

A cross-brain neural mechanism for human-to-human verbal communication

Joy Hirsch,^{1,2,3,4,6} J. Adam Noah,¹ Xian Zhang,¹ Swethasri Dravida,^{4,5} and Yumie Ono^{1,7}

¹Department of Psychiatry, Yale School of Medicine, New Haven, CT 06511, USA, ²Department of Neuroscience, Yale School of Medicine, New Haven, CT 06511, USA, ³Department of Comparative Medicine, Yale School of Medicine, New Haven, CT 06511, USA, ⁴Interdepartmental Neuroscience Program, Yale School of Medicine, New Haven, CT 06511, USA, ⁵Medical Student Training Program, Yale School of Medicine, New Haven, CT 06511, USA, ⁶Department of Medical Physics and Biomedical Engineering, University College London, London, UK and ⁷Department of Electronics and Bioinformatics, School of Science and Technology, Meiji University, Kawasaki, Kanagawa, Japan.

Correspondence should be addressed to Joy Hirsch, Brain Function Laboratory, Yale School of Medicine, 300 George St., Ste. 902 New Haven, CT 06511, USA. E-mail: joy.hirsch@yale.edu.

Abstract

Neural mechanisms that mediate dynamic social interactions remain understudied despite their evolutionary significance. The interactive brain hypothesis proposes that interactive social cues are processed by dedicated brain substrates and provides a general theoretical framework for investigating the underlying neural mechanisms of social interaction. We test the specific case of this hypothesis proposing that canonical language areas are upregulated and dynamically coupled across brains during social interactions based on talking and listening. Functional near-infrared spectroscopy (fNIRS) was employed to acquire simultaneous deoxyhemoglobin (deOxyHb) signals of the brain on partners who alternated between speaking and listening while doing an Object Naming & Description task with and without interaction in a natural setting. Comparison of interactive and non-interactive conditions confirmed an increase in neural activity associated with Wernicke's area including the superior temporal gyrus (STG) during interaction ($P = 0.04$). However, the hypothesis was not supported for Broca's area. Cross-brain coherence determined by wavelet analyses of signals originating from the STG and the subcentral area was greater during interaction than non-interaction ($P < 0.01$). In support of the interactive brain hypothesis these findings suggest a dynamically coupled cross-brain neural mechanism dedicated to pathways that share interpersonal information.

Key words: two-person neuroscience; coupled dynamics; fNIRS; hyperscanning; cross-brain interaction

Introduction

Communication based on natural spoken language and interpersonal interaction is a foundational component of social behavior; however, theoretical representations of the neural underpinnings associated with communicating individuals

remain in the early stages (Hasson *et al.*, 2012; García and Ibáñez, 2014; Hasson and Frith, 2016). Investigation of dynamic social interactions between two individuals extends the fundamental unit of behavior from a single-brain to a two-brain unit, the dyad. For example, canonical human language models are based on single brains and consist of specialized within-brain units for

Received: 28 March 2018; Revised: 26 July 2018; Accepted: 8 August 2018

© The Author(s) 2018. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

language functions such as production of speech (Broca's region), reception/comprehension of auditory signals (Wernicke's region) as well as systems associated with high-level cognitive and linguistic functions (Gabrieli et al., 1998; Binder et al., 2000; Price, 2012). However, understanding how these single within-brain subsystems mediate the rapid and dynamic exchanges of information during live verbal interactions between dyads is an emerging area of investigation. Two-brain studies during natural conversation are challenging not only because it is necessary to simultaneously record synchronized and spontaneous neural activity within the two-brain unit but also because the communicating partners are engaged in different tasks: one talking and the other listening. These joint and non-symmetrical functions within a dyad include complementary transient and adaptive responses as opposed to mirroring or imitation and thus extend the computational complexity of these investigations.

It has been suggested that coupling between neural responses of the speaker and listener represent mutual information transfer functions (Dumas et al., 2014), and these functions implement neural adaptations that dynamically optimize information sharing (Hasson and Frith, 2016). The complexities of synchronized neural activity have been recognized in previous studies where participants recited and subsequently heard the same story (Hasson et al., 2004). A hierarchy of common activation for the compound epoch with both talking and listening functions was associated with multiple levels of perceptual and cognitive processes. These various levels of abstraction were assumed to operate in parallel with distinguishable time scales of representation (Hasson et al., 2008), and it has been proposed that live communication between dyads includes coupling of rapidly exchanged signals between these various levels of representation (Hasson and Frith, 2016).

However, investigation of these putative underlying fine-grained interactive neural processes challenges conventional imaging techniques. The knowledge gap between static and dynamic processes is, in part, a consequence of neuroimaging methods that are generally restricted to single individuals, static tasks and non-verbal responses. Understanding of neural processes that underlie dynamic coupling between individuals engaged in interactive tasks requires the development of novel experimental paradigms, technology and computational methods. These challenges are largely addressed using functional near-infrared spectroscopy (fNIRS) that enables simultaneous acquisitions of brain activity-related signals from two naturally interacting and verbally communicating individuals. fNIRS is based on changes in spectral absorbance of both oxyhemoglobin (OxyHb) and deoxyhemoglobin (deOxyHb) detected by surface-mounted optodes (Villringer and Chance, 1997; Strangman et al., 2002; Cui et al., 2011; Ferrari and Quaresima, 2012; Scholkmann et al., 2014). These hemodynamic signals serve as a proxy for neural activity similar to hemodynamic signals acquired by functional magnetic resonance imaging (fMRI) (Ogawa et al., 1990; Ferrari and Quaresima, 2012; Boas et al., 2014). fNIRS is well suited for simultaneous neuroimaging of two partners because head-mounted signal detectors and emitters are tolerant of limited head movement, and as in fMRI, signal sources are registered to standard brain coordinates.

Human language systems are typically investigated by fMRI and employ non-communicative internal thought processes (covert speech) rather than actual (overt) speaking because of the deleterious effects of head movement in the scanner. Although in some cases actual speaking has been achieved during fMRI (Gracco et al., 2005; Stephens et al., 2010), these studies do

not capture speaking processes as they occur in actual live interactive dialogue with another person. Neural systems engaged by actual speaking have, however, been validated using fNIRS (Scholkmann et al., 2013b; Zhang et al., 2017) and further extend the technological advantages of fNIRS to the investigation of neural systems that underlie the neurobiology of live verbal communication and social interaction (Babiloni and Astolfi, 2014; García and Ibáñez, 2014; Schilbach, 2014; Pinti et al., 2015; Hirsch et al., 2017).

Cross-brain neural coherence of hemodynamic signals originating from fNIRS is established as an objective indicator of synchrony between two individuals for a wide variety of tasks performed jointly by dyads. Representative examples include: coordinated button pressing (Funane et al., 2011), coordinated singing and humming (Osaka et al., 2015), gestural communication (Schippers et al., 2010), cooperative memory tasks (Dommer et al., 2012), and face-to-face unstructured dialogue (Jiang et al., 2012). Similarly, neural synchrony across brains has been shown to index levels of interpersonal interaction, including cooperative and competitive game playing (Cui et al., 2012; Liu et al., 2016; Tang et al., 2016; Piva et al., 2017), imitation (Holper et al., 2012), coordination of speech rhythms (Kawasaki et al., 2013), leading and following (Jiang et al., 2015), group creativity (Xue et al., 2018), and social connectedness among intimate partners using electroencephalography (EEG) (Kinreich et al., 2017), and fNIRS (Pan et al., 2017). Cross-brain synchrony has also been shown to increase during real eye-to-eye contact as compared to mutual gaze at the eyes of a static face picture (Hirsch et al., 2017). Neural coupling of fNIRS signals across the brains of speakers and listeners who separately recited narratives and subsequently listened to the passages were found to be consistent with comprehension of verbally transferred information (Liu et al., 2017). These findings also replicate previous findings using fMRI (Stephens et al., 2010). All contribute to the advancing theoretical framework for two-person neuroscience (Cui et al., 2012; Konvalinka and Roepstorff, 2012; Schilbach et al., 2013; Scholkmann et al., 2013a) and to the emerging proposition that neural coupling between partners underlies mechanisms for reciprocal interactions that mediate the transfer of verbal and non-verbal information between dyads (Saito et al., 2010; Tanabe et al., 2012; Koike et al., 2016).

A recently proposed Interactive Brain Hypothesis provides a general framework by suggesting that natural interactions engage neural processes not engaged without the interaction (Di Paolo and De Jaegher, 2012; De Jaegher et al., 2016). This hypothesis has been previously tested in studies of communicative pointing. Communicative pointing is a human-specific gesture that is intended to share information about a visual item with another person. In one study, neural correlates of pointing were examined with and without communicative intent using positron emission tomography, PET, and showed that pointing when communicating activated the right posterior superior temporal sulcus and right medial prefrontal cortex in contrast to pointing without communication (De Langavant et al., 2011). Further, in addition to augmented activity during communicative gestures, an EEG study found that resonance between cross-brain signals in the posterior mirror neuron system of an observer continuously followed subtle temporal changes in activity of the same region of the sender (Schippers et al., 2010). This fine-grained temporal interplay between cross-brain regions is consistent with both coordinated motor planning functions and synchronized mentalizing during interpersonal communications based on gestures.

