made the time series relatively stationary (Guo et al., 2008). We then converted the preprocessed (task-related) signals into z-scores using the mean and the standard deviation of the signals recorded during the rest (baseline) session. This normalization was performed separately for each channel, in accordance with previous research showing that raw data of NIRS at different channels are not directly comparable (Azuma et al., 2013; Horaguchi et al., 2008; Matsuda and Hiraki, 2006; Schroeter et al., 2003). Second, clean time series from channels associated with significant IBS were averaged as a region of interest (ROI). ROI-based data were submitted to subsequent analyses. Third, segments of data associated with either OBSERVATION or IMITATION during vocal interactions (see above) were concatenated (cf. Handberg and Lund, 2014; Kirchner et al., 2014). Fourth, using Granger Causality Estimation toolbox (Guo et al., 2008; see details in http://www.dcs.warwick.ac.uk/~feng/causality.html), we calculated the mean causalities of the pairs of signals. Specifically, we calculated the conditional Granger Causality, which was used to infer the original direct relationship between multi-variable time series (Chen et al., 2006), for both directions (i.e., from instructor to learner, and from learner to instructor). All signals recorded from channels associated with nonsignificant IBS (from the instructor, latent variable Z1, and from the learner, latent variable Z2) were averaged and included as latent variables. Note that we used the average signal from nonsignificant channels, as opposed to the individual signals from each nonsignificant channel, in order to eliminate global systemic noises. The order of the AR model was determined to be 15 based on the Bayesian information criterion (Schwarz, 1978). There was no significant autocorrelation in the residual with the determined model order based on Ljung-Box Q-tests ($qs < 10.02$, $ps > 0.10$). Finally, we used nonparametric Kolmogorov-Smirnov tests to examine whether each direction differed from zero, and Wilcoxon tests to compare differences between the two directions. We adopted nonparametric tests due to the data being not normally distributed. The Bonferroni correction was used to account for multiple comparisons.

3. Results

3.1. Behavioral results

3.1.1. Learning performance

The overall performance of the PL group and the WL group was compared using an independent sample t-test. This analysis identified a significant group difference ($t_{(22)} = 2.26$, $p < 0.05$, Cohen's $d = 0.92$), indicating that the learning performance of PL group ($M \pm SE$, 28.63 ± 1.93) was higher than that of the WL group (21.63 ± 2.43) (Fig. 3A). Hence, PL led to better learning performance than WL.

3.1.2. Interactive learning activities

The number of interactive learning activities was compared between the two groups. The PL group displayed a significantly larger number of vocal interactions (VI) than the WL group (PL vs. WL: 31.75 ± 2.42 vs. 6.67 ± 1.92; $t_{(22)} = 28.14$, $p < 0.001$, Cohen's $d = 11.48$). Instead, the number of non-vocal interactions (NVI) was comparable across groups (PL vs. WL: 5.25 ± 1.07 vs. 6.67 ± 1.61; $t_{(22)} = 0.73$, $p = 0.47$) (Fig. 3B).

The duration ratio of interactive learning activities was also compared between the two groups. The PL group did not differ significantly from the WL group in the duration ratio of VI (0.84 ± 0.03 vs. 0.85 ± 0.02; $t_{(22)} = 0.39$, $p = 0.70$), and NVI (0.14 ± 0.02 vs. 0.11 ± 0.01;

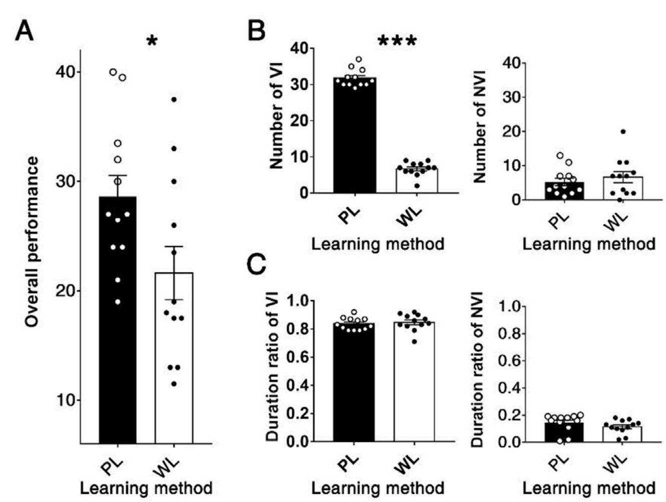

Fig. 3. Learning performance and interactive learning activities. (**A**) Group-averaged learning performance across PL and WL groups. Each point indicates one participant. (**B**) Number of interactive learning activities. Left panel: number of VI (vocal interaction). Right panel: number of NVI (non-vocal interaction). Each point indicates one dyad. **(C)** Duration ratio of interactive learning activities. Left panel: Duration ratio of VI (vocal interaction). Right panel: Duration ratio of NVI (non-vocal interaction). Each point indicates one dyad. PL = part learning, WL = whole learning. Error bars represent standard error. $^{*}p < 0.05$. $^{***}p < 0.001$.

$t_{(22)} = 1.09$, $p = 0.30$) (Fig. 3C).

These behavioral results confirmed that (*i*) PL entailed a more "turn-taking" mode of VI compared to WL and that (*ii*) the PL and WL groups invested a comparable amount of time in VI and NVI throughout the experiment.

3.1.3. Pre- and post-experiment assessment

Pre-experiment assessment. Prior musical skills (as assessed using an online test on musical perception ability before the experiment) were not different across the two groups (PL vs. WL, 0.70 ± 0.10 vs. 0.70 ± 0.15; $t_{(22)} = 0.11$, $p = 0.91$, Cohen's $d = 0.16$).

Post-experiment assessment. The two groups were not significantly different in terms of self-reported nervousness (PL vs. WL, 2.54 ± 0.29 vs. 2.79 ± 0.24), awkwardness (2.58 ± 0.21 vs. 3.17 ± 0.33), difficulty (3.29 ± 0.46 vs. 4.00 ± 0.49), empathy (4.53 ± 0.19 vs. 4.21 ± 0.22), satisfaction (5.79 ± 0.26 vs. 5.17 ± 0.38), and likeability (6.17 ± 0.11 vs. 5.92 ± 0.19), $ps > 0.05$.

3.2. Brain imaging results

3.2.1. Part learning (PL) induces IBS in bilateral inferior frontal cortex (IFC)

During interactive learning, IBS increased (as compared to baseline) in the frequency band comprised between 0.07 and 0.15 Hz (6.61–14.01 s, see methods and Fig. 4A). Within this frequency of interest, a significant interactive-learning-related IBS was identified in the PL group at CH10 (0.08 ± 0.02), CH14 (0.08 ± 0.02), CH31 (0.10 ± 0.02), and CH40 (0.09 ± 0.02), $ts > 4.78$, $ps < 0.05$, Cohen's $ds > 2.56$, FDR-controlled (Fig. 4B). These channels roughly cover the bilateral inferior frontal cortex (IFC, Tzourio-Mazoyer et al., 2002). No interactive-learning-related IBS was detected for any CH of the WL group, $ps > 0.05$.

