

Shared neural representations of syntax during online dyadic communication



Wenda Liu^{a,1}, Holly P. Branigan^{b,1}, Lifen Zheng^a, Yuhang Long^a, Xialu Bai^a, Kanyu Li^a, Hui Zhao^a, Siyuan Zhou^a, Martin J. Pickering^{b,**}, Chunming Lu^{a,c,*}

^a State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, 100875, China

^b Department of Psychology, University of Edinburgh, Edinburgh, EH8 9JZ, United Kingdom

^c IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, 100875, China

ARTICLE INFO

Keywords:

Neural synchronization
Syntax
Shared representation
Communication
fNIRS

ABSTRACT

When people communicate, they come to see the world in a similar way to each other by aligning their mental representations at such levels as syntax. Syntax is an essential feature of human language that distinguishes humans from other non-human animals. However, whether and how communicators share neural representations of syntax is not well understood. Here we addressed this issue by measuring the brain activity of both communicators in a series of dyadic communication contexts, by using functional near-infrared spectroscopy (fNIRS)-based hyperscanning. Two communicators alternatively spoke sentences either with the same or with different syntactic structures. Results showed a significantly higher-level increase of interpersonal neural synchronization (INS) at right posterior superior temporal cortex when communicators produced the same syntactic structures as each other compared to when they produced different syntactic structures. These increases of INS correlated significantly with communication quality. Our findings provide initial evidence for shared neural representations of syntax between communicators.

1. Introduction

When people communicate, they come to see the world in a similar way to each other by aligning their mental representations, for example, concerned with words or meaning (Brennan and Clark, 1996; Garrod and Anderson, 1987). But particularly strong evidence comes from their tendency to use the same syntax as each other (Branigan et al., 2000; Cai et al., 2012; Lu et al., 2001). This behavioral evidence indicates shared syntactic representations between communicators. However, the underlying neural mechanisms are not well understood.

Recent research indicates that communicators synchronize their neural activity when they are involved in real-time communication (Dumas et al., 2010; Jiang et al., 2012). Moreover, the interpersonal neural synchronization (INS) seems to underlie various aspects of communication such as verbal or non-verbal communication, integration of multimodal sensory information, turn-taking, and social engagement, as well as selective processing of target speech in a noisy context (Ahn

et al., 2018; Dai et al., 2018; Dikker et al., 2017; Hirsch et al., 2017; Jiang et al., 2012; Nozawa et al., 2016; Perez-Diaz et al., 2017; Silbert et al., 2014; Stevens et al., 2017). Based on previous behavioral findings that communicators tend to align their syntactic representations (Branigan et al., 2000; Cai et al., 2012; Lu et al., 2001), and previous hyperscanning findings on the relationship between INS and communication, it was hypothesized that a specific pattern of INS might also underlie shared syntactic representations during communication. To localize INS associated with syntax, it is necessary to demonstrate that variations in INS that occur are unambiguously associated with manipulations of syntactic structure rather than sensorimotor properties (i.e., speaking and listening behaviors) or semantics (i.e., meaning of a word or a sentence). In the current study, we achieved this by measuring INS during interactive communication using hyperscanning (Montague and Berns, 2002) while manipulating the prior syntactic context in which utterances were processed (Branigan et al., 2000).

Some research on the single brain suggests that syntactic

* Corresponding author. State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, No. 19 Xijiekouwai Street, Beijing, 100875, PR China.

** Corresponding author. Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, United Kingdom.

E-mail addresses: martin.pickering@ed.ac.uk (M.J. Pickering), luchunming@bnu.edu.cn (C. Lu).

¹ Equal contributions.

representation is exclusively associated with the left hemisphere such as left inferior frontal cortex (IFC) (Atkinson, 2011; Dunn et al., 2011; Pagel et al., 2007) and posterior superior temporal cortex (pSTC) (Friederici et al., 2003, 2006a; Humphries et al., 2006; Papoutsis et al., 2011; Rogalsky and Hickok, 2008; Snijders et al., 2008). However, other evidence suggests that both hemispheres are involved in syntactic representation (Caplan et al., 1996; Linebarger et al., 1983; Schneiderman and Saddy, 1988). One study specifically tested the neural correlates of repeated syntax production by focusing on the single brain, and demonstrated the involvement of not only left IFC and temporal cortices, but also bilateral motor cortices (Segaert et al., 2011). More important, recent hyperscanning research has demonstrated widespread bilateral coupling between speech production and comprehension (Silbert et al., 2014), suggesting that dyadic communication is more likely to be bilaterally distributed. But as far as we know, no studies have examined the neural mechanisms underlying the sharing of syntactic representations between communicators during online dyadic communication.

In this study, we used functional near-infrared spectroscopy (fNIRS)-based hyperscanning, focusing on bilateral IFC, pSTC, and motor cortices. fNIRS is a validated technique that can measure regional changes of hemoglobin concentration in the outer cortex with a spatial resolution of 1–2 cm (Scholkman et al., 2014). It offers considerable benefits over techniques such as fMRI and EEG because it allows research on online dyadic communication (unlike fMRI) alongside a relatively high spatial resolution and good anatomical localization (unlike EEG). fNIRS-based hyperscanning has been successfully used to study dyadic or multi-person communication (Balconi et al., 2018; Cui et al., 2012; Hirsch et al., 2017; Jiang et al., 2012; Lu et al., 2018; Nozawa et al., 2016; Pan et al., 2018).

Specifically, during the current experiment pairs of participants described pictures to each other using one of two possible syntactic structures in a syntactically consistent or inconsistent context. That is, in the syntactically consistent conditions, pairs of participants (participant A and B) alternately produced a completion for a sentence fragment presented below the experimental picture, with the complete sentence always having a double-object (DO) structure (DO condition) or always having a prepositional-object (PO) structure (PO condition) (see Method and materials for example sentences). In the syntactically inconsistent condition, pairs of participants alternately completed sentences with a DO structure and a PO structure (i.e., DO and PO alternated, DP condition). Although this setup was not free communication, it allowed us to test the relationship of INS with syntactic representation while other factors such as sensorimotor properties and semantics were well controlled (for details, see Methods and materials). We predicted that INS that was associated with syntax would be greater when the context was syntactically consistent than when it was syntactically inconsistent. Additionally, we investigated whether such syntactic-related INS increase was affected by integration of multimodal information by examining pairs interacting face-to-face (f2f) or not (Jiang et al., 2012). While f2f communication with eye-contact and back-to-back (b2b) communication modes have been examined previously (Jiang et al., 2012), this study additionally examined a further mode of communication, i.e., f2f without eye-contact. The additional communication mode allowed us to specifically test the roles of eye-contact (f2f with eye-contact vs. f2f without eye-contact) and visual information other than eye-contact (f2f without eye-contact vs. b2b) in dyadic communication. Finally, we investigated whether the effect was associated with left, right, or bilateral IFC/pSTC.

2. Methods and materials

2.1. Participants

One hundred and eighty adults (mean age = 20 years; *S.D.* = 1.6) participated in this study. They were randomly assigned into 90 two-person pairs. In each pair, the members were the same sex (to avoid a potential confound of mixed-sex interactions) (Baker et al., 2016; Daniel

et al., 2011) and were strangers to one another (Aron et al., 1992). All participants were right-handed (Oldfield, 1971), with normal hearing and normal or corrected-to-normal vision, and no language, neurological, or psychiatric disorders. The 90 participant pairs (50 female pairs) were further randomly split into three groups that corresponded to the two syntactically consistent conditions (i.e., DO and PO) and one syntactically inconsistent condition (i.e., DP). During the experiment, 6 pairs (four females and two males) were excluded because of data collection failure, leaving 84 pairs for data analysis (see Table 1 for the final number of pairs in each condition).