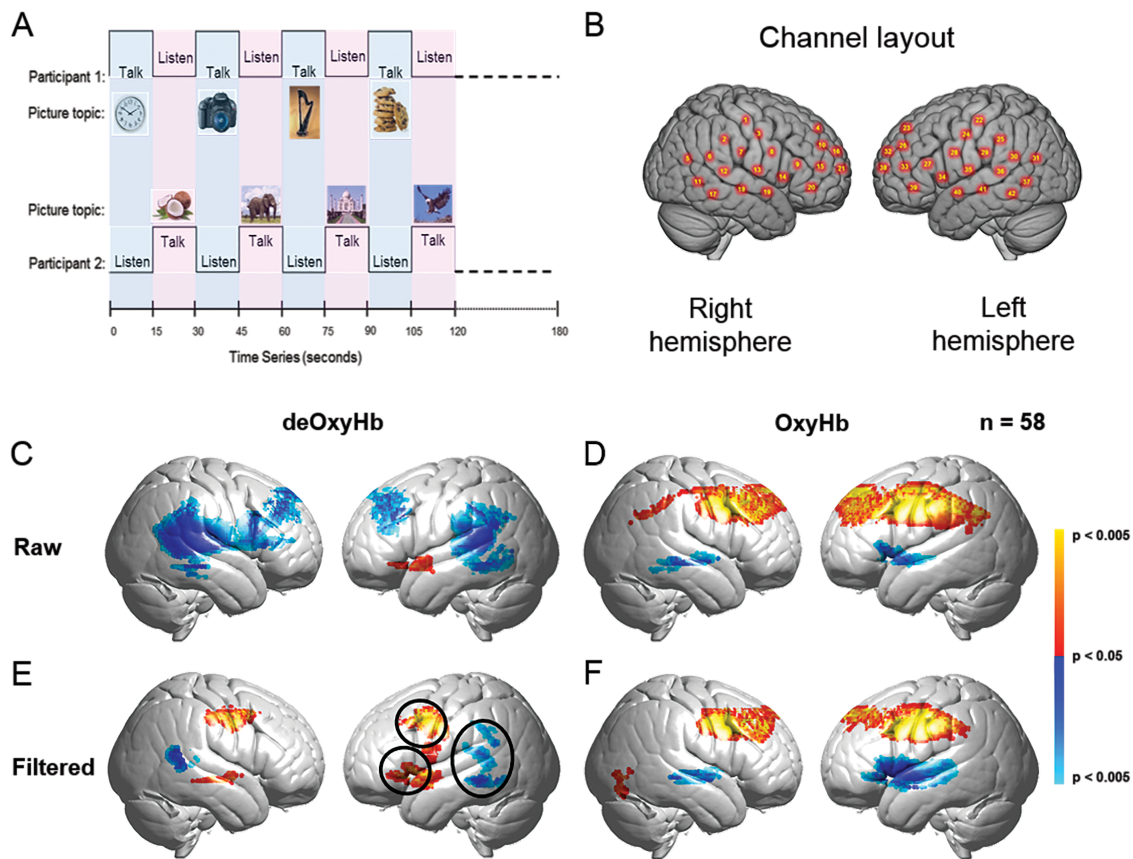


Fig. 1. (A) Experimental paradigm for the two-person Object Naming & Description task. (B) Right and left hemispheres of a single rendered brain illustrate average locations (red circles) for the 42 channels identified by number. See [Table S1](#) for group coordinates. (C–F) Signal selection was based on empirical comparison of voxel-wise contrasts from deOxyHb (left column) and OxyHb (right column) signals without global mean removal (raw, middle row) and with global mean removal using a high pass spatial filter (filtered, bottom row). Red/yellow indicates [talking > listening], and blue/cyan indicates [listening > talking] with levels of significance indicated by the color bar on the right. Images include left and right group sagittal views. The three circles shown in panel E (left hemisphere) represent the canonical language regions: Broca's area (anterior), Wernicke's area (posterior) and articulatory motor system (central). These regions are observed for the deOxyHb signals following global mean removal (Zhang et al., 2016, 2017), but not for the other signal processing approaches, as illustrated in panels C, D and F. This empirical approach supports the decision to use the deOxyHb signal for this study ($n = 58$).

Neural systems engaged during live speaking and listening with intent to communicate are not conventionally differentiated from neural systems engaged during static language functions. However, dyadic observations enable a direct test of the interactive brain hypothesis during verbal communication by isolating the effects of interaction using both contrast comparisons as well as cross-brain coupling methods. Specifically, in this study we predict increases in within-brain neural activity for regions classically associated with language-related functions and increases in cross-brain coherence during live interactions involving dynamic speaking and listening.

Experimental methods

The task and stimuli

The verbal task was based on the well-established Object Naming & Description task frequently employed for clinical applications using fMRI where mapping of the human language system is the goal for neurosurgical planning purposes. In fMRI the task is conventionally performed as covert (silent) speech due to the deleterious effects of head movement in the scanner. The patient is asked to imagine producing the speech related to naming and describing pictured objects. Neural regions activated by this task have been validated by intraoperative recordings of

Broca's and Wernicke's areas (Hirsch et al., 2000; Hart et al., 2007). In Broca's area, speech is disrupted by the intraoperative stimulation. In Wernicke's area, paraphasic errors are observed for naming of objects. These well-defined and objective procedures routinely serve as functional maps to guide these 'standard of care' neurosurgical procedures. This prior use and validation of the Object Naming & Description task guided the task selection and development for this two-person interactive study. Stimulus pictures employed common and unrelated objects selected for clarity and familiarity.

Experimental paradigm

The time series for the two-person Object Naming & Description task is shown in [Figure 1A](#). Participant 'speaker' and 'listener' roles switched every 15 s when a new picture was automatically presented and viewed by each participant. This exchange between talking and listening continued for 3 min and was repeated twice. Participants were occluded from each other to prevent face-to-face interactions thought to confound the language-related interaction. The onset of each block was cued by the appearance of a new picture viewed by both participants. In the control condition, i.e. no interaction, 'monologue', the speaker named and described the picture. Listener and speaker

roles switched with each epoch. The interactive task was identical to the monologue task except that each speaker responded to the comments of the previous speaker before describing the new picture. Participants were instructed to change topics from 'your partner's comments about the previous picture' to 'my comments about the new picture' near the middle of the epoch. The exact time of the topic switch was not specified to assure that communication flowed as naturally as possible.

Participants

Sixty-two healthy adult participants (31 dyads) were enrolled in the study: mean age = 24 ± 6.1 , and range, 18–42 years; 57% female, 97% right-handed (Oldfield, 1971). Twenty-seven dyads provided usable data and four dyads were excluded from coherence analyses as the fNIRS signals from one partner in each pair were unusable because of excessive head movement. However, the four participants without partners and with usable data were included in single-brain analyses ($n = 58$). No individual participated in more than one dyad. Participants were screened prior to the experiment to determine eligibility for the study by performing two tasks: a right-handed finger-thumb-tapping task and passive viewing of a reversing checkerboard during acquisition of fNIRS signals. These tasks were selected because the response patterns are well defined and served as fiducial markers for signal evaluation. In these experiments, active and rest epochs were alternated in 15 s blocks for two 3-min runs for each task. If counter correlated deOxyHb and OxyHb signals were observed in the left motor hand area for the finger-thumb-tapping task ($P < 0.05$) and in the bilateral occipital lobe for the passive viewing task ($P < 0.05$), then the participant was eligible for the two-person speaking and listening experiment. Approximately 30% of possible participants were not eligible by these criteria and did not participate in the experiment. This attrition rate is common for adult subjects, as unknowable factors, such as skull thickness, fat deposits, bone density and blood chemistry are known to affect fNIRS signal strength (Owen-Reece et al., 1999; Okada and Delpy, 2003; Cui et al., 2011). Participant pairs were assigned in order of recruitment and were either strangers prior to the experiment or casually acquainted as classmates in order to prevent possible confounds due to the putative effects of affiliation. Twelve pairs were mixed gender, 10 were female–female and 5 were male–male. Power limitations prevented comparisons of dyad gender types. All participants provided written informed consent in accordance with guidelines approved by the Yale University Human Investigation Committee (HIC #1501015178) and were reimbursed for their participation.

Experimental setup

Participants were positioned ~140 cm across a table from each other and were separated by an occluder (91×122 cm white cardboard) for half of the runs and by two 24-inch monitors with a 16×9 aspect ratio (Asus VG248QE) for the other half of the runs. In both cases, participants did not have a view of their partner. Comparisons between the two types of occluders provided no evidence for differences, and conditions were combined for data analysis. The experimental room was dimly lit with fluorescent light. Windows were covered to prevent stray light in the room, and 'dark values' recorded for each channel on all experiments confirmed the absence of stray light. The experimental room was dedicated to the fNIRS system, and adjacent laboratory noise was minimized in order to protect a

distraction-free experimental environment. Experiments were generally performed during similar time intervals ranging from early to mid-afternoon in order to reduce sources of variance that might arise from chronobiological factors related to human physiology. Video and audio recordings were acquired on all sessions and confirmed compliance with instructions. Audio recordings also confirmed the absence of evidence for a difference in the number of words spoken during the monologue (no interaction condition) and the interactive conditions.