Moreover, interactive-learning-related IBS was significantly different between the PL and WL groups for all the aforementioned channels, $ts > 2.83$, $ps < 0.05$, Cohen's $ds > 0.79$ (Fig. 4C), but not for the remaining CHs. Control analyses confirmed that these patterns of IBS were specifically associated with the real time interaction within dyads as they could not be observed in pseudo dyads ($ts < 1.51$, $ps > 0.05$, Fig. S2). To examine whether such increases in IBS were lateralized, we further explored the IBS difference between two hemispheres in the PL group: left IFC (mean of IBS at CH10 and CH14) vs. right IFC (mean of IBS at CH31 and CH40). However, no significant lateralization was observed, $t_{(11)} = 0.88$, $p > 0.05$.

Average IBS in bilateral IFC (mean of IBS at CH10, CH14, CH31 and CH40) and overall learning performance were positively correlated, $r = 0.58$, $p < 0.05$ (Fig. 4D), indicating that the IBS found in the PL group was associated with learners' performance: better learning was associated with stronger IBS. However, similar correlations were not significant in the WL group, $r = -0.20$, $p = 0.54$ (Fig. 4D).

3.2.2. IBS during vocal interaction (VI) predicts learning performance

The CHs that exhibited significant interactive-learning-related IBS (i.e., CH10, CH14, CH31, and CH40) were averaged separately for specific interactive learning activities (i.e., VI-IBS, NVI-IBS, and NI-IBS) and compared across groups (i.e., PL and WL).

We observed that VI-IBS was significantly higher in the PL compared to the WL group (0.09 ± 0.02 vs. 0.05 ± 0.01; $t_{(22)} = 3.54$, $p < 0.05$, Cohen's $d = 0.74$). In contrast, no significant group differences were found for NVI (0.07 ± 0.02 vs. 0.06 ± 0.02; $t_{(22)} = 1.27$, $p = 0.65$), or NI (0.05 ± 0.02 vs. 0.06 ± 0.02; $t_{(22)} = 0.17$, $p = 0.88$) (Fig. 5A). Thus, somehow in accordance with the behavioral results, the groups' neural dynamics differed particularly during VI.

We next compared the relationship between IBS and learning performance across groups. A significant correlation between VI-IBS and overall learning performance was disclosed only in the PL group, $r = 0.63$, $p = 0.03$, uncorrected (Fig. 5B). This result also appeared to be specific for VI, as the same analysis did not yield significant results for NVI or NI ($ps > 0.58$). Moreover, these analyses performed on the WL group's data yielded no significant results for VI-IBS, NVI-IBS, or NI-IBS ($ps > 0.50$).

3.2.3. IBS associated with OBSERVATION during vocal interaction (VI) predicts learning performance

We explored the specific contribution of the neural processes associated with OBSERVATION and IMITATION during VI and tested which one of those would induce higher IBS contributing to PL.

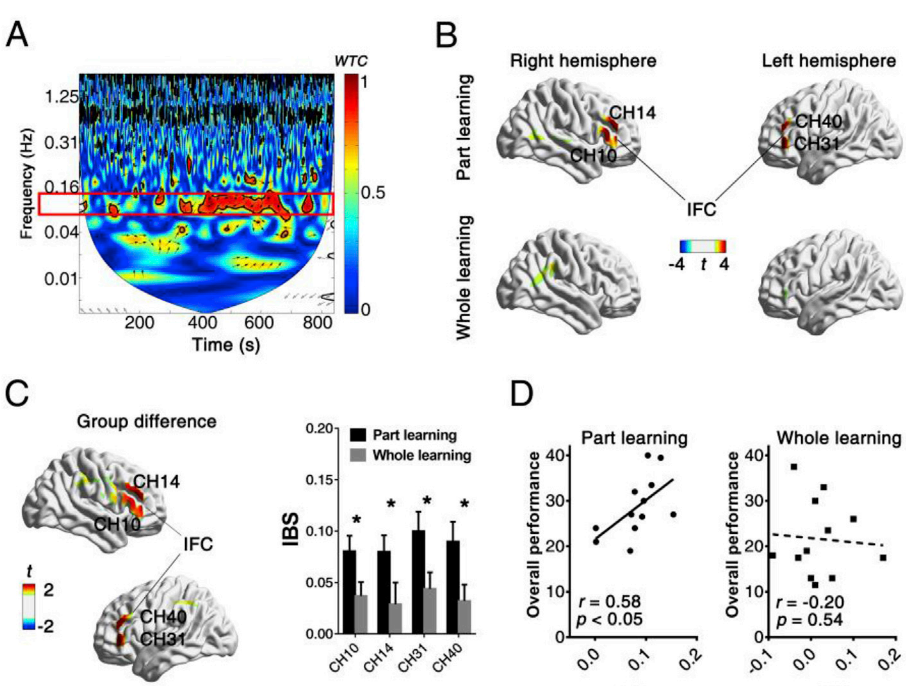

Fig. 4. Interpersonal brain synchronization (IBS). (**A**) Wavelet transform coherence (WTC) estimating IBS. The data displayed is based on raw HbO signal recorded from channel 31 (CH31) of a representative dyad. The red border line represents the frequency band of interest (0.07–0.15 Hz, corresponding to period of 6.61–14.01 s). Higher/lower coherence is encoded by red/blue colors, respectively. (**B**) One-sample *t*-test maps of IBS. The CHs that showed significant IBS are marked with channel labels. (**C**) IBS group differences at CH10, CH14, CH31, and CH40. (D) Correlations between average IBS in the bilateral inferior frontal cortex and learning performance in two groups. Error bars represent standard errors. *$p < 0.05$.

