Speaker–listener neural coupling underlies successful communication

Greg J. Stephens[a,b,1], Lauren J. Silbert[c,1], and Uri Hasson[c,d,2]

[a]Joseph Henry Laboratories of Physics, Princeton University, Princeton, NJ 08544; [b]Lewis–Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544; [c]Neuroscience Institute, Princeton University, Princeton, NJ 08544; and [d]Department of Psychology, Princeton University, Princeton, NJ 08540

Communicated by Charles G. Gross, Princeton University, Princeton, NJ, June 18, 2010 (received for review April 30, 2010)

Verbal communication is a joint activity; however, speech production and comprehension have primarily been analyzed as independent processes within the boundaries of individual brains. Here, we applied fMRI to record brain activity from both speakers and listeners during natural verbal communication. We used the speaker's spatiotemporal brain activity to model listeners' brain activity and found that the speaker's activity is spatially and temporally coupled with the listener's activity. This coupling vanishes when participants fail to communicate. Moreover, though on average the listener's brain activity mirrors the speaker's activity with a delay, we also find areas that exhibit predictive anticipatory responses. We connected the extent of neural coupling to a quantitative measure of story comprehension and find that the greater the anticipatory speaker–listener coupling, the greater the understanding. We argue that the observed alignment of production- and comprehension-based processes serves as a mechanism by which brains convey information.

functional MRI | intersubject correlation | language production | language comprehension

Verbal communication is a joint activity by which interlocutors share information (1). However, little is known about the neural mechanisms underlying the transfer of linguistic information across brains. Communication between brains may be facilitated by a shared neural system dedicated to both the production and the perception/comprehension of speech (1–7). Existing neurolinguistic studies are mostly concerned with either speech production or speech comprehension, and focus on cognitive processes within the boundaries of individual brains (1). The ongoing interaction between the two systems during everyday communication thus remains largely unknown. In this study we directly examine the spatial and temporal coupling between production and comprehension across brains during natural verbal communication.

Using fMRI, we recorded the brain activity of a speaker telling an unrehearsed real-life story and the brain activity of a listener listening to a recording of the story. In the past, recording speech during an fMRI scan has been problematic due to the high levels of acoustic noise produced by the MR scanner and the distortion of the signal by traditional microphones. Thus, we used a customized MR-compatible dual-channel optic microphone that cancels the acoustic noise in real time and achieves high levels of noise reduction with negligible loss of audibility (see *SI Methods* and Fig. 1*A*). To make the study as ecologically valid as possible, we instructed the speaker to speak as if telling the story to a friend (see *SI Methods* for a transcript of the story and Movie S1 for an actual sample of the recording). To minimize motion artifacts induced by vocalization during an fMRI scan, we trained the speaker to produce as little head movement as possible. Next, we measured the brain activity ($n = 11$) of a listener listening to the recorded audio of the spoken story, thereby capturing the time-locked neural dynamics from both sides of the communication. Finally, we used a detailed questionnaire to assess the level of comprehension of each listener.

Our ability to assess speaker–listener interactions builds on recent findings that a large portion of the cortex evokes reliable and selective responses to natural stimuli (e.g., listening to a story), which are shared across all subjects (8–11). These studies use the intersubject correlation method to characterize the similarity of cortical responses across individuals during natural viewing conditions (for a recent review, see ref. 12). Here we extend these ideas by looking at the direct interaction (not induced by shared external input) between brains during communication and test whether the speaker's neural activity during speech production is coupled with the shared neural activity observed across all listeners during speech comprehension.

We hypothesize that the speaker's brain activity during production is spatially and temporally coupled with the brain activity measured across listeners during comprehension. During communication, we expect significant production/comprehension couplings to occur if speakers use their comprehension system to produce speech, and listeners use their production system to process the incoming auditory signal (3, 13, 14). Moreover, because communication unfolds over time, this coupling will exhibit important temporal structure. In particular, because the speaker's production-based processes mostly precede the listener's comprehension-based processes, the listener's neural dynamics will mirror the speaker's neural dynamics with some delay. Conversely, when listeners use their production system to emulate and predict the speaker's utterances, we expect the opposite: the listener's dynamics will precede the speaker's dynamics (14). However, when the speaker and listener are simply responding to the same shared sensory input (both speaker and listener can hear the same utterances), we predict synchronous alignment. Finally, if the neural coupling across brains serves as a mechanism by which the speaker and listener converge on the same linguistic act, the extent of coupling between a pair of conversers should predict the success of communication.

Results

Speaker–Listener Coupling Model. We formed a model of the expected activity in the listeners' brains during speech comprehension based on the speaker's activity during speech production (see Fig. 1*B* and *Methods* for model details). Due to both the spatiotemporal complexity of natural language and an insufficient understanding of language-related neural processes, conventional hypothesis-driven fMRI analysis methods are largely unsuitable for modeling the brain activity acquired during communication. We therefore developed an approach that circumvents the need

Author contributions: G.J.S., L.J.S., and U.H. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

[1]G.J.S. and L.J.S. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: hasson@princeton.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1008662107/-/DCSupplemental.

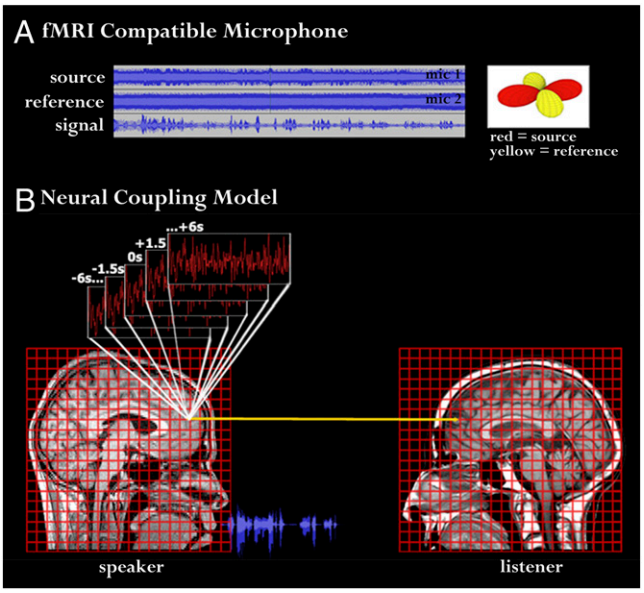


Fig. 1. Imaging the neural activity of a speaker–listener pair during storytelling. (*A*) To record the speaker's speech during the fMRI scan, we used a customized MR-compatible recording device composed of two orthogonal optic microphones (*Right*). The source microphone captures both the background noise and the speaker's speech utterances (upper audio trace), and the reference microphone captures the background noise (middle audio trace). A dual-adaptive filter subtracts the reference input from the source channel to recover the speech (lower audio trace). (*B*) The speaker–listener neural coupling was assessed through the use of a general linear model in which the time series in the speaker's brain are used to predict the activity in the listeners' brains. To capture the asynchronous temporal interaction between the speaker and the listeners, the speaker's brain activity was convolved with different temporal shifts. The convolution consists of both backward shifts (up to −6 s, intervals of 1.5 s, speaker precedes) and forward shifts (up to +6 s, intervals of 1.5 s, listener precedes) relative to the moment of vocalization (0 shift). For each brain area (voxel), the speaker's local response time course is used to predict the time series of the Talairach-normalized, spatially corresponding area in the listener's brain. The model thus captures the extent to which activity in the speaker's brain during speech production is coupled over time with the activity in the listener's brain during speech production.

to specify a formal model for the linguistic process in any given brain area by using the speaker's brain activity as a model for predicting the brain activity within each listener. To analyze the direct interaction of production and comprehension mechanisms, we considered only spatially local models that measure the degree of speaker–listener coupling within the same Talairach location. To capture the temporal dynamics, we first shifted the speaker's time courses backward (up to −6 s, intervals of 1.5 s, speaker precedes) and forward (up to +6 s, intervals of 1.5 s, listener precedes) relative to the moment of vocalization (0 shift). We then combined these nine shifted speaker time courses with linear weights to build a predictive model for the listener brain dynamics. Though correlations between shifted voxel time-courses can complicate the interpretation of the linear weights, here, these correlations are small as shown by the mean voxel autocorrelation function (Fig. S1). The weights are thus an approximately independent measure of the contribution of the speaker dynamics for each shift. To further ensure that the minimal autocorrelations among regressors did not affect the model's temporal discriminability, we decorrelated the regressors within the model and repeated the analysis. Similar results were obtained in both cases (*SI Methods*).