Signal acquisition

Hemodynamic signals were acquired using a 64-fiber (84-channel) continuous-wave fNIRS system (Shimadzu LABNIRS, Kyoto, Japan) configured for hyperscanning of two participants. Figure 1B illustrates the distribution of 42 channels over both hemispheres, and this was the same for both participants. Channel separations were adjusted by individual differences in head size and were either 2.75 cm for small heads or 3.0 cm for large heads (Dravida et al., 2017). This assured that emitters and detectors were optimally placed on the scalp of each subject regardless of head size. Hemodynamic responses that are based on the transportation of light through tissue also depend upon factors including the scattering nature of the tissue, wavelength of the light, age of the subject and skull thickness. The accuracy of measured changes has been shown to vary with models of these effects, referred to as differential pathlength factors (DPFs) (Scholkmann et al., 2013c). Although DPFs are most commonly applied to frontal lobe structures, variations in DPF models have also been shown to improve sensitivity in somatosensory, motor and occipital regions (Zhao et al., 2002). However, given the extended head coverage of this study with optodes located in regions without prior standardization of DPFs, these data were not adjusted for DPF.

Temporal resolution for signal acquisition was 27 ms. In the LABNIRS system, three wavelengths of light (780, 805, and 830 nm) are delivered by each emitter. Each detector measures the absorbance for these wavelengths, which were selected by the manufacturer for differential absorbance properties related to blood oxygenation levels. Using the three wavelengths together, absorption is converted to concentration changes for deOxyHb, OxyHb, and total combined deOxyHb and OxyHb. Conversion of absorbance measures to concentration have been previously described (Matcher, 1995).

Optode localization

Anatomical locations of optodes were determined for each participant in relation to standard head landmarks (inion, nasion, top center (Cz), and left and right tragi) using a Patriot 3D Digitizer (Polhemus, Colchester, VT) and linear transform techniques (Okamoto and Dan, 2005; Eggebrecht et al., 2012; Ferradal et al., 2014). Montreal Neurological Institute (MNI) coordinates for the channels were obtained using NIRS-SPM (Ye et al., 2009) with MATLAB (Mathworks, Natick, MA), and corresponding anatomical locations of each channel were determined. See Figure 1B and supplementary Table S1 for median group channel centroids. Channels were further automatically clustered into anatomical regions based on shared anatomy, and this grouping was employed for cross-brain coherence analyses. The average number of channels in each region was 1.68 ± 0.70 . Grouping was achieved by automatic identification of 12 bilateral regions of interest (ROIs) from the acquired channels, including (i) angular gyrus (BA39), (ii) dorsolateral prefrontal cortex (BA9),

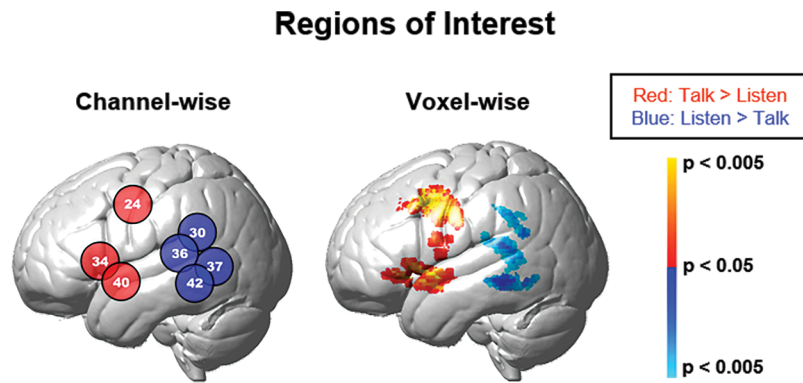


Fig. 2. Functionally determined ROIs for talking and listening. Monologue and interactive conditions are combined to isolate brain regions associated with talking (talk > listen), red/yellow colors, and listening (listen > talk), blue colors. Findings are presented on the left hemisphere showing two computational approaches: channel-wise (left panel, Table 1, $P < 0.05$) and voxel-wise (right panel, Table 2. See color bar on right). Group deOxyHb signals ($n = 58$).

(iii) dorsolateral prefrontal cortex (BA46), (iv) pars triangularis (BA45), (v) supramarginal gyrus (BA40), (vi) fusiform gyrus (BA37), (vii) middle temporal gyrus (BA21), (viii) superior temporal gyrus (STG) (BA22), (ix) somatosensory cortex (BA1, 2 and 3), (x) premotor and supplementary motor cortex (BA6), (xi) subcentral area (SCA) (BA43) and (xii) frontopolar cortex (BA10). See Supplementary Table S2 for median group region centroids.

Signal processing

Baseline drift was removed using wavelet detrending (NIRS-SPM). Any channel with strong noise was identified automatically by the root mean square of the raw data when the signal magnitude was more than 10 times the average signal. These events are assumed to be due to insufficient optode contact with the scalp. Approximately 4% of the channels were automatically excluded based on this criterion.

Global component removal

Systemic global effects (e.g. blood pressure, respiration, and blood flow variation) have previously been shown to alter relative blood hemoglobin concentrations (Kirilina et al., 2012). These effects are represented in fNIRS signals, raising the possible confound of inadvertently measuring hemodynamic responses that are not due to neuro-vascular coupling (Tachtsidis and Scholkmann, 2016). Global components were removed using a principle components analysis spatial filter (Zhang et al., 2016; Zhang et al., 2017) prior to general linear model (GLM) analysis. This technique exploits advantages of distributed optode coverage to distinguish signals originating from local sources (assumed to be specific to the neural events under investigation) by removing signal components due to global factors that originate from systemic cardiovascular functions.

Signal selection

Both deOxyHb and OxyHb signals are acquired by fNIRS. The signal choice for this study was empirically determined by comparisons of both signals with respect to language fiducials for talking (Broca's area) and listening (Wernicke's area) using the total data set combined over both conditions. Both signals were evaluated using raw (unprocessed) signals (Figure 1, middle row, C and D) and for signals with the global component removed using the high-pass spatial filtering (Figure 1, 'filtered', bottom

row, E and F). Circled clusters in panel E document left hemisphere canonical language productive (red) and receptive (blue) fiducial regions not seen for OxyHb signals or raw data of the other panels. Observation of these fiducial regions for this task validates the use of 'filtered' deOxyHb signals for this study. The deOxyHb signal associated with talking and listening tasks using fNIRS has also previously been validated (Scholkmann et al., 2013b; Zhang et al., 2017). Similar comparisons have led to the use of deOxyHb signals in investigations of conflict (Noah et al., 2017), signal reliability (Dravida et al., 2017), eye-to-eye contact (Zhang et al., 2017) and competitive games (Piva et al., 2017). The deOxyHb signal is characterized by lower signal-to-noise than the OxyHb signal, accounting for the more prevalent use of the OxyHb signal. However, validation based on observed fiducial markers confirms that actual speaking and listening tasks for this investigation are best represented by the deOxyHb signal following removal of global mean components and guided the decision to base findings on the deOxyHb signal.

ROIs

ROIs were determined by contrast effects based on comparisons of talking vs listening data combined for both conditions and determined by voxel-wise and channel-wise signal processing approaches. Voxel-wise analysis (Figure 2, right image) used computational tools conventionally applied to fMRI and provided the most precise spatial locations of activity using fine-grained interpolation between the channels. Channel-wise analysis (Figure 2, left image) provided a discrete report of activity averaged within each single channel. Channels were adjusted to cover the same brain regions for all subjects (Zhang et al., 2017). Each method is discussed below. All ROIs were on the left hemisphere.

ROIs by voxel-wise contrast comparisons

The 42-channel fNIRS data sets per subject were reshaped into 3D volume images for the first-level GLM analysis using SPM8. Beta values (i.e. signal amplitudes) were normalized to standard MNI space using linear interpolation. Any voxel that was ≥ 1.8 cm away from the brain surface was excluded. The computational mask consisted of $3753 \times 2 \times 2$ mm voxels that 'tiled' the shell region covered by the 42 channels. Anatomical variation across subjects was used to generate the distributed response maps. This approach provided a spatial resolution advantage

Table 1. Regions of interest: all conditions, channel-wise GLM contrast comparisons (deOxyHb signals)

Contrast	Channel number	MNI coordinates ^a			t	P	Anatomical regions in area	BA ^b	Probability ^c
		X	Y	Z					
[Talk > listen]	24	-62	-4	37	1.86	0.034	Pre-motor and supplementary motor cortex	6	0.98
	34	-60	16	6	2.01	0.025	STG ^d	22	0.21
							Pars opercularis	44	0.37
							Pars triangularis	45	0.31
	40	-65	-1	-6	2.30	0.013	Middle temporal gyrus	21	0.71
STG	22	0.29							
[Listen > talk]	42	-68	-42	-4	-2.20	0.016	Middle temporal gyrus	21	0.81
	42	-68	-42	-4	-2.20	0.016	STG	22	0.19
							STG	22	0.32
	30	-68	-40	24	-2.21	0.016	Supramarginal gyrus	40	0.65
	36	-69	-30	14	-1.82	0.037	STG	22	0.41
							Supramarginal gyrus	40	0.14
	37	-66	-52	6	-1.87	0.033	Auditory primary and association cortex	42	0.46
							Middle temporal gyrus	21	0.56
STG							22	0.43	

^aCoordinates are based on the MNI system and (-x) indicates left hemisphere; ^bBA, Brodmann area; ^cProbability of inclusion; ^dSTG, Superior Temporal Gyrus.

achieved by interpolation across participants and between channels. Results are presented on the normalized brain using images rendered on a standardized MNI template. See Figure 2, right image, and Table 1. Anatomical locations of peak voxel activity were identified using NIRS-SPM (Ye et al., 2009).