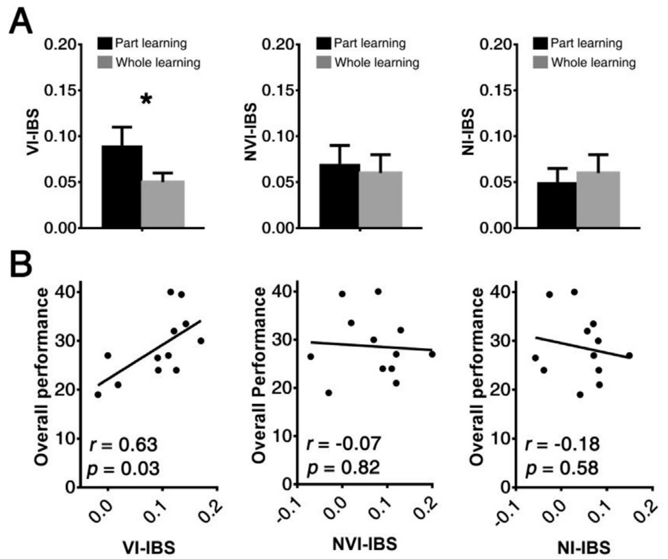

Fig. 5. Interpersonal brain synchronization (IBS) and interactive learning activities. (**A**) IBS during vocal interactions (VI-IBS), non-vocal interactions (NVI-IBS), and no interactions (NI-IBS). Only VI-IBS was larger in the PL group than in the WL group. (**B**) The relationship between IBS and learning performance during VI, NVI, and NI in the part leaning (PL) group (n = 12). Only the correlation between VI-IBS and learning performance was significant. Error bars represent standard errors. $^*p < 0.05$.

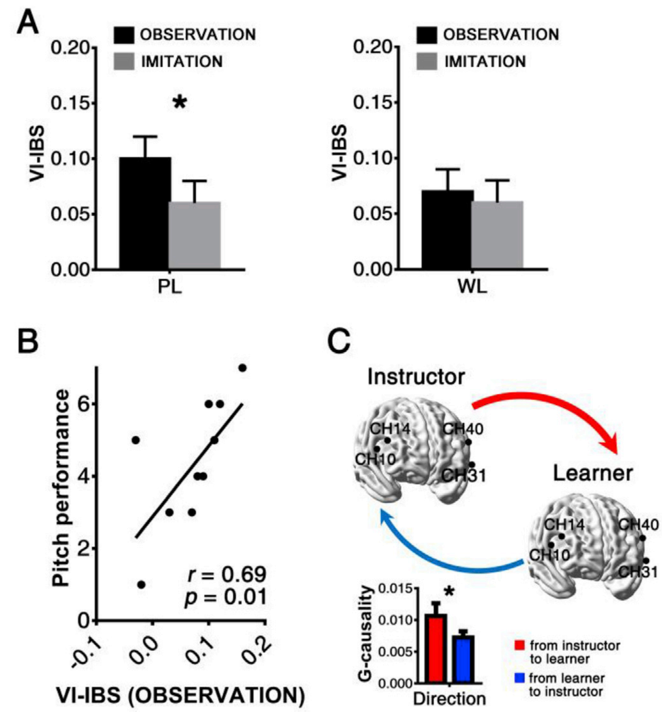

Fig. 6. IBS during vocal interaction. (**A**) IBS during vocal interaction (VI-IBS) for learners' OBSERVATION and IMITATION. OBSERVATION was associated with higher VI-IBS than IMITATION in the part learning (PL) group, but not in the whole learning (WL) group. (**B**) The VI-IBS during OBSERVATION in the PL group was correlated with learners' pitch performance (n = 12). (**C**) Granger causality analysis (GCA) results for VI-IBS during OBSERVATION. The mean G-causality from instructor to learner was significantly larger than that from learner to instructor. Error bars represent standard errors. $^*p < 0.05$.

First, we observed that VI-IBS was higher during OBSERVATION than during IMITATION, specifically in the PL group (0.10 ± 0.02 vs. 0.06 ± 0.02; $t_{(11)} = 3.46$, $p < 0.05$, Cohen's $d = 0.57$), but not in the WL group (0.07 ± 0.02 vs. 0.06 ± 0.01; $t_{(11)} = 0.84$, $p > 0.05$) (Fig. 6A).

Second, we conducted a series of correlational analyses to explore the relationship between the VI-IBS and learning performance in the PL group. A significant correlation between VI-IBS and pitch performance was observed during OBSERVATION ($r = 0.69$, $p = 0.01$, uncorrected) (Fig. 6B), but not IMITATION ($p > 0.11$).

Third, we carried out Granger causality analyses (GCA) to explore the directionality of the coupling during vocal interactions (separately for OBSERVATION and IMITATION) in the PL group. The results of this analysis indicated that, during OBSERVATION, both directions yielded significant increases in the mean G-causality relative to zero: from the instructor to learner (0.011 ± 0.002), as well as from the learner to instructor (0.007 ± 0.001), $ps < 0.05$. The Wilcoxon test further revealed that the mean G-causality from the instructor to the learner was significantly larger than that from the learner to instructor, $p < 0.05$ (Fig. 6C). In contrast, during IMITATION, there was no evidence of coupling directionality, $ps > 0.05$.

These results indicate that IBS was particularly enhanced during VI OBSERVATION – as opposed to VI IMITATION. Furthermore, the results showed that during VI OBSERVATION, the signal measured from the learner's brain was better predicted by that measured from the instructor's brain than vice versa. These results, taken together, could possibly imply that the dyad engagement was maximal while the learners were observing the instructor, and that the instructor brain's activity was driving the learner's one during this phase of the task.

4. Discussion

In the current study, we used a fNIRS-based hyperscanning approach to characterize real-time learner-instructor social interactive learning from a neurophysiological perspective. We measured neural activity and behavioral performance while learners were acquiring a song from an instructor using two different methods. One method – whole learning (WL) – entailed learning through one exposure to the whole song, while the other – part learning (PL) – entailed learning the song part-by-part through several turn-taking interactions. The two methods were meant to entail more (PL) or less (WL) social interactions, hence providing a testbed for the involvement theory (Astin, 1984, 1996), according to which the quantity of interaction between a learner and an instructor is a key factor that facilitates social learning. Supporting this theory, we observed that PL led to better learning performance than WL. Furthermore, we provide a neurophysiological characterization of such enhancement in performance: PL was associated with stronger interpersonal brain synchronization (IBS) between learner and instructor, particularly when the learner was observing (as opposed to imitating) the performance of the instructor. The enhanced IBS was found in bilateral inferior frontal cortex (IFC).

4.1. *Interpersonal brain synchronization (IBS) during turn-taking interactive learning*

We observed IBS during an interactive learning task implying acquisition of a music song. This result was in accordance with other findings reporting that IBS is associated with other interactive activities, such as cooperation (Cui et al., 2012), competition (Liu et al., 2015), imitation (Holper et al., 2012) and dialog (Holper et al., 2013). Taken together with our results, this research collectively indicates that IBS can track the successful transfer of information between two individuals, in this case between a learner and an instructor.

Previous IBS reports have been criticized on the basis of the following issue: One may argue that patterns of IBS could simply reflect functional similarities between two brains processing the same sensory information or performing the same movements at the same time (see Lindenberger et al., 2009; Burgess, 2013; Novembre et al., 2017). From this perspective, IBS would be seen as an epiphenomenon originating from intrinsic similarities reflected in the brain signals of two motorically or perceptually synchronized people, regardless their interactivity (Abrams et al., 2013; Hasson et al., 2008; Nummenmaa et al., 2012).