Speaker and Listener Brain Activity Exhibits Widespread Coupling During Communication. For each brain area, we identified significant speaker–listener couplings by applying an analytical F test to the overall model fit, and controlled for multiple comparisons across the volume using a fixed false discovery rate ($\gamma = 0.05$; see *Methods* for details). Similar results were obtained using a nonparametric permutation test (Fig. S2). Figure 2*A* presents the results in the left hemisphere; similar results were obtained in the right hemisphere (Fig. S3*A*). Significant speaker–listener coupling was found in early auditory areas (A1+), superior temporal gyrus, angular gyrus, temporoparietal junction (these areas are also known as Wernicke's area), parietal lobule, inferior frontal gyrus (also known as Broca's area), and the insula. Although the function of these regions is far from clear, they have been associated with various production and comprehension linguistic processes (15–17). Moreover, both the parietal lobule and the inferior frontal gyrus have been associated with the mirror neuron system (18). Finally, we also observed significant speaker–listener coupling in a collection of extralinguistic areas known to be involved in the processing of semantic and social aspects of the story (19), including the precuneus, dorsolateral prefrontal cortex, orbitofrontal cortex, striatum, and medial prefrontal cortex (see Table S1 for Talairach coordinates).

Brain areas that were coupled across the speaker and listener coincided with brain areas used to process incoming verbal information within the listeners (Fig. 2*B*). To compare the speaker–listener interactions (production/comprehension) with the listener–listener interactions (comprehension only), we constructed a listener–listener coupling map using similar analysis methods and statistical procedures as above. In agreement with previous work, the story evoked highly reliable activity in many brain areas across all listeners (8, 11, 12) (Fig. 2*B*, yellow). We note that the agreement with previous work is far from assured: the story here was both personal and spontaneous, and was recorded in the noisy environment of the scanner. The similarity in the response patterns across all listeners underscores a strong tendency to process incoming verbal information in similar ways. A comparison between the speaker–listener and the listener–listener maps reveals an extensive overlap (Fig. 2*B*, orange). These areas include many of the sensory-related, classic linguistic-related and extralinguistic-related brain areas, demonstrating that many of the areas involved in speech comprehension (listener–listener coupling) are also aligned during communication (speaker–listener coupling).

Speaker–Listener Neural Coupling Emerges Only During Communication. To test whether the extensive speaker–listener coupling emerges only when information is transferred across interlocutors, we blocked the communication between speaker and listener. We repeated the experiment while recording a Russian speaker telling a story in the scanner, and then played the story to non–Russian-speaking listeners ($n = 11$). In this experimental setup, although the Russian speaker is trying to communicate information, the listeners are unable to extract the information from the incoming acoustic sounds. Using identical analysis methods and statistical thresholds, we found no significant coupling between the speaker and the listeners or among the listeners. At significantly lower thresholds we found that the non–Russian-speaking listener–listener coupling was confined to early auditory cortices. This indicates that the reliable activity in most areas, besides early auditory cortex, depends on a successful processing of the incoming information, and is not driven by the low-level acoustic aspects of the stimuli.

As further evidence that extensive speaker–listener couplings rely on successful communication, we asked the same English speaker to tell another unrehearsed real-life story in the scanner. We then compared her brain activity while telling the second story with the brain activity of the listeners to the original story. In this experimental setup, the speaker transmits information and the

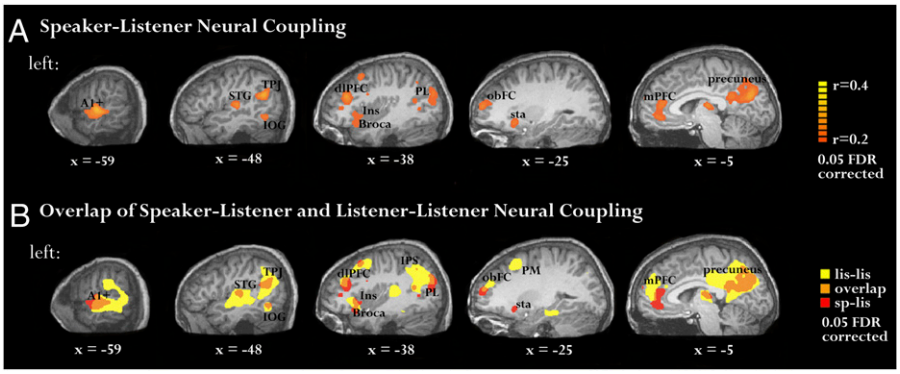


Fig. 2. The speaker–listener neural coupling is widespread, extending well beyond low-level auditory areas. (*A*) Areas in which the activity during speech production is coupled to the activity during speech comprehension. The analysis was performed on an area-by-area basis, with *P* values defined using an *F* test and was corrected for multiple comparisons using FDR methods ($\gamma = 0.05$). The findings are presented on sagittal slices of the left hemisphere (similar results were obtained in the right hemisphere; see Fig. S3). The speaker–listener coupling is extensive and includes early auditory cortices and linguistic and extra-linguistic brain areas. (*B*) The overlap (orange) between areas that exhibit reliable activity across all listeners (listener–listener coupling, yellow) and speaker–listener coupling (red). Note the widespread overlap between the network of brain areas used to process incoming verbal information among the listeners (comprehension-based activity) and the areas that exhibit similar time-locked activity in the speaker's brain (production/comprehension coupling). A1+, early auditory cortices; TPJ, temporal-parietal junction; dlPFC, dorsolateral prefrontal cortex; IOG, inferior occipital gyrus; Ins, insula; PL, parietal lobule; obFC, orbitofrontal cortex; PM, premotor cortex; Sta, striatum; mPFC, medial prefrontal cortex.

listener receives information; however, the information is decoupled across both sides of the communicative act. As in the Russian story, we found no significant coupling between the speaker and the listeners. We therefore conclude that coupling across interlocutors emerges only while engaged in shared communication.

Listeners' Brain Activity Mirrors the Speaker's Brain Activity with a Delay. Natural communication unfolds over time: speakers construct grammatical sentences based on thoughts, convert these to motor plans, and execute the plans to produce utterances; listeners analyze the sounds, build phonemes into words and sentences, and ultimately decode utterances into meaning. In our model of brain coupling, the speaker–listener temporal coupling is reflected in the model's weights, where each weight multiplies a temporally shifted time course of the speaker's brain activity relative to the moment of vocalization (synchronized alignment, zero shift). As expected, and in agreement with previous work (12), the activity among the listeners is time locked to the moment of vocalization (Fig. 3*A*, blue curve). In contrast, in most areas, the activity in the listeners' brains lagged behind the activity in the speaker's brain by 1–3 s (Fig. 3*A*, red curve). These lagged responses suggest that on average the speaker's production-based processes precede and likely induce the mirrored activity observed in the listeners' brains during comprehension. These findings also allay a methodological concern that the speaker–listener neural coupling is induced simply by the fact that the speaker is listening to her own speech.