ROIs by channel-wise contrast comparisons

Although voxel-wise analysis offers precise estimates of spatial localization by virtue of interpolated activity computed from spatially distributed signals, this method provides a highly processed description of the data. Here, we take advantage of the discrete actual spatial sampling of fNIRS by using separate channels as the fundamental unit of analysis. Channel locations for each subject were converted to MNI space and registered to median locations using non-linear interpolation. Once in normalized space, comparisons across conditions were based on discrete channel units as originally acquired rather than interpolated voxel units. See Figure 2, left image, and Table 2.

All reported findings were observed with both analyses and served to identify the ROIs employed in this study. These regions are consistent with expectations of the Object Naming & Description task as determined for neurosurgical applications using covert speech and fMRI (Hirsch et al., 2000) and current models of human language systems (Gabrieli et al., 1998; Binder et al., 2000; Price, 2012; Hagoort, 2014; Poeppel, 2014; Abel et al., 2015). Specifically, for talking (productive language), active regions included left hemisphere pars opercularis, pars triangularis and inferior frontal gyrus, as well as left hemisphere primary, pre-motor and supplementary motor cortex (Tables 1 and 2, top rows). For listening (receptive language), active regions included left hemisphere, STG, supramarginal gyrus, primary and association auditory cortex and fusiform gyrus (Tables 1 and 2, bottom rows). Together, observed ROIs include canonical regions for 'sending' and 'receiving' language-related information and provide an empirical and conventional framework to test the hypothesis that neural systems recruited in interactive talking and listening are distinguished from

those employed in non-interactive (monologue) talking and listening.

Wavelet analysis for cross-brain effects

Cross-brain synchrony (coherence) was evaluated using wavelet analysis (Torrence and Compo, 1998; Cui et al., 2011). The wavelet kernel was a complex Gaussian (Mexican hat-shaped kernel) provided by MATLAB. The number of octaves was 4 and the range of frequencies was 0.4 to 0.025 Hz. The number of voices per octave was also 4, and therefore 16 scales were used for which the wavelength difference was 2.5 s. Methodological details and validation of this technique have been previously described (Zhang et al., 2017). The analysis was conducted by using concatenated segments of the same kind (talking and listening) and coherence values were averaged across the whole data time series, i.e. associated with consecutive and alternating talking and listening tasks. This approach provided a measurement of non-symmetric coupled dynamics (Hasson and Frith, 2016), where the listener's neural signals were synchronized with the speaker's neural signals representing predictable transformations between the two brains. Signals acquired from the predefined anatomical regions (Supplementary Table S2) were decomposed into various temporal frequencies that were correlated across the two brains for each dyad following removal of the task regressor as is conventional for psychophysiological interaction analysis (Friston et al., 1997). Analysis of the residual signal according to this technique theoretically eliminates the correlated and anticorrelated effects induced by different tasks such as talking and listening performed simultaneously within specified blocks of time. Here we apply the residual signal to investigate effects other than the main task-induced effect. For example, cross-brain coherence of multiple signal components (wavelets) is thought to provide an indication of dynamic coupling processes rather than task-specific processes. Coherence during speaking and listening exchanges were compared for the two conditions: with and without interpersonal interaction. This analysis was also applied to scrambled (random pairs) of dyads to control for possible effects of common processes.

Table 2. Regions of interest: all conditions, voxel-wise GLM contrast comparisons (deOxyHb signals)

Contrast	Contrast threshold	Peak Voxels					Anatomical regions in cluster	BA ^c	Probability	n of voxels		
		MNI coordinates ^a			t	P					df ^b	
[Talk > listen]	P = 0.05	-56	4	-4	2.77	0.0038	56	STG ^d	22	0.39	660	
								Middle temporal gyrus	21	0.34		
									Temporopolar area	38		0.15
		64	-16	-6	2.44	0.0089	56	Middle temporal gyrus	21	0.76		133
								Superior temporal gyrus	22	0.20		
		-64	0	16	2.16	0.0175	56	Pre-motor and supplementary motor cortex	6	0.34		131
								STG	22	0.22		
								Pars opercularis	44	0.15		
		-58	0	40	3.51	0.0004	56	Pre-motor and supplementary motor cortex	6	0.75		809
								Dorsolateral prefrontal cortex	9	0.14		
56	2	44	3.06	0.0017	56	Pre-motor and supplementary motor cortex	6	0.73	684			
						Dorsolateral prefrontal cortex	9	0.13				
[Listen > talk]	P = 0.05	-70	-26	12	-3.17	0.0012	56	STG	22	0.36	739	
								Primary and association auditory cortex	42	0.33		
		Supramarginal gyrus	40	0.11								
		68	-44	10	-3.62	0.0003	56	STG	22	0.67		254
								Middle temporal gyrus	21	0.25		

^aCoordinates are based on the MNI system and (-) indicates left hemisphere; ^bdf, degrees of freedom; ^cBA, Brodmann area; ^dSTG, Superior Temporal Gyrus.

Results

Statistical comparisons of ROIs between monologue and interactive conditions

According to the functionally defined regions (Figure 3), the receptive (listening) function, Wernicke's ROI, included channel numbers 30, 36, 37, and 42 and the productive (talking) function, Broca's ROI, included channel numbers 34, 40, and 24 (see inset illustrations and x-axis). Group-averaged signal strength is shown on the y-axis. Statistical comparisons between the two conditions of monologue and interactive (Figure 3) are shown for each of the two ROIs: Broca's area, left panel, red bar; and Wernicke's area, right panel, blue bar. The horizontal dotted line indicates equal average signal strength for the two conditions: monologue and interactive and interactive signal strength greater than monologue signal strength is a positive y-value. The average interactive signal exceeded the average monologue signal ($P < 0.04$, $t = 2.07$, $d = 57$) for Wernicke's ROI (right panel), and there was no evidence for a difference between monologue and interactive conditions for Broca's ROI (left panel). Figure S1 presents the individual subject scatter plots associated with this bar graph for further clarification.

In addition to the average signal strength comparisons above, we also compare the cluster sizes for each of the ROIs for the channel-wise and voxel-wise analyses. The anatomical 'heat maps' for interactive and non-interactive (monologue) conditions are shown on Figure 4A and B, respectively. In the case of cluster sizes for activity associated with the [listen > talking] (blue) condition, the cluster size increases from one channel in the monologue condition to five adjacent channels in the interactive condition. These results are similar to the comparison of the number of voxels in the voxel-wise analysis. Tables 3 (channel-wise, monologue) and 4 (voxel-wise monologue) and 5 (channel-wise, interaction), and 6 (voxel-wise, interaction) provide the cluster/channel locations, anatomical labels, and statistical reports for these findings and further document

that Wernicke's area was most responsive to the interactive condition.

Comparison of cross-brain coherence between monologue and interactive conditions

A component of Wernicke's area, the Superior Temporal Gyrus (STG), and the Subcentral Area (SCA) increased cross-brain coherence during the interactive condition. Specifically, left and right hemispheres of STG (brain 1) were coupled with the left and right hemispheres of the SCA (brain 2) and vice versa. Temporal oscillations of hemodynamic signals decomposed into wavelet components (Figure 5, x-axis) are plotted against the correlation between the signals of partners acquired while engaged in the joint task of talking and listening (y-axis). Cross-brain coherence was greater for the interactive (red) than the monologue (blue) condition for wavelet components between 8 and 14 s for actual partners, shown on the left panel ($P < 0.01$), Figure 5A. However, for scrambled partners (i.e. randomly paired with every other participant except the original partner, right panel, Figure 5B), there was no difference. No other pairs of regions met these statistical criteria and control conditions. The neuroanatomy of the coherent pairs are illustrated in Figure 5C.