Written informed consent was obtained from all participants. The study protocol was approved by the Institutional Review Board of the State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University.

2.2. Experimental materials

The experimental materials were the same as Cai et al. (2012). Specifically, there were 96 experimental pictures, each depicting a ditransitive action that involved an agent, a patient, and a beneficiary. There were 6 different action types (corresponding to 6 different verbs), each associated with 16 experimental pictures. A sentence fragment was presented below each picture that described the contents of the picture (Fig. 1A, B, and C). All experimental pictures were easy to recognize and describe. Additionally, each experimental picture had a corresponding control picture that differed from the experimental picture in one entity.

2.3. Tasks and procedures

In total, there were two syntactically consistent conditions (DO and PO) and one syntactically inconsistent condition (DP). Each condition involved three communication modes. For each condition, the three communication modes were as follows. First, in the *f2f with eye-contact* mode, the two participants sat face-to-face so that they could see each other. Second, in the *f2f without eye-contact* mode, the two participants could see each other but could not make eye-contact. The participants were required to fixate on the screen, which was then confirmed by checking the video recordings of the experiment. Finally, in the *back-to-back (b2b)* mode, the two participants sat back-to-back so that they could not see each other. The sequence of the three communication modes was counterbalanced across participant pairs.

For each communication mode, the two participants in each pair (participants A and B) sat f2f or b2b. A computer screen was placed on a table in front of each participant (Fig. 2A). Each task had two blocks. In one block, the communication started with participant A, whereas in the other block, the communication started with participant B. This sequence was counterbalanced across the participant pairs.

Each block involved 16 pictures that corresponded to one of the 6 actions. For the first block within a communication mode, an initial 15s interval during which the participants did nothing with eyes open was inserted at the beginning of the block to allow the participants to reach a steady state. During this period, both participants' screens remained blank (the data collected during this period were removed during data analyses, see below). An additional 15s interval was inserted at the ending phase of the second block for the same purpose as the initial 15s interval. Then, the experiment began. On the first trial, a picture with a

Table 1
Mean of communication quality in each communication mode of each condition.

	DO (n = 28)	DP (n = 29)	PO (n = 27)
f2f with eye contact	0.951 (0.043)	0.952 (0.039)	0.939 (0.049)
f2f without eye contact	0.941 (0.036)	0.939 (0.043)	0.954 (0.040)
b2b	0.949 (0.039)	0.942 (0.038)	0.938 (0.038)

Note: Italic numbers in the bracket represents standard deviation. “n” indicates the sample size.

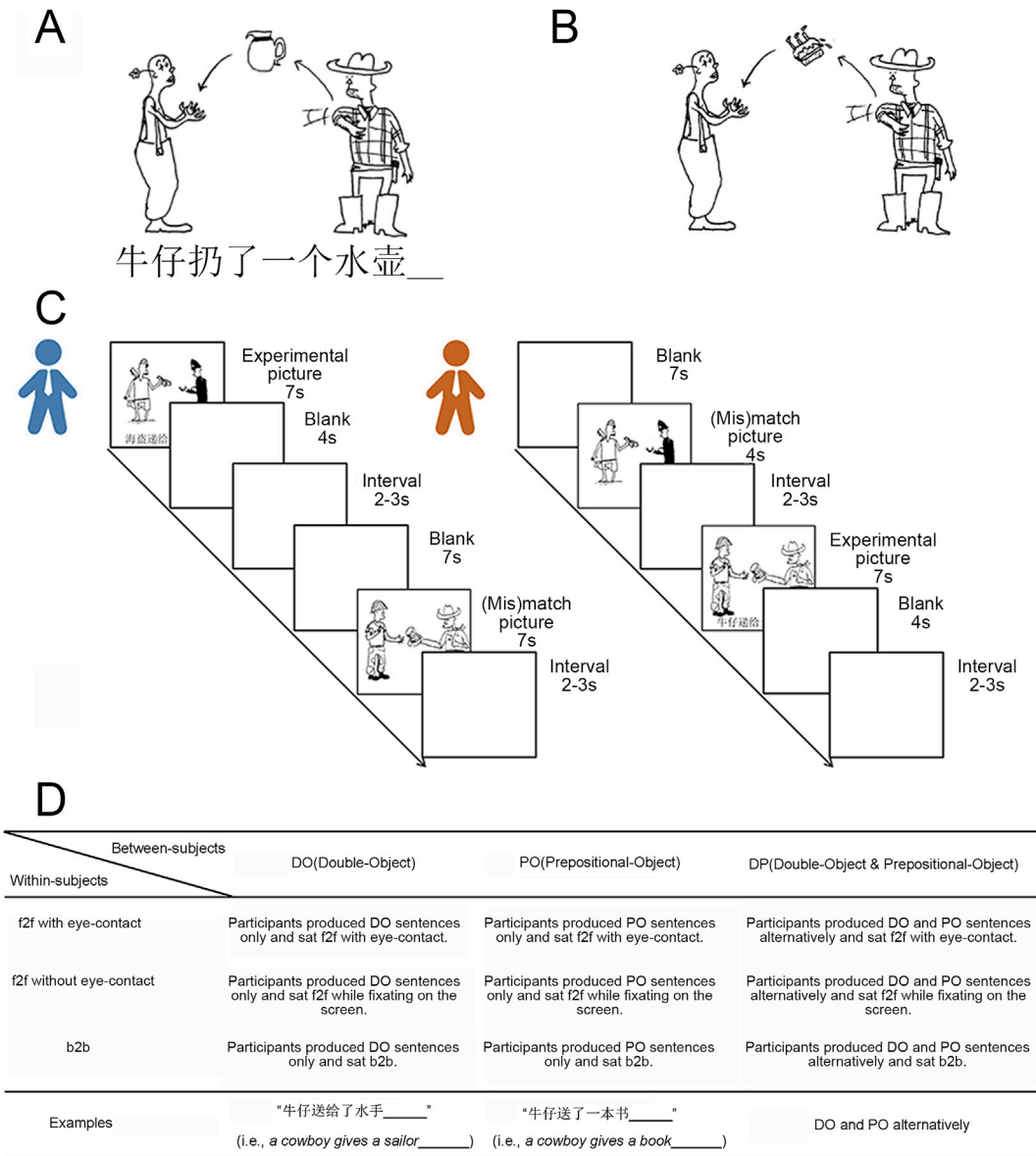


Fig. 1. Experimental materials and procedures. (A) An example of the experimental pictures for the speaker to be described. The sentence fragment in English is “The cowboy throws a jug__”. (B) An example of the control pictures for the comprehender to make judgement whether it matched the picture that was described by the speaker. (C) The experimental procedures for a single trial. The left and right sides are procedures for the speaker and the comprehender respectively. (D) A summary about the design (all conditions and communication modes). Explanations for each mode within each condition are provided, and the corresponding examples are given.