Neural Couplings Display Striking Temporal Differences Across the Brain. The temporal dynamics of the speaker–listener coupling varied across brain areas (Fig. 3*B*). Among significantly coupled brain areas, important differences in dynamics are contained within the weights of the different temporally shifted regressors. To assess how these patterns varied across the brain, we categorized the weights as delayed (speaker precedes, −6 to −3 s), advanced (listener precedes, 3–6 s), or synchronous (−1.5 to 1.5 s). Though such categorizations increase statistical power, they also reduce the temporal resolution of the analysis. Thus, synchronous weights reflect processes that occur both at the point of vocalization (shift 0) as well as ±1.5 s around it, whereas delayed and advanced weights require shifts of over 1.5 s. Next, for each area we performed a contrast analysis to identify brain areas in which the mean weight for each temporal category is statistically greater ($P < 0.05$) than the mean weight over the rest of the couplings (*Methods*). In early auditory areas (A1+) the speaker–listener coupling is aligned to the speech utterances (synchronized alignment; Fig. 3*B*, yellow); in posterior areas, including the right TPJ and the precuneus, the speaker's brain activity preceded the listener's brain activity (speaker precedes; Fig. 3*B*, blue); in the striatum and anterior frontal areas, including the mPFC and dlPFC, the listener's brain activity preceded the speaker's brain activity (listener precedes; Fig. 3*B*, red). To verify that our categorization of temporal couplings was independent of autocorrelations within the speaker's time series, we repeated the analysis after decorrelating the model's regressors. We found nearly exact overlap between the delayed, synchronous, and advanced maps obtained with the original and decorrelated models (97%, 97%, and 94%, respectively). The result that significant speaker–listener couplings include substantially advanced weights may be indicative of predictive processes generated by the listeners before the moment of vocalization to enhance and facilitate the processing of the incoming, noisy speech input (14). Furthermore, the spatial specificity of the temporal coupling shows that it cannot simply be attributed to nonspecific, spatially global effects such as arousal. In comparison to the speaker–listener couplings, the comprehension-based processes in the listeners' brains were entirely aligned to the moment of vocalization (Fig. 3*C*, yellow). Thus, the dynamics of neural coupling between the speaker and the listeners are fundamentally different from the neural dynamics shared among all listeners.

Extent of Speaker–Listener Neural Coupling Predicts the Success of the Communication Humans use speech to convey information across brains. Here we administer a behavioral assessment to each listener at the end of the scan to assess the amount of information transferred from the speaker to each of the listeners (*SI Methods* and Fig. S4). We independently ranked both the listeners' behavioral scores and the spatial extent of significant neural coupling between the speaker and each listener, and found a strong positive correlation. ($r = 0.55$, $P < 0.07$; Fig. 4*A*). The correlation between the neural coupling and the level of comprehension was robust to changes in the exact statistical threshold and remained stable across many statically significant *P* values (Fig. S5*A*). These findings suggest that the stronger the neural coupling between interlocutors, the better the understanding. Finally, we computed behavioral correlations with brain regions that show coupling at

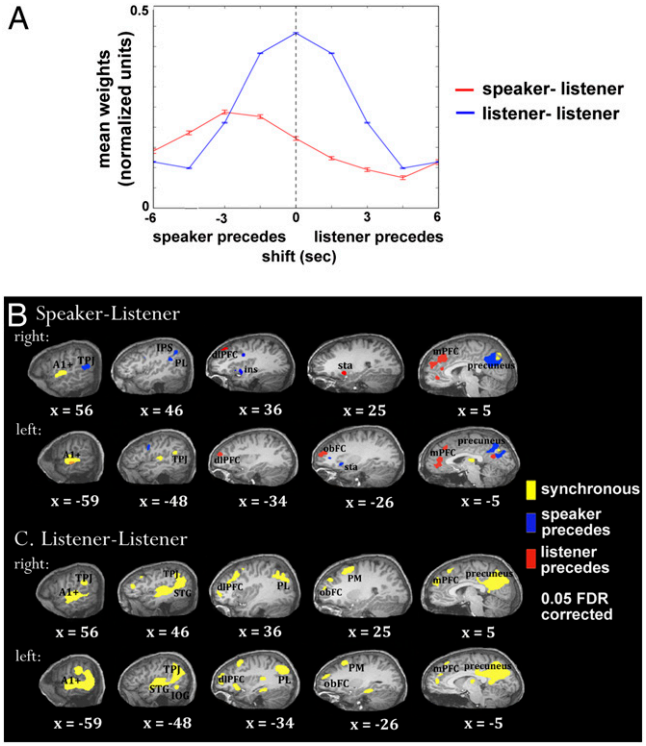


Fig. 3. Temporal asymmetry between speaker–listener and listener–listener neural couplings. (*A*) The mean distribution of the temporal weights across significantly coupled areas for the listener–listener (blue curve) and speaker–listener (red curve) brain pairings. For each area, the weights are normalized to unit magnitude, and error bars denote SEMs. The weight distribution within the listeners is centered on zero (the moment of vocalization). In contrast, the weight distribution between the speaker and listeners is shifted; activity in the listeners' brains lagged activity in the speaker's brain by 1–3 s. This suggests that on average the speaker's production-based processes precede and hence induce the listeners' comprehension-based processes. (*B*) The speaker–listener temporal coupling varies across brain areas. Based on the distribution of temporal weights within each brain area, we divided the couplings into three temporal profiles: the activity in speaker's brain precedes (blue); the activity is synchronized with ±1.5 s around the onset of vocalization (yellow), and the activity in listener's brain precedes (red). In early auditory areas, the speaker–listener coupling is time locked to the moment of vocalization. In posterior areas, the activity in the speaker's brain preceded the activity in the listeners' brains; in the mPFC, dlPFC, and striatum, the listeners' brain activity preceded. Results differ slightly in right and left hemisphere. (*C*) The listener–listener temporal coupling is time locked to the onset of vocalization (yellow) across all brain areas in right and left hemispheres. Note that unique speaker–listener temporal dynamics mitigates the methodological concern that the speaker's activity is similar to the listeners' activity due to the fact that the speaker is merely another listener of her own speech.

different delays (speaker precedes, synchronous, listener precedes). Using these temporal categories, we analyzed the connection between each category and the level of comprehension. Remarkably, the extent of cortical areas where the listeners' activity preceded the speaker's activity (red areas in Fig. 3*B*; contrast $P < 0.03$) provided the strongest correlation with behavior ($r = 0.75, P < 0.01$). This suggests that prediction is an important aspect of successful communication. Furthermore, the behavioral correlation in both cases increases to $r = 0.76$ ($P < 0.01$) and $r = 0.93$ ($P < 0.0001$), respectively, when we remove a single outlier listener (ranked eighth in Fig. 4 *A* and *B*). Finally, the correlation between comprehension and neural coupling was robust to changes in the exact contrast threshold (Fig. S5*B*). Importantly, we note that the correlation with the level of understanding cannot be attributed to low-level processes (e.g., the audibility of the audiofile), as the

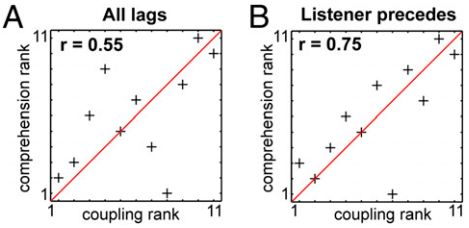


Fig. 4. The greater the extent of neural coupling between a speaker and listener the better the understanding. (*A*) To assess the comprehension level of each individual listener, an independent group of raters ($n = 6$) scored the listeners' detailed summaries of the story they heard in the scanner. We ranked the listeners' behavioral scores and the extent of significant speaker–listener coupling and found a strong positive correlation ($r = 0.54$, $P < 0.07$) between the amount of information transferred to each listener and the extent of neural coupling between the speaker and each listener (Fig. 4*A*). These findings suggest that the stronger the neural coupling between interlocutors, the better the understanding. (*B*) The extent of brain areas where the listeners' activity preceded the speaker's activity (red areas in Fig. 3*B*) provided the strongest correlation with behavior ($r = 0.75$, $P < 0.01$). These findings provide evidence that prediction is an important aspect of successful communication.

correlation with behavior increases when we do not include early auditory areas (synchronous alignment, yellow areas in Fig. 3*B*).