Importantly, this issue does not apply to our study because of the following two reasons. First, our IBS findings were specific to for one group of participants, namely those that undertook PL (as opposed to WL) although the two groups performed similar actions and processed similar sensory information. While IBS in the PL increased markedly from rest session to task session, no significant IBS associated with WL interactions was detected. One might argue that we did not observe such effect in the WL group because our FOI nicely encompassed the temporal structure of the task performed by the PL group but not that of the WL group. Yet, ruling out this criticism, we also compared IBS across rest and task sessions for frequency bands encompassing the temporal structure of the task performed by the WL group, and these tests did not yield any significant result. Second, strongly highlighting the functional significance of IBS, we observed that the IBS increase associated with the PL group was positively correlated with the behavioral learning performance of these participants. Taken together, these results rule out that IBS emerged as a consequence of the similarity between the sensory or motor processes across the participants forming each dyad.

Instead, we believe that our results would be better conceptualized if discussed in relation to social interactive learning in general, and the involvement theory in particular (Astin, 1984, 1996; Zhou and Cole, 2017). According to this theory, "involvement" is manifested through the investment of physical and psychological energy in the academic experience (the involvement theory in fact addresses both quantity and quality components of interaction, but this study focused only on the former component, Astin, 1984, 1996). From this perspective, the key factor facilitating learning is the learner's involvement, as reflected in his-her active participation during a class or level of interaction with the instructor.

This theory fits nicely with our observations. In fact, we observed that learning through PL method, indeed implying more active participation and more turn-taking interactions, led to better learning performance. We further speculate that the notion of "involvement" might be empirically characterized in terms of IBS. In support of this, we observed higher IBS during PL, in particular when the learners were actively observing the instructor's behavior. Thus, IBS might support interpersonal information transfer by aligning neural processes in the learner and the instructor (Gallotti et al., 2017; Hasson and Frith, 2016), particularly when the instructor is conveying information that the learner is expected to acquire.

Examining this interpretation even further, we conducted a Granger causality analysis (GCA) in order to determine whether brain activity recorded from the learner could be predicted by that recorded from the instructor, and to what degree. This analysis revealed that the signal recorded from the instructor's brain better predicted that recorded from the learner's brain than vice versa. Our GCA findings might suggest that while both participants were actively engaged in the interaction, the vocalizations produced by the instructor may reset (or entrain) the phase of learners' neural processes to eventually align them with those of the instructor, thus facilitating IBS. Taken together, these results show that, besides the learner's involvement, the instructor's modeling has great importance during social interactive learning. This speculation is in line with the larger IBS detected during OBSERVATION (vs. IMITATION). Indeed, during the OBSERVATION phase of the task, the instructor was presenting the song to the learner.

The correlational nature of our results does not allow us to make strong conclusions regarding the specific function of this process and/or its causal effects upon learning. Yet, it could be hypothesized that the alignment of neural processes between a learner and an instructor might not only "track" social interactive learning, but also be a sufficient condition for the interpersonal learning process to occur or improve. This hypothesis could be tested in the future by providing causal evidence of the above principle, i.e. by synchronizing the brains of instructors and learners using dual-brain stimulation (cf. Novembre et al., 2017) and monitor its effect upon the learning outcome afterwards.

Last but not the least, it should be noted that the effectiveness of the PL and WL methods may also depend on what is actually learned: e.g. PL could be more beneficial in learning skill-oriented contents (Gobet et al., 2001), whereas WL may be more advantageous in learning theory-oriented contents (Miller, 2005). Future studies might examine the effects of learners' involvement and instructors' modeling on social interactive learning and IBS by comparing different learning contents.

4.2. *The role of the inferior frontal cortex (IFC) during IBS*

The IBS was detected in bilateral inferior frontal cortex (IFC, i.e., channels 10, 14, 31, and 40, Tzourio-Mazoyer et al., 2002). Previous studies have reported IFC activity in the context of interactive tasks entailing communication (Jiang et al., 2012), singing (Osaka et al., 2015), and game playing (Liu et al., 2015). We highlight three possible mechanisms through which IFC might play a crucial role in interactive learning.

First, it is important to emphasize that our task entailed vocal interactions, and that our results were particularly strong during the observation of vocal behavior, somehow in agreement with the notion of vocal interaction being a critical factor affecting learning (Kuhl and Meltzoff, 1997; Locke and Snow, 1997). Because the inferior frontal cortex constitutes a critical "language" hub of the brain, strongly associated with the processing of linguistic structure/syntax in spoken and signed language (Friederici et al., 2003; Homae et al., 2002; Wagner et al., 2001), we could speculate that the observed IBS was mediated by similar IFC activities marking the production and recognition of linguistic structures in both the instructor and the learner. Relatedly to this possibility, it should be mentioned that syntax in language and music share a common set of processes instantiated in frontal areas (Patel, 2003), especially in the left IFG (Levitin and Menon, 2003). From this perspective, learners might have formed representations of the syntactic structure, mostly mediated by auditory perception of the song, which eventually aligned with the ones of the instructor producing it.

A second account of our findings is instead based on evidence suggesting that IFC is also an important hub of the "mirror neuron system" (Iacoboni and Dapretto, 2006; Iacoboni and Mazziotta, 2007). This system has been proposed to facilitate social interactions through "motor resonance" of other's actions and behavior (Gallese, 2003, 2013). A number of studies have shown that IFC is involved with both production and comprehension of both language and action (Cavallo et al., 2015), understanding of others' intention (Hamilton and Grafton, 2008), shared emotional information (Babiloni et al., 2012). From this perspective, IFC might have played a key role in representing the actions produced by the instructor, rather than the syntactic information (as proposed by the previous account).

A final account, potentially integrating the previous two, would suggest that the IFC might parse syntactic information during both auditory perception (as when listening to an utterance) as well as during the perception of a series of movements collectively forming a structure (Novembre and Keller, 2011; Bianco et al., 2016a, 2016b). According to this view, IFC might have played a key role in representing the observed movements using neural resources that are shared across perception and production, and ultimately formed structures that could have helped learners to optimally predict the instructor's behavior and hereby interact successfully (Novembre and Keller, 2014; Sammler et al., 2013; Stephens et al., 2010). This view is consistent with other accounts suggesting that neural processes responsible for action planning are engaged during action observation in order to model others' behaviors (Mattar and Gribble, 2005; Wilson et al., 2004; Mukamel et al., 2010). Certainly, more work is necessary in order to understand the functional significance of the IFC during IBS.