Discussion

Although interpersonal interaction is an essential and evolutionarily conserved human behavior, insight into the underlying neurobiology is sparse. Investigations of interactions between two freely behaving individuals are typically beyond conventional neuroimaging capabilities, which are generally limited to single subjects, intolerant of head movement and restricted to confined and loud environments. In this study, we introduce a solution to many of these limitations using functional near infrared spectroscopy with whole-head surface-mounted

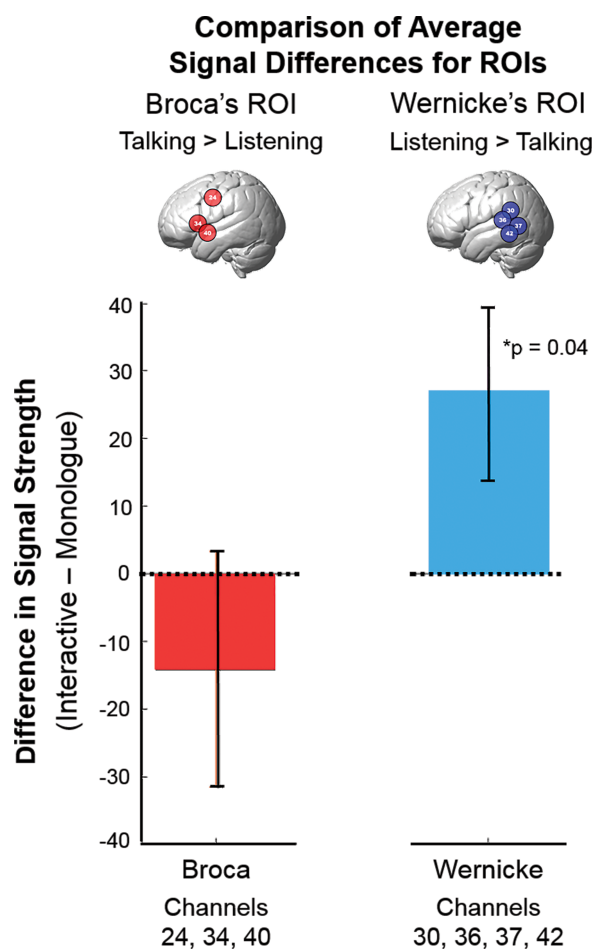


Fig. 3. Statistical comparisons of signal amplitudes for Broca's and Wernicke's ROIs. Channels more active during listening are represented in blue (inset brain) and were associated with Wernicke's ROI (right panel), and channels more active during talking are represented in red (inset brain) and were associated with Broca's ROI (left panel) and are in accordance with classical models of speech reception and production, respectively. Differences in signal strength between the interactive and monologue conditions are indicated on the y-axis, and the colored bars (\pm SEM) indicate the respective conditions. In the case of Wernicke's ROI, signals during the Interactive condition were increased relative to monologue ($P = 0.04$, right panel), whereas there was no evidence for a difference between the two conditions in Broca's ROI. (deOxyHb signal, $n = 58$). See scatterplot (Supplementary Material, Figure 1).

detectors that permit simultaneous dual-brain imaging in natural interactive conditions. Prior interactive investigations using EEG signals have also established a foundation for cross-brain investigations without the disadvantages of conventional scanning with functional MRI. This novel system and paradigm represents a growing shift from the single-brain to a dual-brain frame of reference that enables direct observations of neural signals and coupled dynamics between interacting dyads.

Summary of contrast and coherence findings

The hypotheses that regional neural activity and cross-brain coherence of canonical language areas is modulated by interpersonal interaction was tested in this study by the comparison of two verbal communication conditions, interactive and monologue (non-interactive). Although participants sat across a table from each other, they were occluded from each other so that facial expressions and body gestures were not a part of the

interaction. Further, participants were not acquainted with each other prior to the experiment, which was intended to eliminate possible effects of affiliation.

Increased contrast-based neural activity was associated with the interactive task within the temporal language receptive ROI, Wernicke's area (including the posterior STG), and this region was also associated with cross-brain coherence. However, no evidence supported the hypothesis for Broca's area. Taking a cautionary point of view, an effect of interaction might be expected based on general features of the task differences such as non-specific increased task difficulty, memory load, attention, arousal, etc. However, these factors would be expected to influence both frontal and temporal regions. Here we observe modulations only in the posterior sectors consistent with specific effects of interaction associated with receptive systems.

Cross-brain coherence is taken as an indicator of dynamic coupling between cooperating neural systems engaged in reciprocal exchanges of information. Cross-brain coherence for signals originating within the STG (a subset of Wernicke's area) and the SCA increased during verbal exchanges with interactive narratives as compared to verbal exchanges with non-interactive, monologue, narratives. Although it was predicted that Wernicke's area would be one of the regions associated with cross-brain coherence, the region dynamically coupled with Wernicke's area, the SCA, is not a canonical language region. This finding is consistent with a model for a distinct communicative interaction mechanism that facilitates live interactive communications although the findings suggest that language systems may participate.

A functional role for the SCA

Although the STG is a well-established component of Wernicke's area, the canonical language region generally associated with receptive functions, the SCA (BA43), is not associated with a well-defined function. This area is formed by the union of the pre- and post-central gyri at the inferior end of the central sulcus in face-sensitive topography, with internal projections that extend into the supramarginal area over the inner surface of the operculum with a medial boundary at the insular cortex (Brodmann and Garey, 1999). The finding of cross-brain coupling between the STG and the SCA suggests a functional role for the SCA within a nexus of neural mechanisms associated with receiving and processing live social content. Consistent with this interpretation, the SCA was also found to be active during real eye-to-eye contact between interacting partners compared to mutual gaze at a face picture (Hirsch et al., 2017). The discovery of a role for the SCA in cross-brain neural coupling during non-symmetrical (simultaneous talking and listening) verbal interactions advances the hypothesis of a specific neural substrate that underlies the coupled dynamics between speech comprehension and speech production.

A pathway for cross-brain interaction

Coupled neural activity, as represented by the hemodynamic signal of both the listener and the speaker (non-symmetrical neural coupling, Hasson & Frith, 2016), reflects a similarly synchronous temporal neural pattern during interaction in the range of 8–14 seconds. This speaker-listener neural coupling reveals a shared neural substrate across interlocutors (Hasson et al., 2012). The observed findings are consistent with this

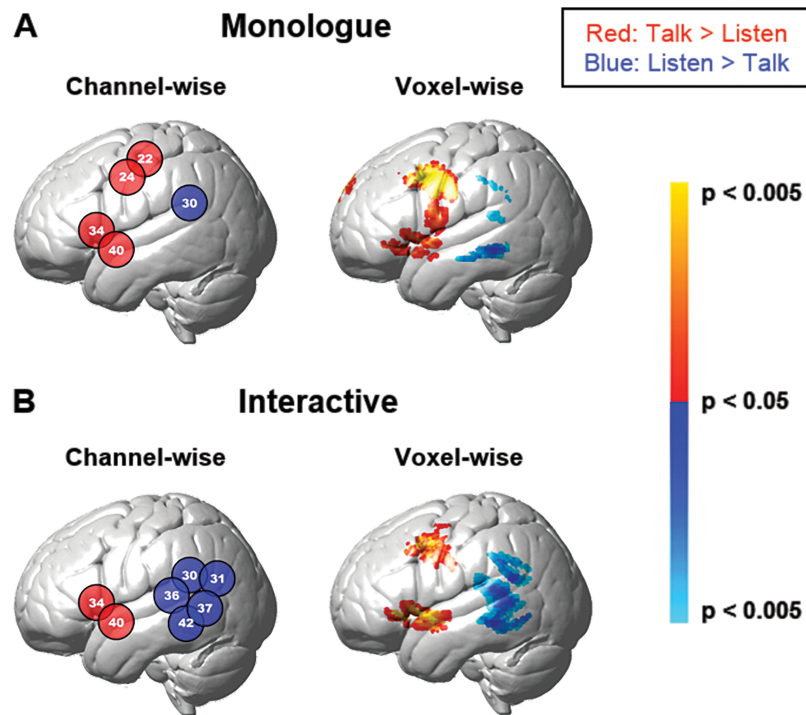


Fig. 4. Contrast effects. Contrast findings for monologue (A, top row) and interactive (B, bottom row) conditions are presented based on two computational approaches: channel-wise (left, $P < 0.05$) and voxel-wise (right, $P < 0.05$). The numbers in the circles indicate the channels. Red and blue colors indicate contrasts [talk > listen] and [listen > talk], respectively. Tables 3 and 4 and Tables 5 and 6 include the monologue and interactive conditions, respectively, and indicate channels and clusters, anatomical labels, Brodmann Area (BA), probability of inclusion and statistical levels (deOxyHb signal, $n = 58$).

