How a speaker herds the audience: Multi-brain neural convergence over time during naturalistic storytelling

Claire H. C. Chang[1,2*], Samuel A. Nastase[1], and Uri Hasson[1]

1. Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey, 08540, USA.

2. The Graduate Institute of Mind, Brain and Consciousness, Taipei Medical University, New Taipei City, 235, Taiwan

*Correspondence: claire.hc.chang@gmail.com (C.C.)

Keywords
Brain-to-brain coupling, fMRI, multi-brain neural dynamics, narratives, naturalistic stimuli

Summary
Sharing narratives is an ancient and efficient way for humans to transmit experiences to each other. The efficacy of a given narrative has been shown to be associated with the similarity between the brain activation patterns of the speaker and the listeners (SL), as well as the neural similarity among the listeners (LL). Operationalizing the pattern (dis)similarity as the distances between participants, this study proposes a “herding hypothesis”. That is, like a group of sheep guided by a shepherd, the more closely the listeners follow the speaker, i.e. higher SL similarity, the more tightly the listeners will tend to cluster together, i.e. higher LL similarity. Using fMRI data collected during the verbal production of two spoken narratives as well as in an audience of listeners, we found that SL and LL similarities are correlated across time, as predicted by the herding hypothesis. In addition, the more “herded” brain regions also show a stronger LL similarity at the more engaging moments of the narrative, supporting an interpretation that the herding effect reflects effective storytelling. By taking both LL and SL neural coupling into consideration in a moment-by-moment manner, this study demonstrates that examining the dynamic multi-brain functional network can potentially reveal when and how the speaker loses the audience; for example, whether they go astray in all directions or they share the same misunderstanding.

Introduction

Humans use narratives to convey complex, temporally structured sequences of thoughts to one another (Bruner et al., 1986; Willems et al., 2020). This kind of communication is thought to rely on a process of neural "alignment" or "coupling" [1,2], whereby the speaker guides the listener(s) through a sequence of brain states to arrive at an understanding of the ideas or events the speaker intends to convey. Spoken stories have been found to drive synchronized neural alignment among listeners (LL coupling) throughout the cortical language network and extending into higher-level areas of the default-mode network thought to support event representation and narrative comprehension [3]. On the other hand, asymmetric, time-lagged coupling has been observed between the speaker and listener(s) (SL coupling) in an overlapping set of high-level cortical areas [4–9].

The efficacy of a given narrative has been shown to vary across individuals. Both higher LL neural coupling and higher SL neural coupling have been separately associated with better behavioral estimates of speech comprehension across individuals [4,5,7,9–12,12–16,16–18]. Individuals performing better in the post-test often showed higher neural coupling with the speaker or other listeners. The efficacy of a narrative may also vary across time: a storyteller may meander or lose focus, and the content of the narrative may fluctuate in terms of how engaging it is or how much it resonates with listeners.

This study proposes a "herding hypothesis" incorporating both LL and SL neural coupling into a unified framework for the multi-brain neural dynamics between speaker and audience. Like a shepherd, a successful speaker guides the listeners toward the same brain states. We operationalize the "distance" between speaker and listeners as the intersubject (dis)similarity of brain activity patterns within different cortex regions: SL dissimilarity reflects the distance between the speaker and the listeners; LL dissimilarity reflects the distance between listeners (Fig. 1). The herding hypothesis proposes that when the listeners follow the speaker closely, they also tend to cluster more closely to each other. On the other hand, when the audience is lost, most of the time, they simply go astray in all directions [19], resulting in low SL coupling, as well as low LL coupling. In other words, we expect SL and LL pattern (dis)similarities to correlate over the course of a narrative.

The shepherd must be a few steps ahead of the sheep. With this in mind, we expect the speaker's brain activity patterns to precede the audience's brain activity within a range of a few seconds (1 TR = 1.5 seconds) based on prior work reporting that the listener's brain activity echoes that of the speaker with a lag of up to 10 seconds [4,5,7,9,12,20,21]. In this case, a significant herding effect would mean that the listeners cluster together in trailing the speaker.