One final detail of our results that might be worth discussing is that IFC activity was observed bilaterally, without clear evidence of lateralization (see Results). This result might be considered alongside the traditional left vs. right hemisphere dominance of language and music, respectively (Jiang et al., 2012; Kimura, 1967; Peretz, 2001; Zatorre et al., 2002). Although some studies do not support the existence of such dichotomy (e.g., Ethofer et al., 2006; Tillmann et al., 2003), the presence of a bilateral activity might be taken as indirect evidence of the language and musical components of this task somehow overlapping in the context of this experiment (Gunji et al., 2007; Osaka et al., 2015; Wan et al., 2010).

4.3. Limitations

We also highlight a few limitations of the present study that should be considered alongside the results we presented. First, the NIRS's optode probe set only covered bilateral fronto-temporo-parietal brain regions, leaving other regions unexplored. This is an important point to make given that human learning is likely governed by several brain regions besides those monitored here (Takeuchi et al., 2017). Secondly, other hyperscanning studies also found IBS in the posterior superior temporal sulcus (STS) during mutual gaze and joint attention (Saito et al., 2010), and in the left temporo-parietal junction (TPJ) during group communication (Jiang et al., 2015). However, we did not observe synchronous brain activity between the instructor and learner in these regions. Given that these temporal/temporo-parietal regions are also known to be crucial for successful social interaction (Dumas et al., 2010; Hasson et al., 2008; Frith and Frith, 2012; Novembre et al., 2016), it is likely that these were interacting with IFC in the context of our task as well. Thus, future studies should also explore whether and how these regions collectively work as a network during social interactive learning. Thirdly, signal contamination due to spontaneous blood flow oscillations (i.e., Mayer waves, ~0.1 Hz) or other global systematic components (e.g., scalp blood flow and changes in blood pressure) are very common issues in NIRS measurements. Here, we addressed this by using a novel and robust principle component analysis (PCA) approach (Zhang et al., 2016). Furthermore, we also utilized the CBSI method to remove physical artifacts (e.g., head motion, jaw muscles, Cui et al., 2010). However, we cannot ascertain whether our preprocessing procedures were sufficient to fully remove these artifacts. A fourth issue relates to the fact that the same instructor participated in all recording sessions. This was chosen in order to minimize inter-instructor variability and optimally control for group differences. However, this comes at the cost of having less independency between the collected samples. Thus, future research should consider recruiting more than one instructor.

5. Conclusions

The present study characterized social interactive learning from a neurobiological perspective. Taking an fNIRS-based hyperscanning approach, we recorded brain activity from learner-instructor dyads during the acquisition of a music song. We observed that both behavioral performance and IBS increased as a function of interactive learning. Specifically, we observed that brain activity in the IFC synchronized across the learner and the instructor, particularly (*i*) when the learner was observing the instructor's vocal behavior and (*ii*) when the learning experience entailed a turn-taking and more active interaction (during PL as opposed to WL). These results suggest that IBS might stand for a neurophysiological characterization of interpersonal learning, increasing as a function of the learner's involvement (Astin, 1984, 1996), and eventually predicting learning performance. These perspectives, which should be investigated further, might open interesting research, clinical and practical perspectives for understanding and potentially treating learning disabilities (Fletcher et al., 2006).

Acknowledgements

We'd like to thank Philippe Peigneux, Cordula Vesper, Yi Zhu, and Nan Zhao for their valuable help and comments on earlier drafts, Qiang Luo and Shuixia Guo for their constructive comments on Granger Causality analyses. This work was sponsored by the National Natural Science Foundation of China (31872783), the China Scholarship Council (201706140082), and the outstanding doctoral dissertation cultivation plan of action of East China Normal University (YB2016011).

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.neuroimage.2018.08.005.

Contribution

Y. P., X. L., and Y. H. designed the experiment. Y. P. and B. S. performed the study. Y. P. analyzed the data. Y. P., G. N., B. S., X. L., and Y. H. wrote the manuscript.

Competing financial interests

The authors declare no competing financial interests.

References

Abrams, D.A., Ryali, S., Chen, T., Chordia, P., Khouzam, A., Levitin, D.J., Menon, V., 2013. Inter-subject synchronization of brain responses during natural music listening. Eur. J. Neurosci. 37 (9), 1458–1469.
Anaya, G., Cole, D.G., 2001. Latina/o student achievement: exploring the influence of student-faculty interactions on college grades. J. Coll. Student Dev. 42 (1), 3–14.
Astin, A.W., 1984. Student involvement: a developmental theory for higher education. J. Coll. Student Person. 25 (4), 297–308.
Astin, A.W., 1996. Involvement in learning revisited: lessons we have learned. J. Coll. Student Dev. 37 (2), 123–134.
Azuma, K., Uchiyama, I., Takano, H., Tanigawa, M., Azuma, M., Bamba, I., Yoshikawa, T., 2013. Changes in cerebral blood flow during olfactory stimulation in patients with multiple chemical sensitivity: a multi-channel near-infrared spectroscopic study. PLoS One 8 (11), e80567.
Babiloni, C., Buffo, P., Vecchio, F., Marzano, N., Del Percio, C., Spada, D., Perani, D., 2012. Brains "in concert": frontal oscillatory alpha rhythms and empathy in professional musicians. Neuroimage 60 (1), 105–116.
Belyk, M., Pfordresher, P.Q., Liotti, M., Brown, S., 2016. The neural basis of vocal pitch imitation in humans. J. Cognit. Neurosci. 28 (4), 621–635.
Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Series B Stat. Methodol. 57, 289–300.
Bianco, R., Novembre, G., Keller, P.E., Kim, S.G., Scharf, F., Friederici, A.D., et al., 2016a. Neural networks for harmonic structure in music perception and action. Neuroimage 142, 454–464.
Bianco, R., Novembre, G., Keller, P.E., Scharf, F., Friederici, A.D., Villringer, A., et al., 2016b. Syntax in action has priority over movement selection in piano playing: an ERP study. J. Cognit. Neurosci. 28 (1), 41–54.
Bjorklund, S.A., Parente, J.M., Sathianathan, D., 2004. Effects of faculty interaction and feedback on gains in student skills. J. Eng. Educ. 93 (2), 153–160.
Burgess, A.P., 2013. On the interpretation of synchronization in EEG hyperscanning studies: a cautionary note. Front. Hum. Neurosci. 7, 881.
Carr, L., Iacoboni, M., Dubeau, M.C., Mazziotta, J.C., Lenzi, G.L., 2003. Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. P. Natl. Acad. Sci. 100 (9), 5497–5502.
Cavallo, A., Lungu, O., Becchio, C., Ansuini, C., Rustichini, A., Fadiga, L., 2015. When gaze opens the channel for communication: integrative role of IFC and MPFC. Neuroimage 119, 63–69.
Chang, C., Glover, G.H., 2010. Time–frequency dynamics of resting-state brain connectivity measured with fMRI. Neuroimage 50 (1), 81–98.
Chen, Y., Bressler, S.L., Ding, M., 2006. Frequency decomposition of conditional granger causality and application to multivariate neural field potential data. J. Neurosci. Meth. 150 (2), 228–237.
Cheng, X., Li, X., Hu, Y., 2015. Synchronous brain activity during cooperative exchange depends on gender of partner: a fNIRS-based hyperscanning study. Hum. Brain Mapp. 36 (6), 2039–2048.
Clark, R.C., Mayer, R.E., 2016. E-learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning. John Wiley & Sons.
Cui, X., Baker, J.M., Liu, N., Reiss, A.L., 2015. Sensitivity of fNIRS measurement to head motion: an applied use of smartphones in the lab. J. Neurosci. Meth. 245, 37–43.
Cui, X., Bray, S., Reiss, A.L., 2010. Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. Neuroimage 49 (4), 3039–3046.
Cui, X., Bryant, D.M., Reiss, A.L., 2012. NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. Neuroimage 59 (3), 2430–2437.