sentence fragment below it appeared on participant A's screen, while participant B saw a blank screen (Fig. 1C). For instance, “牛仔送给了水手” (i.e., a cowboy gives a sailor) was used in the DO condition, whereas “牛仔送了一本书” (i.e., a cowboy gives a book) was used in the PO condition. The sentence fragment and the picture lasted 7s, during which participant A was required to view the picture carefully and then read aloud and complete the sentence fragment to accurately describe the picture (e.g., 牛仔送给了水手一本书, a cowboy gives a sailor a book) (All participants finished this task within 7s). Next, participant A's screen went blank for 4s. During this period, a picture (without a sentence fragment) appeared on participant B's screen. This picture was either the same as (50%) or different from (50%, control picture) the picture described by participant A. Participant B had to decide whether or not the picture that she/he saw matched the description produced by participant A by pressing the button “Yes” or “No”. The sequence of trials requiring “Yes” and “No” responses was randomized. On the next trial, the same procedures were repeated except that participant B produced a picture

description and participant A made a matching decision. The interval between trials was jittered between 2 and 3s (with a blank screen for both participants). The same pictures were used for the DO, PO, and DP conditions; only the sentence fragments varied. In sum, in both the DO and PO conditions, the syntactic structure of the sentences produced by participant A was exactly the same as those produced by participant B. For the DP condition, participants A and B produced sentences with different syntactic structures (Fig. 1D).

2.4. fNIRS data acquisition

During the experiment, participants sat in a quiet room. For each group, an initial resting-state session of 5 min served as a baseline. During this session, the participants were required to keep still with their eyes closed, relax their mind, and remain as motionless as possible (Jiang et al., 2012). The communication sessions immediately followed the resting-state session.

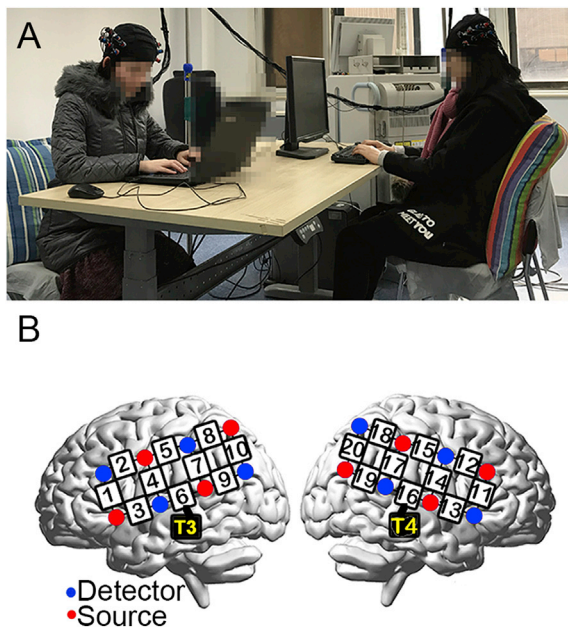


Fig. 2. Experimental setup. (A) Experimental paradigm. Two participants of a pair were seated in a f2f or b2b manner. A computer was placed in front of each participant. (B) fNIRS data acquisition. Customized optodes were placed along Sylvian fissure on both sides of the brain. Each set had 10 measurement channels (CH) that covered bilateral inferior frontal, pre- and post-central, inferior parietal, and superior temporal cortices. CH6 on the left hemisphere and CH16 on the right hemisphere were placed at T3 and T4 respectively according to the international 10–20 system. Measured channels are marked by numbers.

An ETG-4000 optical topography system (Hitachi Medical Company) was used to collect brain functional data from the two participants of each pair simultaneously. Two sets of the same “2 × 4” optode probes were placed along the Sylvian fissure on both sides of the brain (Fig. 2B). Each set had ten measurement channels (CH) that covered bilateral inferior frontal, pre- and post-central, inferior parietal, and superior temporal cortices. CH6 on the left hemisphere and CH16 on the right hemisphere were placed on T3 and T4 respectively according to the international 10–20 system, which was then confirmed by MRI scan on a randomly selected participant. All probe sets were examined and adjusted to ensure consistency of the positions between the two participants of each pair and across the pairs. The absorption of near-infrared light at two wavelengths (695 and 830 nm) was measured with a sampling rate of 10 Hz. The changes in the oxy-hemoglobin (HbO) and deoxy-hemoglobin (HbR) concentrations were recorded in each CH based on the modified Beer–Lambert law.

2.5. Behavioral data analyses

Accuracy for picture-sentence matching was compared between the two participants in each pair using an independent two-sample *t*-test. No significant difference was found ($P > 0.05$). The mean accuracy of the two participants in each pair was then used as an index of communication quality.

To test communication quality across communication modes and syntactic conditions, a two-way mixed ANOVA with a 3 × 3 design was conducted. Syntactic condition (DO, PO and DP) was a between-subjects factor, and communication mode (f2f with eye-contact, f2f without eye-contact, and b2b) was a within-subjects factor.

2.6. fNIRS data analyses

2.6.1. Individual-level analyses

fNIRS data of HbO and HbR concentrations collected during the

resting-state and task sessions were analyzed. During preprocessing, data in the initial and ending interval periods (15s) of each session were removed, leaving 450 s of data for each session. It should be noted that during this step, no filtering or detrending procedures were applied (Cui et al., 2012). Nor did we perform any artifact correction at this level, as wavelet transform coherence (WTC) normalizes the amplitude of the signal according to each time window and thus is not vulnerable to the transient spikes induced by movements (Nozawa et al., 2016). Additional analyses confirmed that our results did not change with and without artifact correction (see the supplementary materials, SM), probably because the probe sets were well-positioned.

Next, a Matlab package was used to perform WTC (Grinstead et al., 2004) in order to assess the cross-correlation between the two fNIRS time series generated by each pair of the participants as a function of frequency and time (Torrence and Compo, 1998). For example, for a specific pair, two time-series of HbO were obtained, one from participant A and the other from participant B. Then, WTC was applied to the two time-series to find regions in the time-frequency space where the two time-series co-varied. This generated a 2-D matrix of the coherence value with both time (column) and frequency (row) information. This analysis was conducted between the same CHs of a pair because shared representations of the same mental process was expected to be associated with INS at the same brain area (Dai et al., 2018; Stolk et al., 2016). Next, the coherence values were time-averaged across the whole communication period, and converted into Fisher *z*-values. These procedures were conducted for each of the communication modes as well as the resting state. According to previous studies (Cui et al., 2012; Jiang et al., 2012), the coherence value increases during the task (i.e., communication) session compared to the resting-state session. Thus, the coherence value from the resting-state session was subtracted from that of the communication session, resulting in an index of INS increase. At this stage, no specific frequency ranges were selected.

2.6.2. Group-level analysis

First, to identify the frequency ranges that were specifically associated with dyadic communication in general, a two-sample *t*-test was conducted between each mode of each condition and the resting-state on the time-averaged coherence value of each CH along the full frequency range (0.01–0.7 Hz, Fig. S1). Data above 0.7 Hz were not included to avoid aliasing of higher frequency physiological noise such as cardiac activity (~0.8–2.5 Hz); data below 0.01 Hz were also not used to remove very low frequency fluctuations; and finally, data within the frequency range of respiratory activity (~0.15–0.3 Hz) were not considered (Guijt et al., 2007; Tong et al., 2011). Frequency ranges were selected based on a center and a range. The center should be a statistically strict threshold that determined the position of the frequency, whereas the range could be a relatively loose threshold that determined the width of the frequency range. In this study the center was set as $P < 0.0005$ whereas the range was $P < 0.05$ (Zheng et al., 2018). The frequency ranges that totally overlapped among modes and conditions were combined, whereas those differing in frequency position or range were considered independently. No further correction for multiple comparisons was applied because this analysis was only used to identify the pattern along the frequency range rather than to obtain the final results.