Table 3. Channel-wise GLM contrast comparisons (deOxyHb signals), monologue

Contrast	Channel number	MNI coordinates ^a			P	t	Anatomical regions	BA ^b	Probability						
		X	Y	Z											
Monologue [talk > listen]	22	-59	-14	48	0.013	2.30	Primary somatosensory cortex	1	0.085						
							Primary somatosensory cortex	2	0.066						
							Primary somatosensory cortex	3	0.206						
							Primary motor cortex	4	0.143						
	24	-62	-4	37	0.003	2.88	Pre-motor and supplementary motor cortex	6	0.500						
							Primary motor cortex	4	0.021						
	34	-60	16	6	0.036	1.84	Pre-motor and supplementary motor cortex	6	0.979						
							STG ^d	22	0.211						
							Pars opercularis	44	0.372						
							Pars triangularis	45	0.306						
							Inferior frontal gyrus	47	0.095						
							Middle temporal gyrus	21	0.707						
40	-65	-1	-6	0.031	1.91	STG	22	0.286							
						Temporopolar area	38	0.007							
						Monologue [listen > talk]	30	-68	-40	24	0.046	-1.71	STG	22	0.323
													Supramarginal gyrus	40	0.652
							Primary and association auditory cortex	42	0.025						

^aCoordinates are based on the MNI system and (-) indicates left hemisphere; ^bBA, Brodmann area; ^dSTG, Superior Temporal Gyrus.

conjecture and advance evidence for a specific cross-brain neural substrate (STG and SCA) that processes or transforms the exchanged neural information within a specific temporal range. This model also proposes that dynamic coupling is a mechanism

by which social information is shared between the receiver and sender (Hasson and Frith, 2016). Accordingly, in this case, the coupled neural signals observed in this study theoretically reflect neural transformations of the individual speaker's and

Table 4. Voxel-wise GLM contrast comparisons (deOxyHb signals), monologue

Contrast	Contrast threshold	Peak voxels					Anatomical regions in cluster	BA ^c	Probability	n of voxels		
		MNI coordinates ^a			t	P					df ^b	
Monologue [talk > listen]	P = 0.05	-62	-2	36	3.81	0.0002	56	Pre-motor and supplementary motor cortex	6	0.75	1575	
								Dorsolateral prefrontal cortex	9	0.14		
		66	-26	-8	2.39	0.0102	56	Middle temporal gyrus	21	0.77		134
								STG ^d	22	0.16		
		8	54	40	3.27	0.0009	56	Dorsolateral prefrontal cortex	9	0.58		156
								Frontal eye fields	8	0.27		
		56	2	42	3.29	0.0009	56	Frontopolar area	10	0.15		824
Pre-motor and supplementary motor cortex	6							0.72				
Dorsolateral prefrontal cortex	9	0.17										
Monologue [listen > talk]	P = 0.05	-64	-36	-8	-2.67	0.0049	56	Middle temporal gyrus	21	0.84	185	
								STG	22	0.65		
		70	-44	8	-3.10	0.0015	56	Middle temporal gyrus	21	0.32	158	
								STG	22	0.67		
		68	-44	10	-2.54	0.0069	56	Supramarginal gyrus	40	0.18		
								Middle temporal gyrus	21	0.14		

^aCoordinates are based on the MNI system and (-) indicates left hemisphere; ^bdf, degrees of freedom; ^cBA, Brodmann area; ^dSTG, Superior Temporal Gyrus.

Table 5. Channel-wise GLM contrast comparisons (deOxyHb signals), interactive

Contrast	Channel number	MNI coordinates ^a			P	t	Anatomical regions	BA ^b	Probability
		X	Y	Z					
Interactive [talk > listen]	34	-60	16	6	0.030	1.91	Pre-motor and supplementary motor cortex	6	0.016
							STG ^d	22	0.211
							Pars opercularis	44	0.372
							Pars triangularis	45	0.306
	40	-65	-1	-6	0.007	2.55	Inferior Frontal gyrus	47	0.095
							Middle temporal gyrus	21	0.707
							STG	22	0.286
Temporopolar area	38	0.007							
Interactive [listen > talk]	30	-68	-40	24	0.009	-2.45	STG	22	0.323
							Supramarginal gyrus	40	0.652
	31	-60	-62	24	0.034	-1.85	Primary and association auditory cortex	42	0.025
							Visual cortex (V3)	19	0.046
							STG	22	0.194
							Angular gyrus	39	0.616
	36	-69	-30	14	0.026	-1.99	Supramarginal gyrus	40	0.144
							STG	22	0.407
							Supramarginal gyrus	40	0.135
	37	-66	-52	6	0.015	-2.23	Primary and association auditory cortex	42	0.456
							SCA	43	0.003
							Middle temporal gyrus	21	0.560
	42	-68	-42	-4	0.009	-2.42	STG	22	0.427
							Fusiform gyrus	37	0.013
Middle temporal gyrus							21	0.808	
STG							22	0.189	
Fusiform gyrus	37	0.003							

^aCoordinates are based on the MNI system and (-) indicates left hemisphere; ^bBA, Brodmann area; ^dSTG, Superior Temporal Gyrus.

listener's neural patterns. We speculate that these putative 'on-line' neural transformations mediate and, perhaps, regulate the continuously adapting stream of interactive information. If

so, then this model predicts an increase in cross-brain coherence between STG and the SCA with increased interactive cues (information). The functional significance of these effects is yet

Table 6. Voxel-wise GLM Contrast comparisons (deOxyHb signals), Interactive

Contrast	Contrast threshold	Peak voxels			t	P	df ^b	Anatomical regions in cluster	BA ^c	Probability	n of voxels
		MNI coordinates ^a									
Interactive [talk > listen]	P = 0.05	-62	6	-4	3.15	0.0013	56	Middle temporal gyrus	21	0.36	658
								STG ^d	22	0.35	
		-58	0	40	3.11	0.0015	56	Temporopolar area	38	0.13	
								Pre-motor and supplementary motor cortex	6	0.77	514
46	-2	44	2.45	0.0087	56	Dorsolateral prefrontal cortex	9	0.16			
						Pre-motor and supplementary motor cortex	6	0.95	412		
Interactive [listen > talk]	p = 0.05	-68	-44	10	-3.57	0.0004	56	STG	22	0.64	1132
								Middle temporal gyrus	21	0.27	
		68	-44	10	-3.77	0.0002	56	STG	22	0.67	474
								Middle temporal gyrus	21	0.25	

^aCoordinates are based on the MNI system and (-) indicates left hemisphere; ^bdf, degrees of freedom; ^cBA, Brodmann area; ^dSTG, Superior Temporal Gyrus.

to be determined but predictably influence functions related to the quality of communication such as comprehension, arousal, social judgments and decision-making.

Dynamic coupling of real and ‘scrambled’ partners

Comparison of coherence between actual partners (Figure 5A) and scrambled partners (Figure 5B) distinguishes between two possible interpretations: (i) cross-brain correlations are due to similar operations performed by both partners or (ii) cross-brain correlations are due to events specific to the partner interactions. If the coherence between a pair of brain areas remains significantly different for the interactive and monologue conditions during the scrambled cases, then we conclude in favor of common processes. If, however, the coherence difference between the dyads is observed for actual partners only, then we conclude in favor of the interaction-specific option. Cross-brain coherence was found to be greater during the interactive condition than the monologue condition for the real partners and not for the scrambled partners and is taken as support for option 2: cross-brain correlations are due to events specific to the partner interactions. These findings also advance an approach for analysis of brain-to-brain coupling using localized wavelets that takes into account the different time scales of putative parallel neural processes and their origins.

Neural processes sensitive to the demands of social interaction and cautionary notes

We predicted increased language-specific neural activity associated with social interaction. Although this hypothesis was supported for left hemisphere regions associated with listening (receptive regions consistent with Wernicke’s area), there was no evidence for either contrast-based (GLM) or coherence-based (wavelet analysis) interaction-modulated activity in the frontal language production ROIs, classically known as Broca’s region. This region includes the supplementary motor cortex (articulatory system), the left inferior frontal gyrus as well as the anterior STG (also part of Broca’s area). The anterior STG, in particular, has recently been shown to be susceptible to possible artifacts

from muscle activity during speaking tasks (Morais et al., 2017). In this investigation there was no effect of interaction in any of the frontal regions including the anterior STG. However, the possibility of a type-II error cannot be ruled out, i.e. a positive true result might not be detected by the analysis due to a motion-related artifact. The same null finding was also observed in other frontal regions known to be associated with speech production and Broca’s area including the inferior frontal gyrus and the supplementary motor cortex. These areas have not been associated with a putative artifact due to muscle movement during speaking. Further, if due to artifact, observations of activity would also be expected on the right hemisphere as well. The right hemisphere showed no significant activity for these contrasts consistent with true findings.

It is well established that the use of real speech as well as inner speech in fNIRS studies can be complicated by changes in hemodynamics and oxygenation (Scholkmann et al., 2013a; Scholkmann et al., 2013c). In this study, there was no evidence for a difference in the number or character of words spoken in either condition suggesting that differences in oxygenation efforts would not differ systematically. Thus, observed interaction effects were most likely due to factors other than changes in end-tidal CO₂. In light of these challenges, however, we completed a ‘proof-of-principle’ experiment prior to this study documenting the validity of the deOxyHb signal and the live speaking techniques. This previous study confirmed reliable observations of known fiducial language regions that were subsequently applied in this investigation (Zhang et al., 2017).