Cui, X., Stetson, C., Montague, P.R., Eagleman, D.M., 2009. Ready… go: amplitude of the fMRI signal encodes expectation of cue arrival time. PLoS Biol. 7 (8), e1000167.
D'Ausilio, A., Novembre, G., Fadiga, L., Keller, P.E., 2015. What can music tell us about social interaction? Trends Cognit. Sci. 19 (3), 111–114.
Decety, J., Jackson, P.L., Sommerville, J.A., Chaminade, T., Meltzoff, A.N., 2004. The neural bases of cooperation and competition: an fMRI investigation. Neuroimage 23 (2), 744–751.
Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., Garnero, L., 2010. Inter-brain synchronization during social interaction. PLoS One 5 (8), e12166.
Edwards-Groves, C., 2017. Teaching and learning as social interaction: salience and relevance in classroom lesson practices. In: Practice Theory Perspectives on Pedagogy and Education. Springer Singapore, pp. 191–213.
Ethofer, T., Anders, S., Erb, M., Herbert, C., Wiethoff, S., Kissler, J., Wildgruber, D., 2006. Cerebral pathways in processing of affective prosody: a dynamic causal modeling study. Neuroimage 30 (2), 580–587.
Fletcher, J.M., Lyon, G.R., Fuchs, L.S., Barnes, M.A., 2006. Learning Disabilities: from Identification to Intervention. Guilford press.
Friederici, A.D., Rueschemeyer, S.A., Hahne, A., Fiebach, C.J., 2003. The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. Cerebr. Cortex 13 (2), 170–177.
Frith, C.D., Frith, U., 2012. Mechanisms of social cognition. Annu. Rev. Psychol. 63 (1), 287–313.
Gallese, V., 2003. The manifold nature of interpersonal relations: the quest for a common mechanism. Philos. T. Roy. Soc. B. 358 (1431), 517–528.
Gallese, V., 2013. Mirror neurons, embodied simulation and a second-person approach to mindreading. Cortex 49 (10), 2954–2956.
Gallotti, M., Fairhurst, M.T., Frith, C.D., 2017. Alignment in social interactions. Conscious. Cognit. 48, 253–261.
Gobet, F., Lane, P.C., Croker, S., Cheng, P.C., Jones, G., Oliver, I., Pine, J.M., 2001. Chunking mechanisms in human learning. Trends Cognit. Sci. 5 (6), 236–243.
Grinsted, A., Moore, J.C., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. Nonlinear Process Geophys. 11 (5/6), 561–566.
Gunji, A., Ishii, R., Chau, W., Kakigi, R., Pantev, C., 2007. Rhythmic brain activities related to singing in humans. Neuroimage 34 (1), 426–434.
Guo, S., Wu, J., Ding, M., Feng, J., 2008. Uncovering interactions in the frequency domain. PLoS Comput. Biol. 4 (5), e1000087.
Hamilton, A.F.D.C., Grafton, S.T., 2008. Action outcomes are represented in human inferior frontoparietal cortex. Cereb. Cortex 18 (5), 1160–1168.
Handberg, R., Lund, M.N., 2014. Automated preparation of Kepler time series of planet hosts for asteroseismic analysis. Mon. Not. Roy. Astron. Soc. 445 (3), 2698–2709.
Hasson, U., Frith, C.D., 2016. Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. Philos. T. Roy. Soc. B. 371 pii: 20150366.
Hasson, U., Furman, O., Clark, D., Dudai, Y., Davachi, L., 2008. Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. Neuron 57 (3), 452–462.
Herrold, R., 2005. New Approaches to Elementary Classroom. Music second.
Hoshi, Y., 2007. Functional near-infrared spectroscopy: current status and future prospects. J. Biomed. Optic. 12 (6), 062106-062106.
Holper, L., Goldin, A.P., Shalóm, D.E., Battro, A.M., Wolf, M., Sigman, M., 2013. The teaching and the learning brain: a cortical hemodynamic marker of instructor–learner interactions in the Socratic dialog. Int. J. Educ. Res. 59, 1–10.
Holper, L., Scholkmann, F., Wolf, M., 2012. Between-brain connectivity during imitation measured by fNIRS. Neuroimage 63 (1), 212–222.
Homae, F., Hashimoto, R., Nakajima, K., Miyashita, Y., Sakai, K.L., 2002. From perception to sentence comprehension: the convergence of auditory and visual information of language in the left inferior frontal cortex. Neuroimage 16 (4), 883–900.
Horaguchi, T., Ogata, Y., Watanabe, N., Yamamoto, M., 2008. Behavioral and near-infrared spectroscopy study of the effects of distance and choice in a number comparison task. Neurosci. Res. 61 (3), 294–301.
Hu, Y., Hu, Y., Li, X., Pan, Y., Cheng, X., 2017. Brain-to-brain synchronization across two persons predicts mutual prosociality. Soc. Cognit. Affect Neurosci. 12 (12), 1835–1844.
Hu, Y., Pan, Y., Shi, X., Cai, Q., Li, X., Cheng, X., 2018. Inter-brain synchrony and cooperation context in interactive decision making. Biol. Psychol. 133, 54–62.
Iacoboni, M., Dapretto, M., 2006. The mirror neuron system and the consequences of its dysfunction. Nat. Rev. Neurosci. 7 (12), 942–951.
Iacoboni, M., Mazziotta, J.C., 2007. Mirror neuron system: basic findings and clinical applications. Ann. Neurol. 62 (3), 213–218.
Ikeda, S., Nozawa, T., Yokoyama, R., Miyazaki, A., Sasaki, Y., Sakaki, K., Kawashima, R., 2017. Steady beat sound facilitates both coordinated group walking and inter-subject neural synchrony. Front. Hum. Neurosci. 11, 147.
Jiang, J., Chen, C., Dai, B., Shi, G., Ding, G., Liu, L., et al., 2015. Leader emergence through interpersonal neural synchronization. Proc. Natl. Acad. Sci. Unit. States Am. 112 (14), 4274–4279.
Jiang, J., Dai, B., Peng, D., Zhu, C., Liu, L., Lu, C., 2012. Neural synchronization during face-to-face communication. J. Neurosci. 32 (45), 16064–16069.
Kahn, C.H., 2013. Plato and the Post-Socratic Dialogue: the Return to the Philosophy of Nature. Cambridge University Press.
Keller, P.E., König, R., Novembre, G., 2017. Simultaneous cooperation and competition in the evolution of musical behavior: sex-related modulations of the singer's formant in human chorusing. Front. Psychol. 8, 1559.
Kimura, D., 1967. Functional asymmetry of the brain in dichotic listening. Cortex 3 (2), 163–178.