Discussion

Communication is a shared activity resulting in a transfer of information across brains. The findings shown here indicate that during successful communication, speakers' and listeners' brains exhibit joint, temporally coupled, response patterns (Figs. 2 and 3). Such neural coupling substantially diminishes in the absence of communication, such as when listening to an unintelligible foreign language. Moreover, more extensive speaker–listener neural couplings result in more successful communication (Fig. 4). We further show that on average the listener's brain activity mirrors the speaker's brain activity with temporal delays (Fig. 3 *A* and *B*). Such delays are in agreement with the flow of information across communicators and imply a causal relationship by which the speaker's production-based processes induce and shape the neural responses in the listener's brain. Though the sluggish BOLD response masks the exact temporal speaker–listener coupling, the delayed and advanced timescales (~1–4 s) coincide with the timescales of some rudimentary linguistic processes (e.g., in this study, it took the speaker on average around 0.5 ± 0.6 s to produce words, 9 ± 5 s to produce sentences, and even longer to convey ideas). Moreover, we recently demonstrated that some high-order brain areas, such as the TPJ and the parietal lobule, have the capacity to accumulate information over many seconds (8).

Our analysis also identifies a subset of brain regions in which the activity in the listener's brain *precedes* the activity in the speaker's brain. The listener's anticipatory responses were localized to areas known to be involved in predictions and value representation (20–23), including the striatum and medial and dorsolateral prefrontal regions (mPFC, dlPFC). The anticipatory responses may provide the listeners with more time to process an input and can compensate for problems with noisy or ambiguous input (24). This hypothesis is supported by the finding that comprehension is facilitated by highly predictable upcoming words (25). Remarkably, the extent of the listener's anticipatory brain responses was highly correlated with the level of understanding (Fig. 4*B*), indicating that successful communication requires the active engagement of the listener (26, 27).

The notion that perception and action are coupled has long been argued by linguists, philosophers, cognitive psychologists, social psychologists, and neurophysiologists (2, 3, 24, 28–33). Our findings document the ongoing dynamic interaction between two

brains during the course of natural communication, and reveal a surprisingly widespread neural coupling between the two, a priori independent, processes. Such findings are in agreement with the theory of interactive linguistic alignment (1). According to this theory, production and comprehension become tightly aligned on many different levels during verbal communication, including the phonetic, phonological, lexical, syntactic, and semantic representations. Accordingly, we observed neural coupling during communication at many different processing levels, including low-level auditory areas (induced by the shared input), production-based areas (e.g., Broca's area), comprehension-based areas (e.g., Wernicke's area and TPJ), and high-order extralinguistic areas (e.g., precuneus and mPFC) that can induce shared contextual model of the situation (34). Interestingly, some of these extralinguistic areas are known to be involved in processing social information crucial for successful communication, including, among others, the capacity to discern the beliefs, desires, and goals of others (15, 16, 31, 35–38)

The production/comprehension coupling observed here resembles the action/perception coupling observed within mirror neurons (35). Mirror neurons discharge both when a monkey performs a specific action and when it observes the same action performed by another (39). Similarly, during the course of communication the production-based and comprehension-based processes seem to be tightly coupled to each other. Currently, however, direct proof of such a link remains elusive for two main reasons. First, mirror neurons have been recorded mainly in the ventral premotor area (F5) and the intraparietal area (PF/IPL) of the primate brain during observation and execution of rudimentary motor acts such as reaching or grabbing food. The speaker–listener neural coupling observed here extends far beyond these two areas. Furthermore, although area F5 in the macaque has been suggested to overlap with Broca's area in humans, a detailed characterization of the links between basic motor acts and complex linguistic acts is still missing (see refs. 40 and 41). Second, based on the fMRI activity recorded during production and comprehension of the same utterances, we cannot tell whether the speaker–listener coupling is generated by the activity of the same neural population that produces and encodes speech or by the activity of two intermixed but independent populations (42). Nevertheless, our findings suggest that, on the systems level, the coupling between action-based and perception-based processes is extensive and widely used across many brain areas.

The speaker–listener neural coupling exposes a shared neural substrate that exhibits temporally aligned response patterns across communicators. Previous studies have shown that during free viewing of a movie or listening to a story, the external shared input can induce similar brain activity across different individuals (8–11, 43, 44). Verbal communication enables us to convey information across brains, independent of the actual external situation (e.g., telling a story of past events). Such phenomenon may be reflected in the ability of the speaker to directly induce similar brain patterns in another individual, via speech, in the absence of any other stimulation. Finally, the recording of the neural activity from both the speaker brain and the listener brain opens a new window into the neural basis of interpersonal communication, and may be used to assess verbal and nonverbal forms of interaction in both human and other model systems (45). Further understanding of the neural processes that facilitate neural coupling across interlocutors may shed light on the mechanisms by which our brains interact and bind to form societies.

Methods

Subject Population. One native-English speaker, one native-Russian speaker, and 12 native-English listeners, ages 21–30 y, participated in one or more of the experiments. Procedures were in compliance with the safety guidelines for MRI research and approved by the Princeton University Committee on Activities Involving Human Subjects. All participants provided written informed consent.

Experiment and Procedure. To measure neural activity during communication, we first used fMRI to record the brain activity of a speaker telling a long, unrehearsed story. The speaker had three practice sessions inside the scanner telling real-life unrehearsed stories. This allowed for the opportunity for the speaker to familiarize herself with the conditions of storytelling inside the scanner and to learn to minimize head movements without compromising storytelling effectiveness. In the final fMRI session, the speaker told a new, nonrehearsed, real-life, 15-min account about an experience she had as a freshman in high school (see *SI Methods* for the transcript). The story was recorded using a MR-compatible microphone (see below). The speech recording was aligned with the scanner's TTL backtick received at each TR. The same procedure was followed for the Russian speaker, telling a nonrehearsed, real-life story in Russian. In focusing on a personally relevant experience, we strove both to approach the ecological setting of natural communication and to ensure an intention to communicate by the speaker.

We measured listeners' brain activity during audio playback of the recorded story. We synchronized the functional time series to the speaker's vocalization through the use of a Matlab code (MathWorks Inc.) written to start the speaker's recording at the onset of the scanner's TTL backtick. Eleven listeners listened to the recording of the English story. Ten of the listeners (and one new subject) listened to the recording of the Russian story. None of the listeners understood Russian. Our experimental design thus allows access to both sides of the simulated communication. Participants were instructed before the scan to attend as best as possible to the story, and further that they would be asked to provide a written account of the story immediately following the scan.

Recording System. We recorded the speaker's speech during the fMRI scan using a customized MR-compatible recording system (FOMRI II; Optoacoustics Ltd.). More details are described in *SI Methods* and in Fig. 1*A*.