Second, the coherence values were averaged within each of the selected frequency ranges. Further group-level statistical tests were conducted on the time-averaged and frequency-averaged data. A two-way mixed model ANOVA with a 3 × 3 design was conducted on the INS increase over all CHs, where communication mode (f2f with eye-contact, f2f without eye-contact, and b2b) was a within-subjects factor, and syntactic condition (DO, PO, and DP) was a between-subjects factor.

Results were corrected with an false discovery rate (FDR) method that implemented the Benjamini-Hochberg approach (Benjamini et al., 2006; Benjamini and Yekutieli, 2001) across all CHs ($P < 0.05$). As a general approach to the multiple comparisons problem, an FDR threshold is determined from the observed *P*-value distribution, and hence is adaptive to the amount of signal in the data (Genovese et al., 2002; Nichols and

Hayasaka, 2003). Only the frequency range of 0.02–0.05 Hz showed significantly statistical results (see Results and Fig. S2). Thus, the following analyses were applied to this frequency range only.

2.6.3. Validating the INS increase through a permutation test

To investigate whether the INS increase was specific to pairs of interacting participants, a validation approach was applied. That is, for each communication mode of each condition, all participants were randomly assigned to form new 2-member pairs (i.e., pairs of participants who had been in the same condition but had not communicated with one another), and then the INS increase was re-computed. Next, the INS increase for the DP condition was subtracted from that for the DO or the PO condition respectively. This permutation test was conducted 1000 times to yield normal distributions of the differences between the DO and DP conditions, and between the PO and DP conditions, for each CH which was then compared with the mean value of differences in the original pair of participants. This procedure was applied to all CHs.

2.6.4. Validating the INS increase by excluding the potential contributions of physiological noises to the fNIRS signals

To test whether physiological noises had significantly contributed to the fNIRS signals and thus had affected the syntactic-related INS increase (Kirilina et al., 2012; Tachtsidis and Scholkmann, 2016), the global mean of INS increase across all CHs were introduced as a covariate when performing syntax-by-communication mode ANCOVA. Next, to further test the spatial sensitivity of the syntactic-related INS increase at CH19 (Scholkmann et al., 2014), we introduced the regional mean of INS increase across CHs (CH16, 17, 18, 19 and 20) that were close to CH19 (see Fig. 2 for the positions of these CHs) as a covariate when performing syntax-by-communication mode ANCOVA.

2.6.5. Analyses on communication processes

To test whether the increase of INS was related to the effect of the prior syntactic context (i.e., consistent vs. inconsistent), or to speaking-listening behaviors, each trial was split into two phases: the first 7s (participant A viewed a picture and described the picture aloud, participant B listened to the speech of participant A) and the next 4s (participant A viewed a blank screen and participant B made a “Yes” or “No” decision). The coherence values were then averaged across all trials for each phase after adjusting for the delay-to-peak effect in the fNIRS signals (about 6s). ANOVAs as described above were applied to the averaged coherence values. If the identified INS increase was associated with only the speaking-listening behaviors or both the speaking-listening behaviors and the effect of the prior syntactic context, the two phases would produce different patterns of INS increase. Moreover, no significant INS increase would be found in the next 4s. Alternatively, if the identified INS increase was associated with only the effect of the prior syntactic context, the two phases would produce a similar pattern of INS increase.

2.6.6. Time-lag analyses between the time courses of the speaker and that of the comprehender

To explore whether there was still a significant INS increase when one participant's brain activity preceded that of the other participant (i.e., a time-lag effect, Stephens et al., 2010), the coherence value was recalculated by shifting the time course of one participant forward or backward by 1–6s (step = 1s), respectively. According to previous studies (Dai et al., 2018; Stephens et al., 2010), the time-lag effect in a communication task is usually 1–3s, which can be well covered by our time-lag ranges. Then, a three-way ANOVA was conducted by adding a within-subjects factor of the time-lag (the speaker's brain activity preceded the comprehender's by 1–6s, and vice versa). The other two factors remained syntactic condition and communication mode.

2.7. Correlation between the INS increase and communication quality

To investigate whether the INS increase was related to quality of

communication, the INS increase was correlated with communication quality using the Pearson correlation method across all CHs. For this, the coherence value was averaged across the three communication modes as no significant difference was found among them (see below).

2.8. Data and code availability statement

The data and code are available from the corresponding authors upon reasonable request.

3. Results

3.1. Behavioral results of communication quality

Results indicated a high level (>90%) of communication quality (Table 1). ANOVA did not show any significant effects of syntactic condition or communication mode, nor was there a significant interaction between syntactic condition and communication mode ($P_s > 0.05$). These findings suggest that the three randomly assigned groups did not differ significantly in communication quality, irrespective of the specific communication mode.

3.2. INS associated with shared syntactic representations and communication mode

ANOVA on HbO concentration showed a significant main effect of syntactic condition at right pSTC (CH19, $F(2, 84) = 10.37$, $P < 0.0001$, $\eta^2 = 0.09$) (Fig. 3). Surprisingly, no significant effect was found at any CHs of the left hemisphere, nor IFC (e.g., CH11) on the right hemisphere.

Further post-hoc analyses were conducted to clarify the patterns of differences across the three conditions. For a conservative analysis, comparisons were conducted across all measurement CHs rather than only on CH19, with an FDR correction at $P < 0.05$ level. The INS increase was significantly higher in the DO condition compared to the DP condition ($P = 0.0003$) at CH19. In addition, a significant difference was found between the DO condition and the PO condition at CH19, though both conditions involved syntactic consistency ($P = 0.0008$). No significant difference was found between the PO condition and the DP condition at any CHs, nor were there any other significant differences at any other CHs ($P_s > 0.05$).

Second, a significant main effect of communication mode was found at CH10 that roughly corresponded to left TPJ ($F(2, 84) = 6.95$, $P = 0.001$, $\eta^2 = 0.031$). Pair-wise comparisons showed that the INS increase at CH10 was significantly higher in the f2f with eye-contact mode than in the f2f without eye-contact mode ($P = 0.014$) or the b2b mode ($P = 0.003$). However, no significant difference was found between the f2f without eye-contact mode and the b2b mode ($P > 0.05$). No other significant effects were found at any other CHs ($P_s > 0.05$). No significant interaction between syntactic condition and communication mode was found at CH10 or any other CHs ($P_s > 0.05$). As this result did not appear relevant to syntax, no further analyses were conducted on the INS increase at CH10.

Finally, the data of HbR concentration were also analyzed in order to confirm the findings on HbO concentration. No significant effect of syntax was found for HbR concentration (see SM text and Fig. S3). Thus, no further analyses were conducted the HbR concentration.

3.3. Validating the INS increase through a permutation test

The permutation results showed that the INS increase of the original pairs at pSTC (CH19) was significantly higher than those of the random pairs at $P < 0.01$ level in the DO vs. DP comparison. Thus, the INS increase at pSTC (CH19) was specific to shared representations of syntax in the original pairs who interacted with each other during communication. No significant results were found in the PO vs. DP comparison ($P > 0.05$, Fig. 4).