Kirchner, M., Schubert, P., Liebherr, M., Haas, C.T., 2014. Detrended fluctuation analysis and adaptive fractal analysis of stride time data in Parkinson's disease: stitching together short gait trials. PLoS One 9 (1), e85787.
Klinger, R., Campbell, P.S., Goolsby, T., 1998. Approaches to children's song acquisition: immersion and phrase-by-phrase. J. Res. Music Educ. 46 (1), 24–34.
Köhler, W., 2015. The Task of Gestalt Psychology. Princeton University Press.
Koike, T., Tanabe, H.C., Okazaki, S., Nakagawa, E., Sasaki, A.T., Shimada, K., Sadato, N., 2016. Neural substrates of shared attention as social memory: a hyperscanning functional magnetic resonance imaging study. Neuroimage 125, 401–412.
Kuhl, P.K., Meltzoff, A.N., 1997. Evolution, nativism, and learning in the development of language and speech. In: Gopnik, M. (Ed.), The Inheritance and Innateness of Grammars. Oxford Univ. Press, New York, pp. 7–44.
Levitin, D.J., Menon, V., 2003. Musical structure is processed in "language" areas of the brain: a possible role for Brodmann Area 47 in temporal coherence. Neuroimage 20 (4), 2142–2152.
Lindenberger, U., Li, S.C., Gruber, W., Müller, V., 2009. Brains swinging in concert: cortical phase synchronization while playing guitar. BMC Neurosci. 10 (1), 1–12.
Liu, T., Saito, H., Oi, M., 2015. Role of the right inferior frontal gyrus in turn-based cooperation and competition: a near-infrared spectroscopy study. Brain Cognit. 99, 17–23.
Liao, Y., Acar, Z.A., Makeig, S., Deak, G., 2015. EEG imaging of toddlers during dyadic turn-taking: mu-rhythm modulation while producing or observing social actions. Neuroimage 112, 52–60.
Light, R.J., 2001. Making the Most of College: Students Speak Their Minds. Harvard University Press, Cambridge, MA.
Locke, J.L., Snow, C., 1997. Social influences on vocal learning in human and nonhuman primates. In: Snowdon, C.T., Hausberger, M. (Eds.), Social Influences on Vocal Development. Cambridge University Press, Cambridge, pp. 274–292.
Lundberg, C.A., Schreiner, L.A., 2004. Quality and frequency of faculty-student interaction as predictors of learning: an analysis by student race/ethnicity. J. Coll. Student Dev. 45 (5), 549–565.
Mandell, J., Schulze, K., Schlaug, G., 2007. Congenital amusia: an auditory-motor feedback disorder? Restor. Neurol. Neurosci. 25, 323–334.
Marchiori, D., Warglien, M., 2008. Predicting human interactive learning by regret-driven neural networks. Science 319 (5866), 1111–1113.
Matsuda, G., Hiraki, K., 2006. Sustained decrease in oxygenated hemoglobin during video games in the dorsal prefrontal cortex: a NIRS study of children. Neuroimage 29 (3), 706–711.
Mattar, A.A., Gribble, P.L., 2005. Motor learning by observing. Neuron 46 (1), 153–160.
Miller, G.A., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol. Rev. 63 (2), 81–97.
Miller, J.P., 2005. Holistic Learning and Spirituality in Education: Breaking New Ground. SUNY Press.
Montague, P.R., Berns, G.S., Cohen, J.D., Mcclure, S.M., Pagnoni, G., Dhamala, M., et al., 2002. Hyperscanning: simultaneous fMRI during linked social interactions. Neuroimage 16 (4), 1159–1164.
Mooney, R., 2009. Neurobiology of song learning. Curr. Opin. Neurobiol. 19 (6), 654–660.
Mukamel, R., Ekstrom, A.D., Kaplan, J., Iacoboni, M., Fried, I., 2010. Single-neuron responses in humans during execution and observation of actions. Curr. Biol. 20 (8), 750–756.
Novembre, G., Keller, P.E., 2011. A grammar of action generates predictions in skilled musicians. Conscious. Cognit. 20 (4), 1232–1243.
Novembre, G., Keller, P.E., 2014. A conceptual review on action-perception coupling in the musicians' brain: what is it good for? Front. Hum. Neurosci. 8, 603.
Novembre, G., Knoblich, G., Dunne, L., Keller, P.E., 2017. Interpersonal synchrony enhanced through 20 Hz phase-coupled dual brain stimulation. Soc. Cognit. Affect Neurosci. 12 (4), 662–670.
Novembre, G., Sammler, D., Keller, P.E., 2016. Neural alpha oscillations index the balance between self-other integration and segregation in real-time joint action. Neuropsychologia 89, 414–425.
Nozawa, T., Sasaki, Y., Sakaki, K., Yokoyama, R., Kawashima, R., 2016. Interpersonal frontopolar neural synchronization in group communication: an exploration toward fNIRS hyperscanning of natural interactions. Neuroimage 133, 484–497.
Nummenmaa, L., Glerean, E., Viinikainen, M., Jääskeläinen, I.P., Hari, R., Sams, M., 2012. Emotions promote social interaction by synchronizing brain activity across individuals. Proc. Natl. Acad. Sci. Unit. States Am. 109 (24), 9599–9604.
Okamoto, M., Dan, H., Sakamoto, K., Takeo, K., Shimizu, K., Kohno, S., Dan, I., 2004. Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10–20 system oriented for transcranial functional brain mapping. Neuroimage 21 (1), 99–111.
Osaka, N., Minamoto, T., Yaoi, K., Azuma, M., Shimada, Y.M., Osaka, M., 2015. How two brains make one synchronized mind in the inferior frontal cortex: fNIRS-based hyperscanning during cooperative singing. Front. Psychol. 6, 1811.
Pan, Y., Cheng, X., Zhang, Z., Li, X., Hu, Y., 2017. Cooperation in lovers: an fNIRS-based hyperscanning study. Hum. Brain Mapp. 38 (2), 831–841.
Patel, A.D., 2003. Language, music, syntax and the brain. Nat. Neurosci. 6 (7), 674–681.
Patel, A.D., 2006. Musical rhythm, linguistic rhythm, and human evolution. Music Percept. 24 (1), 99–104.
Patel, A.D., Iversen, J.R., Bregman, M.R., Schulz, I., 2009. Experimental evidence for synchronization to a musical beat in a nonhuman animal. Curr. Biol. 19 (10), 827–830.
Peretz, I., 2001. Brain specialization for music. Ann. N. Y. Acad. Sci. 930 (1), 153–165.
Persellin, D., Bateman, L., 2009. A comparative study on the effectiveness of two song-teaching methods: holistic vs. phrase-by-phrase. Early Child. Dev. Care 179 (6), 799–806.