MRI Acquisition. Subjects were scanned in a 3T head-only MRI scanner (Allegra; Siemens). More details are described in *SI Methods*.

Data Preprocessing. fMRI data were preprocessed with the BrainVoyager software package (Brain Innovation, version 1.8) and with additional software written with Matlab. More details are described in *SI Methods*.

Model Analysis. The coupling between speaker–listener and listener–listener brain pairings was assessed through the use of a spatially local general linear model in which temporally shifted voxel time series in one brain are linearly summed to predict the time series of the spatially corresponding voxel in another brain. Thus for the speaker–listener coupling we have

$$v^{\text{model}}_{\text{listener}}(t) = \sum_{\tau=-\tau_{\max}}^{\tau=\tau_{\max}} \beta_i v_{\text{speaker}}(t+\tau), \qquad [1]$$

where the weights $\vec{\beta}$ are determined by minimizing the RMS error and are given by $\vec{\beta} = (c)^{-1}\langle \vec{v}\, v_{listener}\rangle$. Here, C is the covariance matrix $c_{mn} = \langle v_m v_n\rangle$ and $\vec{v}$ is the vector of shifted voxel times series, $v_m = v_{\text{speaker}}(t-m)$.We choose $\tau_{\max} = 4$, which is large enough to capture important temporal processes while also minimizing the overall number of model parameters to maintain statistical power. We obtain similar results with $\tau_{\max} = (3, 5)$.

We calculated the neural couplings for three brain pairings: (a) speaker, individual listener; (b) speaker, average listener; and (c) and listener, average listener. In all three cases, the first brain in the pairing provides the independent variables in Eq. **1**. The average listener dynamics was constructed by averaging the functional time series of the (n = 11) listeners at each location in the brain. The (listener, average listener) pairing was constructed by first building, for each listener, the [listener, (N–1) listener average] pairing. For each listener, the (N–1) average listener is the average listener constructed from all other listeners. We then solved the coupling model (Eq. **1**). Finally, to connect our findings to behavioral variability, we constructed the (speaker–listener) coupling separately for the N individual listeners.

Statistical Analysis. We indentified statistically significant couplings by assigning P values through a Fisher's F test. In detail, the model in Eq. **1** has $\delta_{\text{model}} = 9$ degrees of freedom, while $\delta_{\text{null}} = T - \delta_{\text{model}} - 1$, where T is the number of time points in the experiment. For the prom story, T = 581, and T = 451 for the Russian story. For each model fit we construct the F statistic and associated P value $p = 1 - f(F, \delta_{\text{model}}, \delta_{\text{null}})$ where f is the cumulative distribution function of the F statistic. We also assigned nonparametric P values by using a null model based on randomly permuted data (n = 1,000) at each brain location. The nonparametric null model produced P values very close

to those constructed from the *F* statistic (Fig. S2). We correct for multiple statistical comparisons when displaying volume maps by controlling the false discovery rate (FDR). Following ref. 46, we place the *P* values in ascending order $(p_1 \ldots p_q^* \ldots p_{n_{vox}})$ and choose the maximum value p_q^* such that $p_q^* q / n_{vox} < \gamma$, where $\gamma = 0.05$ is the FDR threshold.

To identify significant listener–listener couplings, we applied the above statistical analysis to the model fits across all (n = 11) listener–average listener pairs. This is a statistically conservative approach aimed to facilitate comparison with the speaker–listener brain pairing. The greater statistical power contained within the (n = 11) different listener/average listener pairs can be fully exploited using nonparametric bootstrap methods for estimating the null distribution and calculating significant couplings (Fig. S6).

Coupling Categorization. For each cortical location and brain pairing, the parameters $\vec{\beta}$ of Eq. **1** fully characterize the local neural coupling. As seen in Fig. 3, temporal differences among the couplings reveal important differences between action and perception. To explore these differences we categorize each coupling as delayed (speaker precedes, −6 to −3 s), advanced (listener precedes 3–6 s), or synchronous (−1.5 to 1.5 s) based on the difference of the mean weight within each category relative to the mean weight outside the category. We define the statistical significance of each category through a contrast analysis. For example, for the delayed category we define the contrast $\vec{c} = (2, 2, 2, -1, -1, -1, -1, -1, -1)$ and associated t statistic $t = \langle \vec{c} \cdot \vec{\beta} \rangle / \sqrt{(1 - r^2) \cdot (\vec{c}^{\,\mathrm{T}} \cdot \mathrm{cov}(\vec{v}) \cdot \vec{c})}$, where r^2 is the model variance and $\mathrm{cov}(\vec{v})$ is the covariance of the shifted time-series. The contrasts for other categories are defined similarly. A large contrast indicates that the coupling is dominated by weights within that particular temporal category.

Behavioral Assessment. Immediately following the scan, the participants were asked to record the story they heard in as much detail as possible. Six independent raters scored each of these listener records accordingly, and the resulting score was used as a quantitative and objective measure of the listener's understanding. More details are provided in *SI Methods* and Fig. S4.

ACKNOWLEDGMENTS. We thank our colleagues Forrest Collman, Yadin Dudai, Bruno Galantucci, Asif Ghazanfar, Adele Goldberg, David Heeger, Chris Honey, Ifat Levy, Yulia Lerner, Rafael Malach, Stephanie E. Palmer, Daniela Schiller, and Carrie Theisen for helpful discussion and comments on the manuscript. G.J.S. was supported in part by the Swartz Foundation.