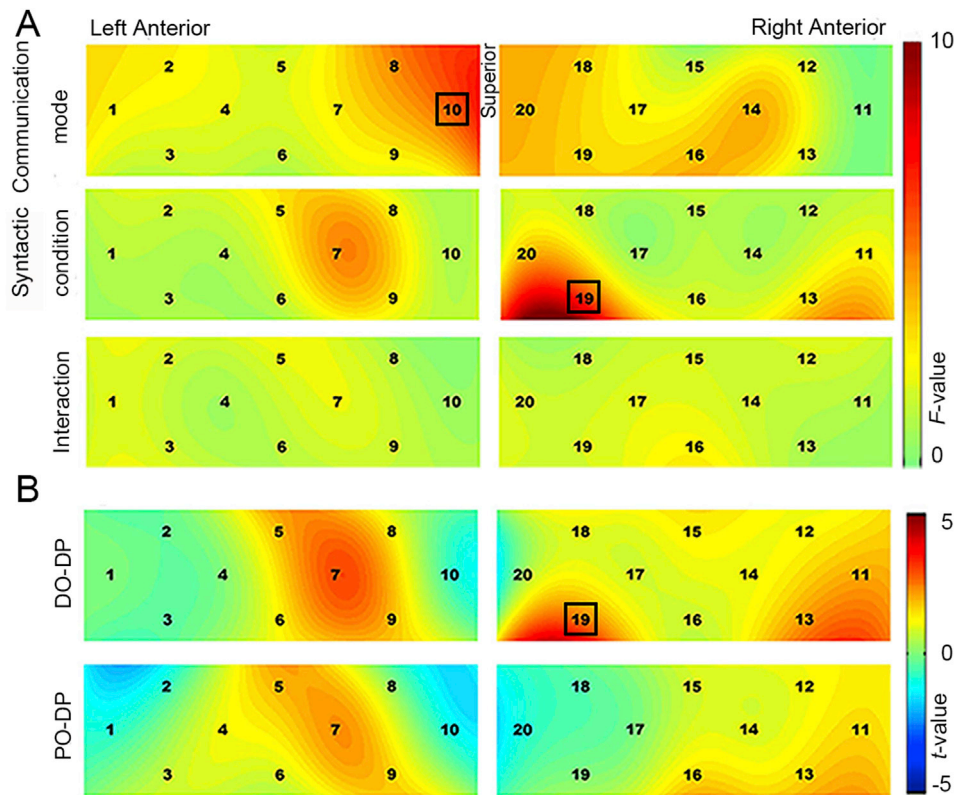


Fig. 3. Results of ANOVA. (A) The main effects and interaction. (B) The syntactically consistent conditions were compared to the syntactically inconsistent condition using post-hoc comparisons. The comparisons were conducted across all CHs rather than only on CH that survived the ANOVA. The numbers represent the measurement channels. Significant results are highlighted using black rectangles.

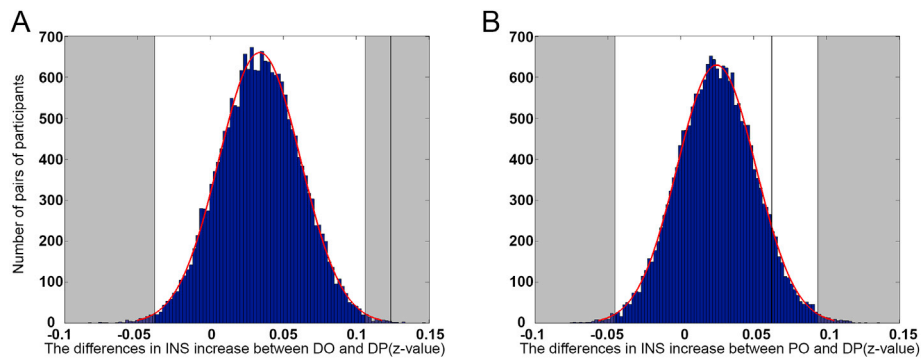


Fig. 4. Results of the permutation test. (A) Distribution of the difference in the INS increase between the DO condition and the DP condition at CH19. The gray areas indicate the top and bottom 1%. The black solid line indicates the position of the original pair's results at CH19. (B) The same as (A) but shows results between the PO condition and the DP condition. The x-axis represents the mean and standard deviation of the distribution, while the y-axis represents number of samples (N).

3.4. Validating the INS increase by excluding the potential contributions of physiological noises to fNIRS signals

When the global mean of INS increase across all CHs were included as a covariate, ANCOVA produced results that were exactly the same as before, i.e., a significant main effect of syntax was found at right pSTC (CH19, $F(2, 84) = 8.62, P = 0.0004, \eta^2 = 0.200$, Fig. 5). Further post-hoc analyses showed that the INS increase was significantly higher in the DO condition than in the DP condition ($P = 0.001$) or in the PO condition ($P = 0.002$). No significant difference was found between the PO condition and the DP condition ($P > 0.05$). No significant main effects of communication mode were found, nor were there significant interactions between syntactic condition and communication mode at any CHs ($P_s > 0.05$).

When the regional mean of INS across CHs that were close to CH19 was included as a covariate, results showed a significant main effect of syntax ($F(2, 84) = 8.85, P = 0.0003, \eta^2 = 0.181$). Further post-hoc analyses showed that the INS increase was significantly higher in the DO condition than in the DP condition ($P = 0.0005$) or in the PO condition ($P = 0.004$), but no significant difference was found between the PO condition and the DP condition ($P > 0.05$). Also, no significant main effect of communication mode was found, nor was there a significant interaction between syntactic condition and communication mode at CH19 ($P_s > 0.05$).

In sum, these results suggested that neither the global nor the regional physiological noises contributed significantly to the syntactic-related INS increase.