Persellin, D., Smith, L., Klein, M., Taguiam, E., 2002. An investigation of the effects of vocal modeling on the development of singing ability in kindergarten children. Texas Music Educ. Res. 20 (1), 2–10.

Quaresima, V., Ferrari, M., 2016. Functional near-infrared spectroscopy (fNIRS) for assessing cerebral cortex function during human behavior in natural/social situations: a concise review. Organ. Res. Meth. 1–23.

Rennie, F., Morrison, T., 2013. E-learning and Social Networking Handbook: Resources for Higher Education. Routledge.

Saito, D.N., Tanabe, H.C., Izuma, K., Hayashi, M.J., Morito, Y., Komeda, H., Sadato, N., 2010. "Stay tuned": inter-individual neural synchronization during mutual gaze and joint attention. Front. Integr. Neurosci. 4, 127.

Sammler, D., Novembre, G., Koelsch, S., Keller, P.E., 2013. Syntax in a pianist's hand: ERP signatures of "embodied" syntax processing in music. Cortex 49 (5), 1325–1339.

Schroeter, M.L., Zysset, S., Kruggel, F., Von Cramon, D.Y., 2003. Age dependency of the hemodynamic response as measured by functional near-infrared spectroscopy. Neuroimage 19 (3), 555–564.

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6 (2), 461–464.

Simmons-Stern, N.R., Budson, A.E., Ally, B.A., 2010. Music as a memory enhancer in patients with Alzheimer's disease. Neuropsychologia 48 (10), 3164–3167.

Singh, A.K., Okamoto, M., Dan, H., Jurcak, V., Dan, I., 2005. Spatial registration of multichannel multi-subject fNIRS data to MNI space without MRI. Neuroimage 27 (4), 842–851.

Stephens, G.J., Silbert, L.J., Hasson, U., 2010. Speaker–listener neural coupling underlies successful communication. Proc. Natl. Acad. Sci. Unit. States Am. 107 (32), 14425–14430.

Takeuchi, N., Mori, T., Suzukamo, Y., Izumi, S.I., 2017. Integration of teaching processes and learning assessment in the prefrontal cortex during a video game teaching–learning task. Front. Psychol. 7, 2052.

Tang, H., Mai, X., Wang, S., Zhu, C., Krueger, F., Liu, C., 2015. Interpersonal brain synchronization in the right temporo-parietal junction during face-to-face economic exchange. Soc. Cognit. Affect Neurosci. 11 (1), 23–32.

Tauber, R.T., 1997. Self-fulfilling Prophecy: a Practical Guide to its Use in Education. Greenwood Publishing Group.

Thepsoonthorn, C., Yokozuka, T., Miura, S., Ogawa, K., Miyake, Y., 2016. Prior knowledge facilitates mutual gaze convergence and head nodding synchrony in face-to-face communication. Sci. Rep. 6, 38261.

Tillmann, B., Janata, P., Bharucha, J.J., 2003. Activation of the inferior frontal cortex in musical priming. Cognit. Brain Res. 16 (2), 145–161.

Tsuzuki, D., Jurcak, V., Singh, A.K., Okamoto, M., Watanabe, E., Dan, I., 2007. Virtual spatial registration of stand-alone fNIRS data to MNI space. Neuroimage 34 (4), 1506–1518.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15 (1), 273–289.

Umbach, P.D., Wawrzynski, M.R., 2005. Faculty do matter: the role of college faculty in student learning and engagement. Res. High. Educ. 46 (2), 153–184.

Verga, L., Kotz, S.A., 2013. How relevant is social interaction in second language learning? Front. Hum. Neurosci. 7, 550.

Verga, L., Kotz, S.A., 2017. Help me if I can't: social interaction effects in adult contextual word learning. Cognition 168, 76–90.

Wagner, A.D., Paré-Blagoev, E.J., Clark, J., Poldrack, R.A., 2001. Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. Neuron 31 (2), 329–338.

Wan, C.Y., Rüüber, T., Hohmann, A., Schlaug, G., 2010. The therapeutic effects of singing in neurological disorders. Music Percept. 27 (4), 287–295.

Werts, C.E., Linn, R.L., Jöreskog, K.G., 1974. Intraclass reliability estimates: testing structural assumptions. Educ. Psychol. Meas. 34 (1), 25–33.

Williamson, R.A., Meltzoff, A.N., Markman, E.M., 2008. Prior experiences and perceived efficacy influence 3-year-olds' imitation. Dev. Psychol. 44 (1), 275–285.

Williamson, R.A., Jaswal, V.K., Meltzoff, A.N., 2010. Learning the rules: observation and imitation of a sorting strategy by 36-month-old children. Dev. Psychol. 46 (1), 57–65.

Wilson, S.M., Saygin, A.P., Sereno, M.I., Iacoboni, M., 2004. Listening to speech activates motor areas involved in speech production. Nat. Neurosci. 7 (7), 701.

Xia, M., Wang, J., He, Y., 2013. BrainNet Viewer: a network visualization tool for human brain connectomics. PLoS One 8 (7), e68910.

Xue, H., Lu, K., Hao, N., 2018. Cooperation makes two less-creative individuals turn into a highly-creative pair. Neuroimage 172, 527–537.

Zatorre, R.J., Belin, P., Penhune, V.B., 2002. Structure and function of auditory cortex: music and speech. Trends Cognit. Sci. 6 (1), 37–46.

Zhang, X., Noah, J.A., Hirsch, J., 2016. Separation of the global and local components in functional near-infrared spectroscopy signals using principal component spatial filtering. Neurophotonics 3 (1), 015004.

Zhou, J., Cole, D., 2017. Comparing international and American students: involvement in college life and overall satisfaction. High Educ. 73 (5), 655–672.

Zhou, Z., Wang, X., Klahr, N.J., Liu, W., Arias, D., Liu, H., Liu, Y., 2011. A conditional Granger causality model approach for group analysis in functional magnetic resonance imaging. Magn. Reson. Imaging 29 (3), 418–433.