1. Pickering MJ, Garrod S (2004) Toward a mechanistic psychology of dialogue. *Behav Brain Sci*, 27:169–190 , discussion 190–226.
2. Hari R, Kujala MV (2009) Brain basis of human social interaction: From concepts to brain imaging. *Physiol Rev* 89:453–479.
3. Liberman AM, Mattingly IG (1985) The motor theory of speech perception revised. *Cognition* 21:1–36.
4. Branigan HP, Pickering MJ, Cleland AA (2000) Syntactic co-ordination in dialogue. *Cognition* 75:B13–B25.
5. Levelt WJM (1989) *Speaking: From Intention to Articulation* (MIT Press, Cambridge, MA).
6. Wilson M, Knoblich G (2005) The case for motor involvement in perceiving conspecifics. *Psychol Bull* 131:460–473.
7. Chang F, Dell GS, Bock K (2006) Becoming syntactic. *Psychol Rev* 113:234–272.
8. Hasson U, Yang E, Vallines I, Heeger DJ, Rubin N (2008) A hierarchy of temporal receptive windows in human cortex. *J Neurosci* 28:2539–2550.
9. Golland Y, et al. (2007) Extrinsic and intrinsic systems in the posterior cortex of the human brain revealed during natural sensory stimulation. *Cereb Cortex* 17:766–777.
10. Hasson U, Nir Y, Levy I, Fuhrmann G, Malach R (2004) Intersubject synchronization of cortical activity during natural vision. *Science* 303:1634–1640.
11. Wilson SM, Molnar-Szakacs I, Iacoboni M (2008) Beyond superior temporal cortex: Intersubject correlations in narrative speech comprehension. *Cereb Cortex* 18: 230–242.
12. Hasson U, Malach R, Heeger D (2010) Reliability of cortical activity during natural stimulation. *Trends Cogn Sci* 14:40–48.
13. Galantucci B, Fowler CA, Turvey MT (2006) The motor theory of speech perception reviewed. *Psychon Bull Rev* 13:361–377.
14. Pickering MJ, Garrod S (2007) Do people use language production to make predictions during comprehension? *Trends Cogn Sci* 11:105–110.
15. Fletcher PC, et al. (1995) Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition* 57:109–128.
16. Völlm BA, et al. (2006) Neuronal correlates of theory of mind and empathy: A functional magnetic resonance imaging study in a nonverbal task. *Neuroimage* 29: 90–98.
17. Sahin NT, Pinker S, Cash SS, Schomer D, Halgren E (2009) Sequential processing of lexical, grammatical, and phonological information within Broca's area. *Science* 326: 445–449.
18. Fadiga L, Craighero L, D'Ausilio A (2009) Broca's area in language, action, and music. *Ann N Y Acad Sci* 1169:448–458.
19. Xu J, Kemeny S, Park G, Frattali C, Braun A (2005) Language in context: Emergent features of word, sentence, and narrative comprehension. *Neuroimage* 25: 1002–1015.
20. Koechlin E, Corrado G, Pietrini P, Grafman J (2000) Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. *Proc Natl Acad Sci USA* 97: 7651–7656.
21. Krueger F, Grafman J (2008) The human prefrontal cortex stores structured event complexes. *Understanding Events: From Perception to Action*, eds Shipley T, Zacks JM (Oxford Univ Press, New York), pp 617–638.
22. Gilbert SJ, et al. (2006) Functional specialization within rostral prefrontal cortex (area 10): A meta-analysis. *J Cogn Neurosci* 18:932–948.
23. Craig AD (2009) How do you feel—now? The anterior insula and human awareness. *Nat Rev Neurosci* 10:59–70.
24. Garrod S, Pickering MJ (2004) Why is conversation so easy? *Trends Cogn Sci* 8:8–11.
25. Schwanenflugel PJ, Shoben EJ (1985) The influence of sentence constraint on the scope of facilitation for upcoming words. *J Mem Lang* 24:232–252.
26. Clark HH (1996) *Using Language* (Cambridge Univ Press, Cambridge, UK).
27. Clark HH, Wilkes-Gibbs D (1986) Referring as a collaborative process. *Cognition* 22: 1–39.
28. Galantucci B, Fowler CA, Goldstein L (2009) Perceptuomotor compatibility effects in speech. *Atten Percept Psychophys* 71:1138–1149.
29. Merleau-Ponty M (1945) *The Phenomenology of Perception* (Gallimard, Paris, France).
30. Gibson JJ (1979) *The Ecological Approach to Visual Perception* (Houghton Mifflin, Boston).
31. Amodio DM, Frith CD (2006) Meeting of minds: The medial frontal cortex and social cognition. *Nat Rev Neurosci* 7:268–277.
32. Rizzolatti G, Fadiga L, Gallese V, Fogassi L (1996) Premotor cortex and the recognition of motor actions. *Brain Res Cogn Brain Res* 3:131–141.
33. Arbib M (2010) Mirror system activity for action and language is embedded in the integration of dorsal and ventral pathways. *Brain Lang* 112:12–24.
34. Johnson-Laird PN (1995) *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness* (Harvard Univ Press, Cambridge, MA).
35. Gallagher HL, et al. (2000) Reading the mind in cartoons and stories: An fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia* 38:11–21.
36. Gallagher HL, Frith CD (2003) Functional imaging of 'theory of mind'. *Trends Cogn Sci* 7:77–83.
37. Saxe R, Carey S, Kanwisher N (2004) Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annu Rev Psychol* 55: 87–124.
38. Saxe R, Kanwisher N (2003) People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *Neuroimage* 19:1835–1842.
39. Rizzolatti G, Fogassi L, Gallese V (2001) Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat Rev Neurosci* 2:661–670.
40. Rizzolatti G, Arbib MA (1998) Language within our grasp. *Trends Neurosci* 21: 188–194.
41. Arbib MA, Liebal K, Pika S (2008) Primate vocalization, gesture, and the evolution of human language. *Curr Anthropol*, 49:1053–1063 , discussion 1063–1076.
42. Dinstein I, Thomas C, Behrmann M, Heeger DJ (2008) A mirror up to nature. *Curr Biol* 18:R13–R18.
43. Hanson SJ, Gagliardi AD, Hanson C (2009) Solving the brain synchrony eigenvalue problem: Conservation of temporal dynamics (fMRI) over subjects doing the same task. *J Comput Neurosci* 27:103–114.
44. Jääskeläinen IP, et al. (2008) Inter-subject synchronization of prefrontal cortex hemodynamic activity during natural viewing. *Open Neuroimaging J* 2:14–19.
45. Schippers MB, Roebroeck A, Renken R, Nanetti L, Keysers C (2010) Mapping the information flow from one brain to another during gestural communication. *Proc Natl Acad Sci USA* 107:9388–9393.
46. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc B* 57:289–300.

Supporting Information

Stephens et al. 10.1073/pnas.1008662107

SI Methods

Recording System. We recorded the speaker's speech during the fMRI scan using a customized MR-compatible recording system (FOMRI II; Optoacoustics Ltd.). The MR recording system uses two orthogonally oriented optical microphones. The typical 3D spherical polarity of the dual-channel sensor is shown in Fig. 1*A*. The reference microphone captures the background noise, and the source microphone captures both background noise and the speaker's speech utterances (signal). A dual-adaptive filter subtracts the reference input from the source channel (using least mean square approach). To achieve an optimal subtraction, the reference signal is adaptively filtered where the filter gains are learned continuously from the residual signal and the reference input. To prevent divergence of the filter when speech is present, a voice activity detector (VAD) is integrated into the algorithm. An additional speech-enhancement spectral-filtering algorithm further preprocesses the speech output to achieve a real-time speech enhancement.

MRI Acquisition. Subjects were scanned in a 3T head-only MRI scanner (Allegra; Siemens). A custom radiofrequency coil was used for the structural scans (NM-011 transmit head coil; Nova Medical). For fMRI scans, a time series of volumes was acquired using a T2*-weighted EPI pulse sequence [repetition time (TR) = 1,500 ms; echo time (TE) = 30 ms; flip angle = 80°]. The volume included 25 slices of 3-mm thickness with 1-mm interslice gap (in-plane resolution = 3×3 mm^2). T1-weighted high-resolution ($1 \times 1 \times 1$ mm) anatomical images were acquired for each observer with an MPRAGE pulse sequence to allow accurate cortical segmentation and 3D surface reconstruction. To minimize head movement, subjects' heads were stabilized with foam padding. Stimuli were presented using Psychophysics Toolbox in Matlab. High-fidelity MRI-compatible headphones (MR Confon) were fitted to provide considerable attenuation to the scanner noise and to present the audio stimuli to the subjects.

Data Preprocessing. fMRI data were preprocessed with the BrainVoyager software package (Brain Innovation, version 1.8) and with additional software written with Matlab. Preprocessing of functional scans included linear trend removal and high-pass filtering (up to six cycles per experiment). To correct for head motion, we used a 3D algorithm that adjusts for small head movements by rigid body transformations of all slices to the first reference volume. Detected head motions were less than 1 mm in size, which is well within the range of typical movements observed in other imaging studies. All functional images were transformed into a shared Talairach coordinate system so that corresponding brain regions are roughly spatially aligned. To further overcome misregistration across subjects, the data were spatially smoothed with a Gaussian filter of 6-mm full width at half-maximum value. To remove transient nonspecific signal elevation effects at the beginning of the experiment, and some preprocessing artifacts at the edges of the time courses, we excluded the first 15 and the last three timepoints of the experiments.

Autocorrelations. In principle, the correlations between shifted voxel time courses can complicate the interpretation of the $\vec{\beta}$ weights. In practice, these correlations are small, as shown by the mean voxel autocorrelation function, averaged across all brain locations in the speaker (Fig. S1). The autocorrelations are strongest for neighboring timepoints, but are small overall and decay rapidly. In addition, we decorrelated the model speaker regressors and repeated the coupling categorization analysis reported below with no substantial changes. The weights β_i are thus an approximate independent measure of the contribution of the speaker dynamics for each shift. The stability of the model was assessed by calculating the condition number $c = \lambda_1/\lambda_2$ of the covariance matrix, where λ_1 and λ_2 are the largest and smallest eigenvalues, respectively. Across the entire brain volume we find $c_{max} = 60$, which is well below machine precision limits.