Relevance to models of developmental disorders

These findings are relevant to models of language and social interaction deficits reported in developmental disorders such as autism spectrum disorder (ASD). Comparisons of neural activity for ASD and typically developing children acquired during passive listening to recorded speech reveal reductions in activation of the STG (Wernicke’s region) in ASD (Lai et al., 2011). Although the findings may suggest that language deficits in ASD are associated with disrupted linguistic comprehension, in view of

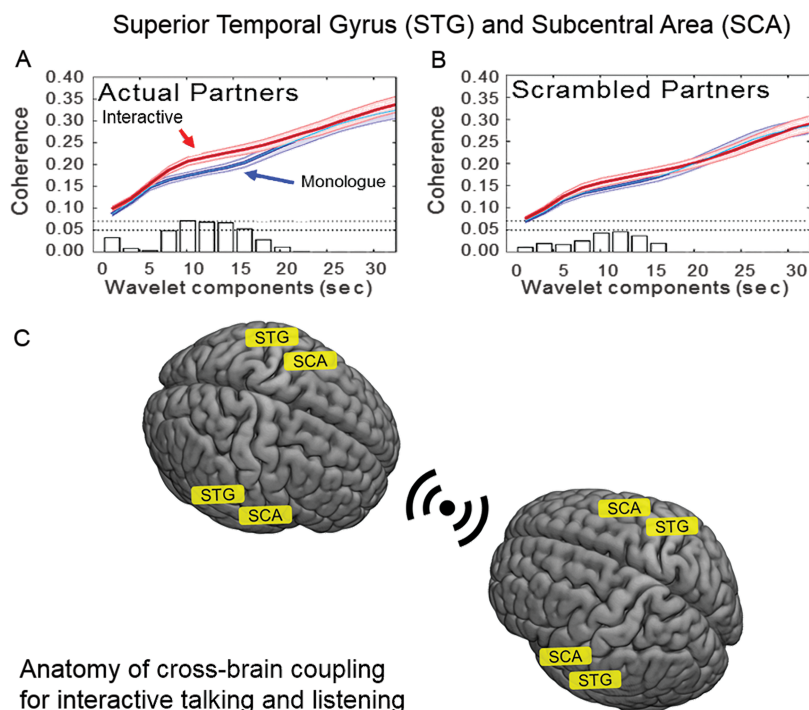


Fig. 5. Cross-brain coherence. Signal coherence between the STG and SCA of participant dyads (y-axis) is plotted against the wavelet components (seconds, x-axis). The functions represent interactive (red) and monologue (blue) conditions (shaded areas: ± 1 SEM). Bar graphs along x-axis indicate significance levels for the separations between the two conditions for each of the period values on the x-axis. The upper horizontal dashed line indicates $P \leq 0.01$ and the lower line indicates $P \leq 0.05$. Panel A shows coherence between actual partners, and panel B shows coherence between scrambled partners. The coherence for periods (wavelet components) between 8 and 14 s is greater for the interactive condition than for monologue in the case of the actual partners ($P < 0.01$). There is no difference between the coherence functions for interactive and monologue conditions in the case of the scrambled partners. (deOxyHb signals, $n = 27$ pairs). C. A graphical illustration of the coupled brain areas, STG to SCA, during simultaneous talking and listening epoch pairs.

the dual-brain interactive findings of this study, an alternative hypothesis emerges that hypoactivity of the temporal-parietal complex may be associated with social interaction disabilities. Together, these findings shed new light on the investigations of language and social disabilities. Impairments related to interpersonal interactions may also be associated with specific neurophysiological sub-systems dedicated to cross-brain interactions that intersect with the language system. This hypothesis suggests a novel future direction for investigation.

Supplementary data

Supplementary data are available at SCAN online.

Funding

This research was partially supported by the National Institute of Mental Health of the National Institutes of Health under award number R01MH107513 (J.H.), the National Institutes of Health Medical Scientist Training Program Training Grant T32GM007205 (S.D.) and the Japan Society for the Promotion of Science grants in Aid for Scientific Research (KAKENHI) JP15H03515 and JP16K01520 (Y.O.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. All data reported in this paper are available upon request from the corresponding author.

Acknowledgements

The authors are grateful for the significant contributions of Jeiyoun Park and Pawan Lapborisuth, Yale University undergraduates, for data collection and operations support; Dr Ilias Tachtsidis, Department of Medical Physics and Biomedical Engineering; Dr Antonia Hamilton and Prof. Paul Burgess, Institute for Cognitive Neuroscience; and Maurice Biriotti, SHM and Department of Medical Humanities, University College London for insightful comments and guidance.

Conflict of Interest. None declared.

References

- Abel, T.J., Rhone, A.E., Nourski, K.V., et al. (2015). Direct physiologic evidence of a heteromodal convergence region for proper naming in human left anterior temporal lobe. *Journal of Neuroscience*, 35(4), 1513–20.
- Babiloni, F., Astolfi, L. (2014). Social neuroscience and hyper-scanning techniques: Past, present and future. *Neuroscience & Biobehavioral Reviews*, 44, 76–93.
- Binder, J.R., Frost, J.A., Hammeke, T.A., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, 10(5), 512–28.
- Boas, D.A., Elwell, C.E., Ferrari, M., Taga, G. (2014). Twenty years of functional near-infrared spectroscopy: Introduction for the special issue. *Neuroimage*, 85(1), 1–5.

- Brodmann, K., Garey, L. (1999). *Brodmann's Localisation in the Cerebral Cortex*, London: Imperial College Press.
- Cui, X., Bray, S., Bryant, D.M., Glover, G.H., Reiss, A.L. (2011). A quantitative comparison of nirs and fmri across multiple cognitive tasks. *Neuroimage*, **54**(4), 2808–21.
- 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), 2430–7.
- De Jaegher, H., Di Paolo, E., Adolphs, R. (2016). What does the interactive brain hypothesis mean for social neuroscience? A dialogue. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **371**(1693), 20150379.
- De Langavant, L.C., Remy, P., Trinkler, I., et al. (2011). Behavioral and neural correlates of communication via pointing. *PLoS One*, **6**(3), e17719.
- Di Paolo, E., De Jaegher, H. (2012). The interactive brain hypothesis. *Frontiers in Human Neuroscience*, **6**(163), 1–16.
- Dommer, L., Jäger, N., Scholkmann, F., Wolf, M., Holper, L. (2012). Between-brain coherence during joint n-back task performance: a two-person functional near-infrared spectroscopy study. *Behavioural Brain Research*, **234**(2), 212–22.
- Dravida, S., Noah, J.A., Zhang, X., Hirsch, J. (2017). Comparison of oxyhemoglobin and deoxyhemoglobin signal reliability with and without global mean removal for digit manipulation motor tasks. *Neurophotonics*, **5**(1), 011006.
- Dumas, G., de Guzman, G.C., Tognoli, E., Kelso, J.S. (2014). The human dynamic clamp as a paradigm for social interaction. *Proceedings of the National Academy of Sciences*, **111**(35), E3726–34.
- Engelbrecht, A.T., White, B.R., Ferradal, S.L., et al. (2012). A quantitative spatial comparison of high-density diffuse optical tomography and fmri cortical mapping. *Neuroimage*, **61**(4), 1120–8.
- Ferradal, S.L., Engelbrecht, A.T., Hassanpour, M., Snyder, A.Z., Culver, J.P. (2014). Atlas-based head modeling and spatial normalization for high-density diffuse optical tomography: In vivo validation against fmri. *Neuroimage*, **85**(1), 117–26.
- Ferrari, M., Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fnirs) development and fields of application. *Neuroimage*, **63**(2), 921–35.
- Friston, K., Buechel, C., Fink, G., Morris, J., Rolls, E., Dolan, R. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage*, **6**(3), 218–29.
- Funane, T., Kiguchi, M., Atsumori, H., Sato, H., Kubota, K., Koizumi, H. (2011). Synchronous activity of two people's prefrontal cortices during a cooperative task measured by simultaneous near-infrared spectroscopy. *Journal of Biomedical Optics*, **16**(7), 077011.
- Gabrieli, J.D., Poldrack, R.A., Desmond, J.E. (1998). The role of left prefrontal cortex in language and memory. *Proceedings of the National Academy of Sciences*, **95**(3), 906–13.
- García, A.M., Ibáñez, A. (2014). Two-person neuroscience and naturalistic social communication: The role of language and linguistic variables in brain-coupling research. *Frontiers in Psychiatry*, **5**(124), 1–6.
- Gracco, V.L., Tremblay, P., Pike, B. (2005). Imaging speech production using fmri. *Neuroimage*, **26**(1), 294–301.
- Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region and beyond. *Current Opinion in Neurobiology*, **28**, 136–41.
- Hart, J., Jr., Rao, S.M., Nuwer, M. (2007). Clinical functional magnetic resonance imaging. *Cognitive and Behavioral Neurology*, **20**(3), 141–4.
- Hasson, U., Frith, C.D. (2016). Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **371**(1693), 20150366.
- Hasson, U., Ghazanfar, A.A., Galantucci, B., Garrod, S., Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends in Cognitive Sciences*, **16**(2), 114–21.
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, **303**(5664), 1634–40.
- Hasson, U., Yang, E., Vallines, I., Heeger, D.J., Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *Journal of Neuroscience*, **28**(10), 2539–50.
- Hirsch, J., Ruge, M.I., Kim, K.H., et al. (2000). An integrated functional magnetic resonance imaging procedure for preoperative mapping of cortical areas associated with tactile, motor, language, and visual functions. *Neurosurgery*, **47**(3), 711–22.
- 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**, 314–30.
- Holper, L., Scholkmann, F., Wolf, M. (2012). Between-brain connectivity during imitation measured by fnirs. *Neuroimage*, **63**(1), 212–22.
- Jiang, J., Chen, C., Dai, B., et al. (2015). Leader emergence through interpersonal neural synchronization. *Proceedings of the National Academy of Sciences*, **112**(14), 4274–9.
- Jiang, J., Dai, B., Peng, D., Zhu, C., Liu, L., Lu, C. (2012). Neural synchronization during face-to-face communication. *Journal of Neuroscience*, **32**(45), 16064–9.
- Kawasaki, M., Yamada, Y., Ushiku, Y., Miyauchi, E., Yamaguchi, Y. (2013). Inter-brain synchronization during coordination of speech rhythm in human-to-human social interaction. *Scientific Reports*, **3**, 1692.
- Kinreich, S., Djalovski, A., Kraus, L., Louzoun, Y., Feldman, R. (2017). Brain-to-brain synchrony during naturalistic social interactions. *Scientific Reports*, **7**(1), 17060.
- Kirilina, E., Jelzow, A., Heine, A., et al. (2012). The physiological origin of task-evoked systemic artefacts in functional near infrared spectroscopy. *Neuroimage*, **61**(1), 70–81.
- Koike, T., Tanabe, H.C., Okazaki, S., et al. (2016). Neural substrates of shared attention as social memory: a hyperscanning functional magnetic resonance imaging study. *Neuroimage*, **125**, 401–12.
- Konvalinka, I., Roepstorff, A. (2012). The two-brain approach: how can mutually interacting brains teach us something about social interaction? *Frontiers in Human Neuroscience*, **6**.
- Lai, G., Schneider, H.D., Schwarzenberger, J.C., Hirsch, J. (2011). Speech stimulation during functional mr imaging as a potential indicator of autism. *Radiology*, **260**(2), 521–30.
- Liu, N., Mok, C., Witt, E.E., et al. (2016). Nirs-based hyperscanning reveals inter-brain neural synchronization during cooperative jenga game with face-to-face communication. *Frontiers in Human Neuroscience*, **10**(82), 11.
- Liu, Y., Piazza, E.A., Simony, E., et al. (2017). Measuring speaker-listener neural coupling with functional near infrared spectroscopy. *Scientific Reports*, **7**, 43293.
- Matcher, S.J.E., Cooper, C.E., Cope, M., Delpy, D.T. (1995). Performance comparison of several published tissue near-infrared spectroscopy algorithms. *Analytical Biochemistry*, **227**(1), 54–68.
- Morais, G.A.Z., Scholkmann, F., Balardin, J.B., et al. (2017). Non-neuronal evoked and spontaneous hemodynamic changes in the anterior temporal region of the human head may lead to misinterpretations of functional near-infrared spectroscopy signals. *Neurophotonics*, **5**(1), 011002.