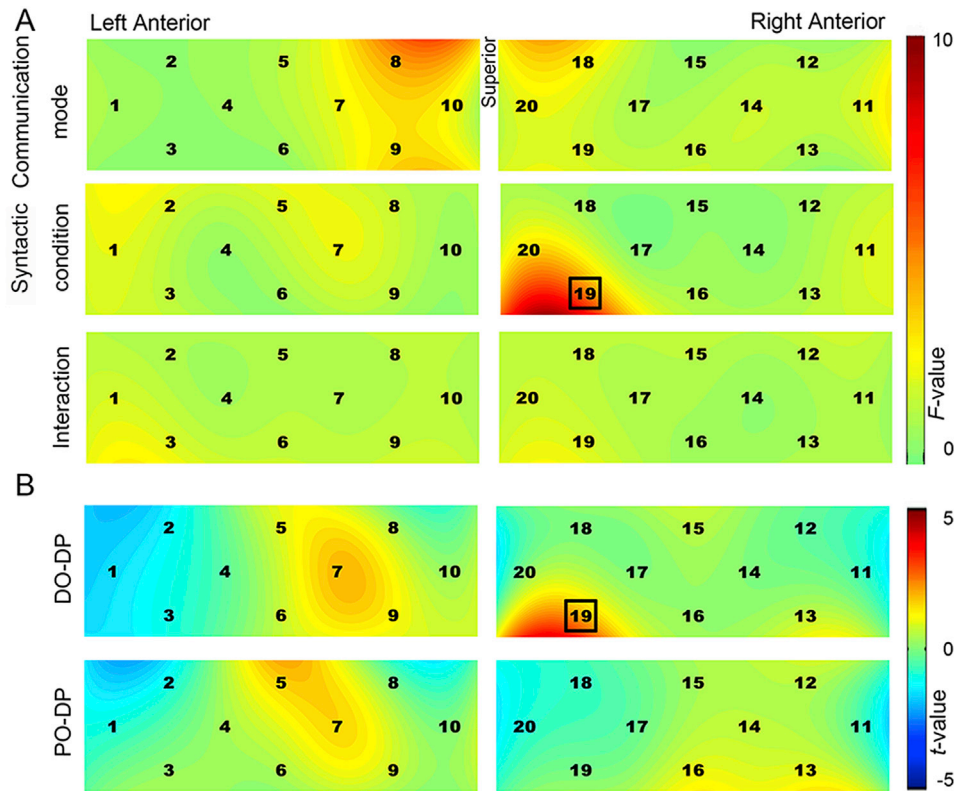


Fig. 5. Results of ANCOVA with the global mean of INS increase as a covariant. (A) The main effects and interaction. (B) The syntactically consistent conditions were compared to the syntactically inconsistent condition using post-hoc comparisons. The comparisons were conducted across all CHs rather than only on CH that survived the ANOVA. The numbers represent the measurement channels. Significant results are highlighted using black rectangles.

3.5. Analyses on communication processes

The INS result for the next 4s (Fig. 6) was very similar to that for the first 7s (Fig. 7), as well as those obtained from the whole time-course. Both for the first 7s and the second 4s, there were main effects of syntactic condition at right pSTC (CH19, the first 7s: $F(2, 84) = 7.86$, $P = 0.0008$, $\eta^2 = 0.078$; the next 4s: $F(2, 84) = 7.96$, $P = 0.0007$, $\eta^2 = 0.022$). Again, no such effect was found in the left hemisphere, nor right IFC ($P_s > 0.05$).

Post-hoc comparisons across all CHs demonstrated that the INS increase at pSTC (CH19) was significantly higher in the DO condition than in the DP condition (the first 7s: $P = 0.001$; the next 4s: $P = 0.001$) or the PO condition (the first 7s: $P = 0.007$; the next 4s: $P = 0.007$). No significant difference was found between the PO condition and the DP condition at pSTC, nor were there any significant differences at other CHs ($P_s > 0.05$).

Neither the first 7s or the next 4s data showed a significant effect of communication mode or a significant interaction between syntactic condition and communication mode ($P_s > 0.05$). These findings confirmed the association between the INS increase at pSTC (CH19) and shared syntactic representations, suggesting an important role of right pSTC in shared neural representations of syntax.

3.6. Time-lag analyses between the time courses of the speaker and that of the comprehender

In this part, we focused on only the 3-way interaction in order to test whether different conditions/modes had any different time-lag effects. Significant 3-way interactions were found in several different CHs that covered left IFC (CH1 and CH3) and pSTC (CH7, CH9, and CH10), and right parietal and sensorimotor cortices (CH15 and CH18). However, further pair-wise comparisons showed that only left TPJ (CH7,

$P = 0.041$) had a significantly higher INS increase in the DO condition than in the DP condition when the speaker's brain activity preceded that of the comprehender by 4s in the f2f without eye-contact mode; there were no significant differences between the DO and the PO conditions, nor were there differences between the PO condition and the DP condition ($P_s > 0.05$). No other significant syntactic effects were found in any other modes or any other time-lags at any other CHs ($P_s > 0.05$).

3.7. Correlation between the INS increase and communication quality

Significant correlations were found between the INS increase at right pSTC (CH19) and communication quality in the DO condition ($r = 0.465$, $P = 0.01$, Pearson correlation, Fig. 8). However, no significant correlations were found either in the PO or in the DP conditions at CH19 ($P_s > 0.05$). Also, no significant correlations were found at CH7 (left TPJ) or at any other CHs in any conditions ($P_s > 0.05$, FDR correction).

4. Discussion

Recent research has suggested that shared representations of syntax between communicators plays a central role in promoting mutual understanding in a dyadic context, but there has been little investigation of the neural mechanism of such representations. This study extended previous studies that focused on only the speaker or the comprehender by examining INS during dyadic communication. Using a manipulation of syntactic context in a picture-description/-matching task, we showed that INS at right pSTC underlies shared representations of syntax, and is closely associated with communication quality. This effect was found when participants consistently produced DO sentences but not PO sentences, which may reflect the lower frequency of DO structures than PO structures in Mandarin (Liu, 2001).

The INS increase was found at right pSTC, but not at left IFC or left

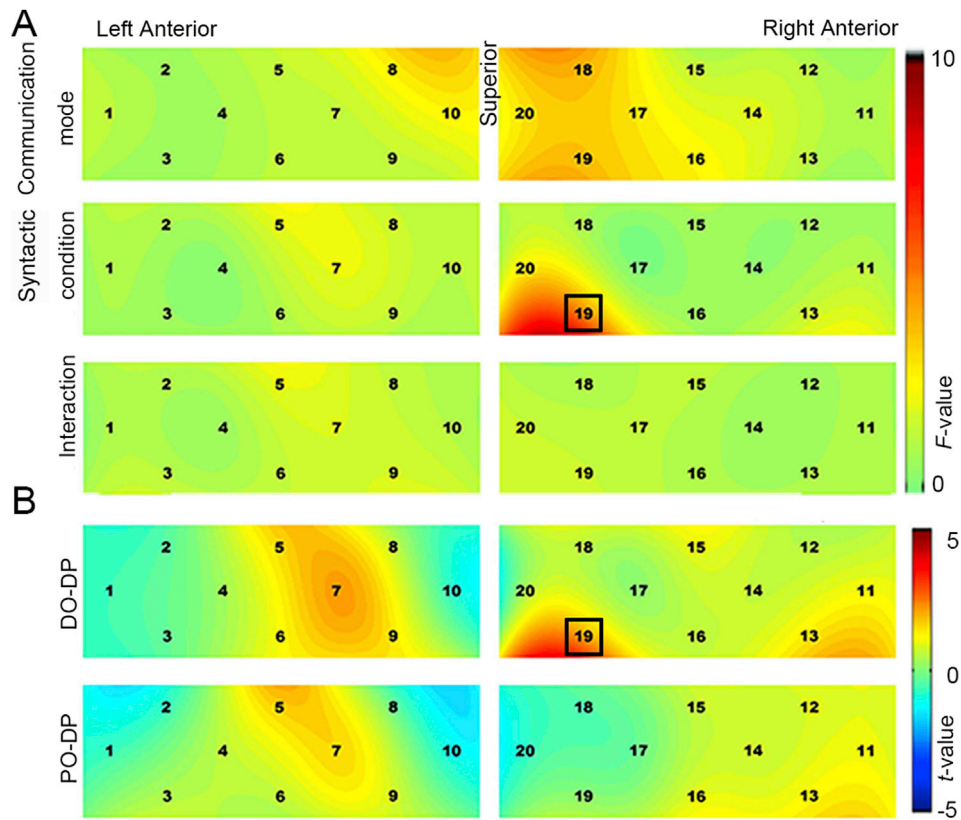


Fig. 6. Results of ANOVA in the next 4 s of a trial. (A) The main effects and interaction. (B) The syntactically consistent conditions were compared to the syntactically inconsistent condition using post-hoc comparisons. The comparisons were conducted across all CHs rather than only on CH that survived the ANOVA. The numbers represent the measurement channels. Significant results are highlighted using black rectangles.