Behavioral Assessment. Immediately following the scan, the participants were asked to record the story they heard in as much detail as possible. Six independent raters scored each of these listener records according to a 115-point true/false questionnaire (*SI Methods*), and the resulting score was used as a quantitative and objective measure of the listener's understanding. Among the raters, there was little variance in the scores assigned to each listener (Fig. S4), and thus the variance in behavioral scores among listeners accurately reflects the variability of their natural experience. To connect behavioral scores to neural activity, we analyzed the couplings derived from the ($n = 11$) pairings between the speaker and the individual listeners. For each listener, we defined the spatial extent of the speaker–listener coupling in two distinct ways: (*i*) the number of significant couplings identified through Eq. **1**, and (*ii*) the number of significantly advanced couplings (a subset of *i* with advanced contrast *P* value, $P < 0.03$). In each case, we then ranked both the listener's behavioral scores and the extent of significant speaker–listener coupling and computed the correlation between ranks.

A Transcript of the Story. I am going to tell you a story about my prom experience. I know everybody has some crazy prom stories, but, well, just wait. I was a freshman in high school in Miami, Florida, and I'm new to the freshman scene. I'm new to the high school scene, I should say, and it's almost December so I've been in high school for about 3 months, and this boy Charles asks me out. He's British, he's a junior, and he's really cute but sort of shy but just, well, it doesn't matter. So I say yes, I'm excited. He comes and picks me up in this really old, beat-up white Toyota, and he has two tapes in his car: Queen and Oasis. And so I ask him if he only listens to English bands when I get in the car and he gets a little offended. So I knew that, um, it just wasn't gonna work out with us. But so we go out and it's sort of awkward and whatever, and nothing really happens. And then we actually go out again and it's also awkward. And then the next day at school, this is like two weeks later, we're sort of with other people talking about prom, which is now months and months away. And I leave to go to class, but he comes with me and we're still sort of talking about prom, in like a general sense about what prom is more than, you know, specifics about this prom. But for some reason, out of the blue, he asks me to go to his prom with him. And, I mean, I'm a freshman girl, and it's really exciting, I never would have thought of going to the prom as a freshman, so of course I say yes, not really thinking about it. I think it's pretty weird but I just go with it. Um, and that's sort of that. And then that weekend I was at a party on the beach at this big house that was like the party house of my high school, at least freshman and sophomore year because his house was right on the beach and so the police were not roaming around those beaches. Um, anyway, so I'm at this party and I'm getting a beer out of the keg and I had never really done that much so of course I'm horrible at it and it's like 85% foam, and I notice that the guy behind me is this guy Amir. He's this really good-looking senior and well, my high school was a really big swimming high school. A lot of people went there

for swimming from all over the world, and so it was half a boarding school also, so there was a boys' dorm and a girls' dorm. Amir was a swimmer and he lived in the boys' dorm. And he was sort of a popular senior guy. And so he takes over and pours me this beer, of course perfectly foamless. And as he's passing it back to me, the guy behind me gets pushed or something, and pushes into me, and of course I spill the beer all over Amir, and that's how we meet. And it's sort of this high school romance fantasy, where we were just totally smitten. And so we start going out, and it's really fun. I don't have my license at this point, I have a permit, and so he teaches me how to drive. He lives in the boys' dorm and so his friends and him put me in a box, a cardboard box, and wrapped it up, and actually carried me, snuck me into the dorm as like a package. It was very funny, it was all fun, I was very happy. And then so we've been together for over two months or so and it comes time for people to kind of start talking about prom. So at this point I've basically forgotten that Charles was even there and just assumed that we both realized that it wasn't really working out and went our separate ways. And he knew that I was now with Amir and I just assumed it was fine. But so I'm at school and I'm sitting in the quad and Charles comes over to me and he sits down and says 'I just wanted to you know discuss with you our plans for prom'. And I'm like 'um, I, um . . .' I just didn't know what to say. So, I feel really badly about it, he obviously was, uh, hurt, and not going to let it go. And I felt badly, so I decide that I'm going to go to the prom with Charles and just meet up with Amir after. So then the day of the prom comes around, and on top of all of this sort of drama with this boy situation, my family—we're a big sailing family. We like to sail to see places, or maybe vice versa. So we take a lot of trips on sailboats, just the family sailing, like around Greece or France or wherever. And that particular summer we had planned a trip to go sailing in the British Virgin Islands, which is really known for its scuba diving. So we all decide to get certified for scuba diving. And of course the only weekend we could do it—it's like a two-weekend consecutive course—the only weekends we can do it are the weekend before the prom and the weekend of the prom. So the first weekend is all sort of in the pool, and the second weekend you actually go out into the ocean and do a real dive off of Miami Beach. And we left at like 8 in the morning, we're supposed to get back at three. So I think ok, this is fine, I'll get back home by like 4 and I'll have like two hours to get ready for this prom. Which is a really big deal, especially as a freshman girl, I was really nervous about prom in general. So we go scuba diving, and that's a whole nother story, my family is just really crazy on this boat. My mother is terrified, literally terrified, but just trying to stay calm, and when she's under any sort of pressure she sort of reverts back to this very thick New York accent, so she's just freaking out about her tank or whatever in this thick New York accent. And my father, who's sort of big tall and prominent, tough, manly man, and he's completely seasick, just trying not to throw up the whole time. Um, but, it was just really this crazy experience. So anyway, we're getting back, we've scuba dived, we've gone through the whole thing, and of course the boat breaks down. So we have to wait for like a tugboat to come bring us to shore, and we finally get back, and we're pulling into my house at like 6 o'clock, like two and a half hours late, just as Charles, who's always on time, of course, is pulling up. We pull up together. And, I don't know if you've ever been scuba diving, but pretty much the worst you'll ever look is after you go scuba diving. You've been under 60 feet of water, which is two atmospheres of pressure, for an hour and a half. You have a goggle mark permanently sketched into your face, which takes like 5 hours to get rid of that. And um, just your hair, it's just a mess, you're just a mess. And now I have approximately 5 minutes to get ready for the prom. So I'm like trying to put on make-up while my sister is shaving my legs, while my mom is brushing my hair. And I sort of put it all together as quickly as possible, throw on a dress, and leave. So now I'm on my way to the prom with Charles and it's very awkward, very awkward. We go to this dinner thing at one of his friends and it's just sort of an awkward situation in general. But we finally get to the prom and I'm looking around for Amir, and I can't find him anywhere. So I go to the bathroom, and I come back and I hear shouting in the back corner. And Amir is just totally drunk and is starting a fight with Charles. He's like about to hit him. And I run over and I grab Amir and we leave. And it's fine and we decide to go to the after party pretty much immediately. So we're walking to the car, and Amir had drove but he obviously couldn't drive 'cause he was drunk, so I drive. And Amir is going through his pockets to find his keys to give me the keys and he trips in the parking lot over one of those parking space indicator things, and his hands are in his pockets and he's not fast enough when he's drunk to pull his hands out so he just falls flat on his face. And he looks up and there's blood running out of his nose and all over and he's just a bloody mess. So I run back to the bathroom and I get paper towels and whatnot and I clean him up a bit. And we get in the car and just start driving and I'm thinking this has been the most absurd day, but of course I thought that way too soon because we're driving and Amir is like playing air guitar or something in the passenger seat and I'm just trying to get us there and get out of the car. And so I'm driving on US1, and there's two lanes going one way and two lanes going the other, and every once in a while there's a turning lane. And there was a car accident in one of these turning lanes. And I'm driving slowly, there's lights and traffic and stuff so I'm not going quickly at all. And there was this fender bender and Amir, I'm not exactly sure what he does, but he sort of grabs my arm and I sort of turn, and I end up crashing into this accident that's already there. And so, um, so by already there I mean like nobody's hurt, it was a very light light hit, no cars were ruined or anything, but the police are already there, and all of the people are already there, like they're all watching it happen. So the policeman comes over, and I don't have my license, by the way. I have a permit so I can be driving with Amir cause he's 18, but he's wasted, I mean, he's still playing air guitar when the policeman comes over and his face is bloody. And so I give him my permit and the registration and he's like uh I need his license, so I give him Amir's license as well. And I'm thinking I'm done for. I'm thinking I'm going to jail for diving without a license with this drunk dude and I'm thinking it's all over for me. So the policeman leaves and I'm really freaking out. And he comes back like 20 minutes later, and, I'm not lying, he hands me back my license and he says to me, 'I'm sorry Miss Silbert, but somehow your registration has blown away and I cannot find it'. And he was so embarrassed about losing my registration that he tells me to drive on my way. For real. So I drive away as quickly as I can and I finally get to the after party house and I get out of the car and Amir passes out on the beach and I call my mom and have her pick me up. And that's it.