- Noah, J.A., Dravida, S., Zhang, X., Yahil, S., Hirsch, J. (2017). Neural correlates of conflict between gestures and words: a domain-specific role for a temporal-parietal complex. *PLoS One*, *12*(3), e0173525.
- Okawa, S., Lee, T.M., Kay, A.R., Tank, D.W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, *87*(24), 9868–72.
- Okada, E., Delpy, D.T. (2003). Near-infrared light propagation in an adult head model. II. Effect of superficial tissue thickness on the sensitivity of the near-infrared spectroscopy signal. *Applied Optics*, *42*(16), 2915–21.
- Okamoto, M., Dan, I. (2005). Automated cortical projection of head-surface locations for transcranial functional brain mapping. *Neuroimage*, *26*(1), 18–28.
- Oldfield, R.C. (1971). The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.
- 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. *Frontiers in Psychology*, *6*(1811), 1–11.
- Owen-Reece, H., Smith, M., Elwell, C., Goldstone, J. (1999). Near infrared spectroscopy. *British Journal of Anaesthesia*, *82*(3), 418–26.
- Pan, Y., Cheng, X., Zhang, Z., Li, X., Hu, Y. (2017). Cooperation in lovers: an fnirs-based hyperscanning study. *Human Brain Mapping*, *38*(2), 831–41.
- Pinti, P., Aichelburg, C., Lind, F., et al. (2015). Using fiberless, wearable fnirs to monitor brain activity in real-world cognitive tasks. *Journal of Visualized Experiments*, *106*, 53336.
- Piva, M., Zhang, X., Noah, A., Chang, S.W., Hirsch, J. (2017). Distributed neural activity patterns during human-to-human competition. *Frontiers in Human Neuroscience*, *11*, 571.
- Poeppl, D. (2014). The neuroanatomic and neurophysiological infrastructure for speech and language. *Current Opinion in Neurobiology*, *28*, 142–9.
- Price, C.J. (2012). A review and synthesis of the first 20 years of pet and fmri studies of heard speech, spoken language, and reading. *Neuroimage*, *62*(2), 816–47.
- Saito, D.N., Tanabe, H.C., Izuma, K., et al. (2010). “Stay tuned”: inter-individual neural synchronization during mutual gaze and joint attention. *Frontiers in Integrative Neuroscience*, *4*(127), 1–11.
- Schilbach, L. (2014). On the relationship of online and offline social cognition. *Frontiers in Human Neuroscience*, *8*(278), 1–8.
- Schilbach, L., Timmermans, B., Reddy, V., et al. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, *36*(04), 393–414.
- Schippers, M.B., Roebroek, A., Renken, R., Nanetti, L., Keysers, C. (2010). Mapping the information flow from one brain to another during gestural communication. *Proceedings of the National Academy of Sciences*, *107*(20), 9388–93.
- Scholkmann, F., Gerber, U., Wolf, M., Wolf, U. (2013a). End-tidal co₂: an important parameter for a correct interpretation in functional brain studies using speech tasks. *Neuroimage*, *66*, 71–9.
- Scholkmann, F., Holper, L., Wolf, U., Wolf, M. (2013b). A new methodical approach in neuroscience: assessing inter-personal brain coupling using functional near-infrared imaging (fniri) hyperscanning. *Frontiers in Human Neuroscience*, *7*, 813.
- Scholkmann, F., Kleiser, S., Metz, A.J., et al. (2014). A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *Neuroimage*, *85*(1), 6–27.
- Scholkmann, F., Wolf, M., Wolf, U. (2013c). The effect of inner speech on arterial CO₂ and cerebral hemodynamics and oxygenation: a functional nirs study. *Advances in Experimental Medicine and Biology*, *789*, 81–7.
- Stephens, G.J., Silbert, L.J., Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences*, *107*(32), 14425–30.
- Strangman, G., Culver, J.P., Thompson, J.H., Boas, D.A. (2002). A quantitative comparison of simultaneous bold fmri and nirs recordings during functional brain activation. *Neuroimage*, *17*(2), 719–31.
- Tachtsidis, I., Scholkmann, F. (2016). False positives and false negatives in functional near-infrared spectroscopy: issues, challenges, and the way forward. *Neurophotonics*, *3*(3), 031405.
- Tanabe, H.C., Kosaka, H., Saito, D.N., et al. (2012). Hard to “tune in”: Neural mechanisms of live face-to-face interaction with high-functioning autistic spectrum disorder. *Frontiers in Human Neuroscience*, *6*, 268.
- Tang, H., Mai, X., Wang, S., Zhu, C., Krueger, F., Liu, C. (2016). Interpersonal brain synchronization in the right temporo-parietal junction during face-to-face economic exchange. *Social Cognitive and Affective Neuroscience*, *11*(1), 23–32.
- Torrence, C., Compo, G.P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, *79*(1), 61–8.
- Villringer, A., Chance, B. (1997). Non-invasive optical spectroscopy and imaging of human brain function. *Trends in Neuroscience*, *20*(10), 435–42.
- Xue, H., Lu, K., Hao, N. (2018). Cooperation makes two less-creative individuals turn into a highly-creative pair. *Neuroimage*, *172*, 527–37.
- Ye, J.C., Tak, S., Jang, K.E., Jung, J., Jang, J. (2009). Nirs-spm: Statistical parametric mapping for near-infrared spectroscopy. *Neuroimage*, *44*(2), 428–47.
- Zhang, X., Noah, J.A., Dravida, S., Hirsch, J. (2017). Signal processing of functional nirs data acquired during overt speaking. *Neurophotonics*, *4*(4), 041409.
- 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.
- Zhao, H., Tanikawa, Y., Gao, F., et al. (2002). Maps of optical differential pathlength factor of human adult forehead, somatosensory motor and occipital regions at multi-wavelengths in nir. *Physics in Medicine & Biology*, *47*(12), 2075.