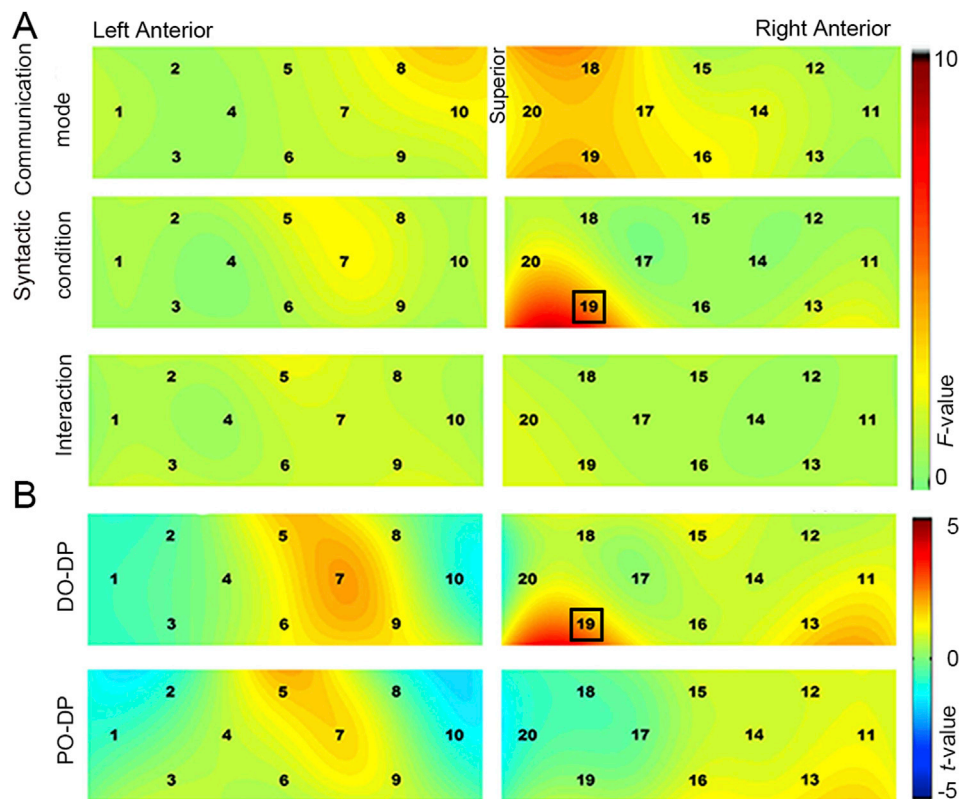


Fig. 7. The same as Fig. 6, but results of ANOVA in the first 7 s of a trial are presented.

- Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* 29, 1165–1188.
- Bornkessel, I., Zysset, S., Friederici, A.D., Von Cramon, D.Y., Schlesewsky, M., 2005. Who did what to whom? The neural basis of argument hierarchies during language comprehension. *Neuroimage* 26, 221–233.
- Branigan, H.P., Pickering, M.J., Cleland, A.A., 2000. Syntactic co-ordination in dialogue. *Cognition* 75, B13–B25.
- Brennan, S.E., Clark, H.H., 1996. Conceptual pacts and lexical choice in conversation. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 1482–1493.
- Cai, Z.G., Pickering, M.J., Branigan, H.P., 2012. Mapping concepts to syntax: evidence from structural priming in Mandarin Chinese. *J. Mem. Lang.* 66, 833–849.
- Caplan, D., Hildebrandt, N., Makris, N., 1996. Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain* 119, 933–949.
- Cui, X., Bryant, D.M., Reiss, A.L., 2012. fNIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *Neuroimage* 59, 2430–2437.
- Dai, B., Chen, C., Long, Y., Zheng, L., Zhao, H., Bai, X., Liu, W., Zhang, Y., Liu, L., Guo, T., 2018. Neural mechanisms for selectively tuning in to the target speaker in a naturalistic noisy situation. *Nat. Commun.* 9, 2405.
- Daniel, B., Li, N.P., Macfarlan, S.J., Mark, V.V., 2011. Sex differences in cooperation: a meta-analytic review of social dilemmas. *Psychol. Bull.* 137, 881–909.
- Dikker, S., Wan, L., Davidesco, I., Kaggan, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., Van Bavel, J.J., Ding, M., 2017. Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Curr. Biol.* 27, 1375–1380.
- Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., Garnero, L., 2010. Inter-brain synchronization during social interaction. *PLoS One* 5, e12166.
- Dunn, M., Greenhill, S.J., Levinson, S.C., Gray, R.D., 2011. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473, 79–82.
- Friederici, A.D., 2002. Towards a neural basis of auditory sentence processing. *Trends Cognit. Sci.* 6, 78–84.
- Friederici, A.D., 2011. The brain basis of language processing: from structure to function. *Physiol. Rev.* 91, 1357–1392.
- Friederici, A.D., Bahlmann, J., Heim, S., Schubotz, R.I., Anwander, A., 2006a. The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proc. Natl. Acad. Sci. U. S. A.* 103, 2458–2463.
- Friederici, A.D., Fiebach, C.J., Schlesewsky, M., Bornkessel, I.D., Von Cramon, D.Y., 2006b. Processing linguistic complexity and grammaticality in the left frontal cortex. *Cerebr. Cortex* 16, 1709–1717.
- 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, 170–177.
- Gagnon, L., Yucel, M.A., Boas, D.A., Cooper, R.J., 2014. Further improvement in reducing superficial contamination in fNIRS using double short separation measurements. *Neuroimage* 85 Pt 1, 127–135.
- Garrod, S., Anderson, A., 1987. Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition* 27, 181–218.
- Ge, J., Peng, G., Lyu, B., Wang, Y., Zhuo, Y., Niu, Z., Tan, L.H., Leff, A.P., Gao, J.-H., 2015. Cross-language differences in the brain network subserving intelligible speech. *Proc. Natl. Acad. Sci. Unit. States Am.* 112, 2972–2977.
- Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15, 870–878.
- 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, 561–566.
- Grodzinsky, Y., Amunts, K., 2006. Broca's Region. Oxford University Press.
- Grodzinsky, Y., Friederici, A.D., 2006. Neuroimaging of syntax and syntactic processing. *Curr. Opin. Neurobiol.* 16, 240–246.
- Guijt, A.M., Sluiter, J.K., Frings-Dresen, M.H.W., 2007. Test-retest reliability of heart rate variability and respiration rate at rest and during light physical activity in normal subjects. *Arch. Med. Res.* 38, 113–120.
- Hamilton, A.F.d.C., 2016. Gazing at me: the importance of social meaning in understanding direct-gaze cues. *Phil. Trans. R. Soc. B* 371, 20150080.
- Hirsch, J., Zhang, X., Noah, J.A., Ono, Y., 2017. Frontal temporal and parietal systems synchronize within and across brains during live eye-to-eye contact. *Neuroimage* 157, 314–330.
- 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, 883–900.
- Humphries, C., Binder, J.R., Medler, D.A., Liebenthal, E., 2006. Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *J. Cogn. Neurosci.* 18, 665–679.
- Humphries, C., Love, T., Swinney, D., Hickok, G., 2005. Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Hum. Brain Mapp.* 26, 128–138.
- Jiang, J., Chen, C., Dai, B., Shi, G., Ding, G., Liu, L., Lu, C., 2015. Leader emergence through interpersonal neural synchronization. *Proc. Natl. Acad. Sci. Unit. States Am.* 112, 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, 16064–16069.
- Khalid, S., Deska, J.C., Hugenberg, K., 2016. The eyes are the windows to the mind: direct eye gaze triggers the ascription of others' minds. *Pers. Soc. Psychol. Bull.* 42, 1666–1677.
- Kirilina, E., Jelzow, A., Heine, A., Niessing, M., Wabnitz, H., Bruhl, R., Ittermann, B., Jacobs, A.M., Tachtsidis, I., 2012. The physiological origin of task-evoked systemic artefacts in functional near infrared spectroscopy. *Neuroimage* 61, 70–81.
- Linebarger, M.C., Schwartz, M.F., Saffran, E.M., 1983. Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition* 13, 361–392.
- Liu, D., 2001. A Typological Study of Giving Type Ditransitive Patterns in Chinese. *Zhong Guo Yu Wen*, pp. 387–479.
- Liu, Y., Piazza, E.A., Simony, E., Shewokis, P.A., Onaral, B., Hasson, U., Ayaz, H., 2017. Measuring speaker–listener neural coupling with functional near infrared spectroscopy. *Sci. Rep.* 7, 43293.
- Lu, C.-C., Bates, E., Hung, D., Tzeng, O., Hsu, J., Tsai, C.-H., Roe, K., 2001. Syntactic priming of nouns and verbs in Chinese. *Lang. Speech* 44, 437–471.
- Lu, K., Xue, H., Nozawa, T., Hao, N., 2018. Cooperation Makes a Group Be More Creative. *Cerebral cortex* bhy215.
- Maguire, E.A., Frith, C.D., 2004. The brain network associated with acquiring semantic knowledge. *Neuroimage* 22, 171–178.
- Montague, P.R., Berns, G.S., 2002. Neural economics and the biological substrates of valuation. *Neuron* 36, 265–284.
- Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S.F., Fazio, F., 2001. Syntax and the brain: disentangling grammar by selective anomalies. *Neuroimage* 13, 110–118.
- Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Büchel, C., Weiller, C., 2003. Broca's area and the language instinct. *Nat. Neurosci.* 6, 774–781.
- Nichols, T., Hayasaka, S., 2003. Controlling the familywise error rate in functional neuroimaging: a comparative review. *Stat. Methods Med. Res.* 12, 419–446.
- 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.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- Pagel, M., Atkinson, Q.D., Meade, A., 2007. Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* 449, 717–720.
- Pan, Y., Novembre, G., Song, B., Li, X., Hu, Y., 2018. Interpersonal synchronization of inferior frontal cortices tracks social interactive learning of a song. *Neuroimage* 183, 280–290.
- Papoutsis, M., Stamatakis, E.A., Griffiths, J., Marslen-Wilson, W.D., Tyler, L.K., 2011. Is left fronto-temporal connectivity essential for syntax? Effective connectivity, tractography and performance in left-hemisphere damaged patients. *Neuroimage* 58, 656–664.
- Perez-Diaz, F., Zillmer, R., Groß, R., 2017. Control of synchronization regimes in networks of mobile interacting agents. *Phys. Rev. Appl.* 7, 054002.
- Pickering, M.J., Garrod, S., 2004. Toward a mechanistic psychology of dialogue. *Behav. Brain Sci.* 27, 169–190.
- Rogalsky, C., Hickok, G., 2008. Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. *Cerebr. Cortex* 19, 786–796.
- Schneiderman, E.L., Saddy, J.D., 1988. A linguistic deficit resulting from right-hemisphere damage. *Brain Lang.* 34, 38–53.
- Scholkmann, F., Kleiser, S., Metz, A.J., Zimmermann, R., Mata Pavia, J., Wolf, U., Wolf, M., 2014. A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *Neuroimage* 85, 6–27.
- Segaert, K., Menenti, L., Weber, K., Petersson, K.M., Hagoort, P., 2011. Shared syntax in language production and language comprehension—an fMRI study. *Cerebr. Cortex* 22, 1662–1670.
- Silbert, L.J., Honey, C.J., Simony, E., Poeppel, D., Hasson, U., 2014. Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proc. Natl. Acad. Sci. U. S. A.* 111, E4687–E4696.
- Snijders, T.M., Vosse, T., Kempen, G., Van Berkum, J.J., Petersson, K.M., Hagoort, P., 2008. Retrieval and unification of syntactic structure in sentence comprehension: an fMRI study using word-category ambiguity. *Cerebr. Cortex* 19, 1493–1503.
- Stephens, G.J., Silbert, L.J., Hasson, U., 2010. Speaker-listener neural coupling underlies successful communication. *Proc. Natl. Acad. Sci. U. S. A.* 107, 14425–14430.
- Stevens, R., Galloway, T., Lamb, J., Steed, R., Lamb, C., 2017. Linking Team Neurodynamic Organizations with Observational Ratings of Team Performance. *Innovative Assessment of Collaboration*. Springer, pp. 315–330.
- Stolk, A., Verhagen, L., Toni, I., 2016. Conceptual alignment: how brains achieve mutual understanding. *Trends Cognit. Sci.* 20, 180–191.
- Stowe, L.A., Broere, C.A., Paans, A.M., Wijers, A.A., Mulder, G., Vaalburg, W., Zwartz, F., 1998. Localizing components of a complex task: sentence processing and working memory. *Neuroreport* 9, 2995–2999.
- Tachtsidis, I., Scholkmann, F., 2016. False positives and false negatives in functional near-infrared spectroscopy: issues, challenges, and the way forward. *Neurophotonics* 3, 030401.
- Tong, Y., Lindsey, K.P., Frederick, B.d., 2011. Partitioning of physiological noise signals in the brain with concurrent near-infrared spectroscopy and fMRI. *J. Cerebr. Blood Flow Metab.* 31, 2352–2362.
- Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. *Bull. Am. Meteorol. Soc.* 79, 61–78.
- Vandenbergh, R., Nobre, A.C., Price, C., 2002. The response of left temporal cortex to sentences. *J. Cogn. Neurosci.* 14, 550–560.
- Wirth, J.H., Sacco, D.F., Hugenberg, K., Williams, K.D., 2010. Eye gaze as relational evaluation: averted eye gaze leads to feelings of ostracism and relational devaluation. *Pers. Soc. Psychol. Bull.* 36, 869–882.
- Zaccarella, E., Meyer, L., Makuuchi, M., Friederici, A.D., 2017a. Building by syntax: the neural basis of minimal linguistic structures. *Cerebr. Cortex* 27, 411–421.
- Zaccarella, E., Schell, M., Friederici, A.D., 2017b. Reviewing the functional basis of the syntactic Merge mechanism for language: a coordinate-based activation likelihood estimation meta-analysis. *Neurosci. Biobehav. Rev.* 80, 646–656.
- Zheng, L., Chen, C., Liu, W., Long, Y., Zhao, H., Bai, X., Zhang, Z., Han, Z., Liu, L., Guo, T., 2018. Enhancement of teaching outcome through neural prediction of the students' knowledge state. *Hum. Brain Mapp.* 39, 3046–3057.