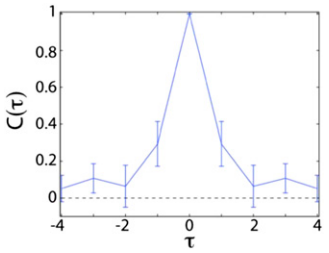


Fig. S1. Decay of temporal correlations. We show the voxel autocorrelation function averaged across all brain regions. Error bars denote SDs. Temporal correlations are strongest for neighboring timepoints (±1) but are small overall and decay rapidly in time.

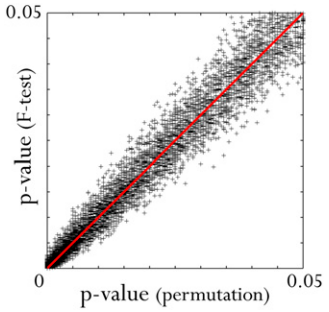


Fig. S2. The equivalence between *F* test (parametric) and shuffled (nonparametric) hypothesis testing in our study. We plot model *P* values obtained by an *F* test against *P* values obtained although a null model constructed by random permutation (n = 1,000) of the time series.

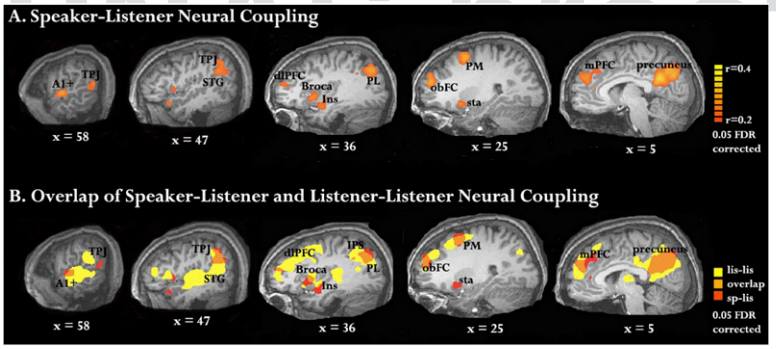


Fig. S3. The speaker–listener neural coupling within the right hemisphere. Figure layout is identical to that in Fig. 2*A*.

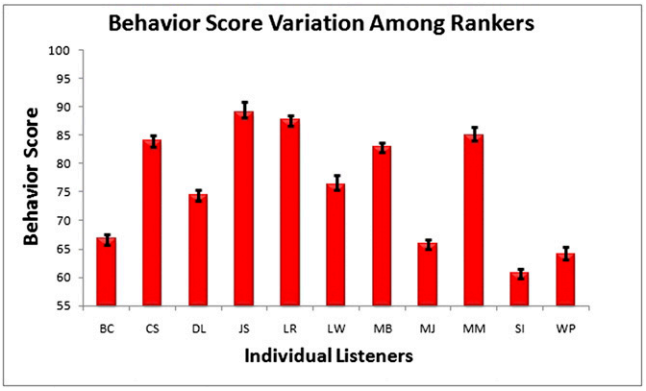


Fig. S4. The quantitative measure of story comprehension varies little among the different rankers who score the written answers. Error bars on the average score for each listener reflect the SD among the rankers.

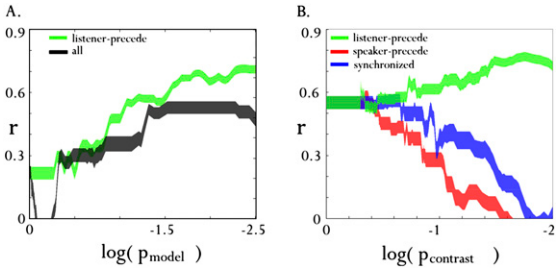


Fig. S5. The rank correlation between the extent of neural coupling and story comprehension grows with both temporal asymmetry (smaller contrast *P* values) and overall significance threshold (smaller model *P* values). (*A*) Plot of the activation/understanding rank correlation as a function of the model *P* values (black) and listener precedes (green) couplings (contrast *P* value fixed to $P < 0.03$). (*B*) Plot of the activation/understanding rank correlation as a function of the temporal asymmetry (contrast *P* value) for speaker precedes (red), synchronous (blue), and listener precedes (green) couplings (model *P* value fixed to FDR, $\gamma = 0.05$).

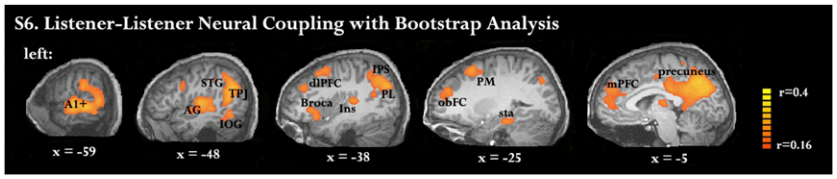
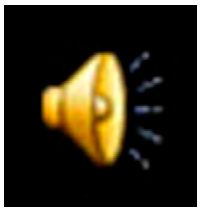


Fig. S6. Significant listener–listener couplings identified using shuffled bootstrap estimates of the null distribution, resulting in more extensive listener–listener alignment. Figure layout is identical to that in Fig. 2*A*.

Table S1. Talairach coordinates of ROIs

Anatomical region	Hemisphere	Talairach coordinates X	Y	Z
	Right	57	−9	6
A1+	Left	−57	−20	5
Parietal lobule	Right	42	−46	31
	Right	56	−49	17
TPJ	Left	−47	−55	21
	Right	42	11	7
Broca	Left	−34	19	−10
	Right	35	8	3
Insula	Left	−32	13	−4
Anterior STG	Right	44	15	−15
Orbitofrontal cortex	Right	25	52	22
	Right	24	8	52
Premotor cortex	Left	−13	−4	69
	Right	6	−59	29
Precuneus	Left	−6	−65	26
	Right	6	42	27
mPFC	Left	−8	42	13
Insula	Left	−39	−77	25
mPFC	Left	−40	29	21
Inferior occipital gyrus	Left	−50	−57	−5



Movie S1. A segment of the audio recording of the story told by the speaker in the fMRI scanner and then played for 11 listeners.

Movie S1