Aiming for a better understanding of the multi-brain neural dynamics underlying storytelling, this study first verifies the herding hypothesis: the listeners cluster together when they follow the speaker and disperse in different directions when they deviate from the speaker, and then illustrates the potential of this approach in identifying not only successful listeners but also moments in a speech where the audience resonates with the speaker.

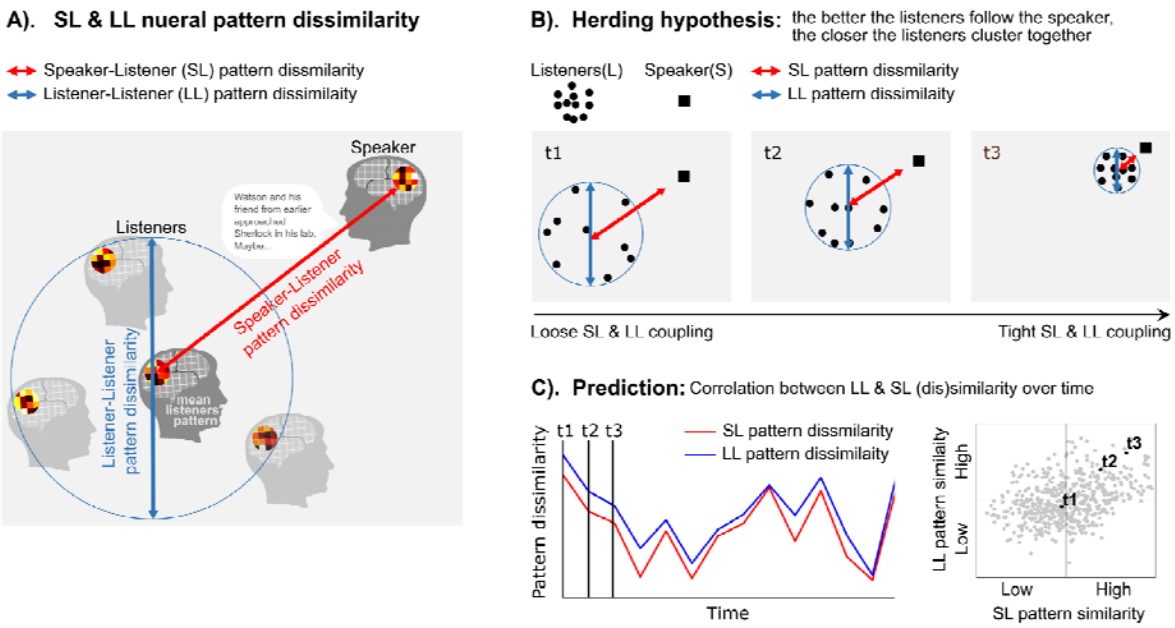

Figure 1. The herding hypothesis. A). We use neural pattern dissimilarities to quantify the distances between the speaker and listeners (SL) and the distances between listeners (LL). B). The herding hypothesis proposes that listeners cluster together when they follow the speaker closely, like a group of sheep guided by a shepherd. C). In other words, the herding hypothesis proposes that SL pattern dis(similarity) correlates with LL (dis)similarity across time. Note that SL coupling was computed using a lag of -10 to -1 TRs (speaker activity precedes for 15 to 1.5 seconds).

Results

The herding hypothesis predicts that listeners will more closely cluster together at moments of the story where they more closely follow the speaker (Fig. 1). We quantify the distance between speaker and listeners by computing the moment-by-moment intersubject (dis)similarity of neural activity patterns. The resulting dynamic LL and SL couplings indicate how tightly the listeners are clustered together and aligned to the speaker, respectively. We calculate the SL dissimilarity at different lags from -15 to -1.5 seconds. We first verify that listener activity patterns echo those of the speaker with a lag of several seconds; that is, SL similarities peak at negative lags (Fig. 2). We then reveal a significant herding effect in which the LL coupling is correlated with the strength of SL coupling in the DMN and language network (Fig. 3). Finally, we show higher LL coupling at moments of the story behaviorally reported as more engaging (Fig. 4A). This effect is stronger in brain regions showing higher herding effect, such as DMN (Fig. 4B), providing behavioral evidence that the herding effect reflects how well the audience follows the speaker.

SL and LL neural similarities: Listeners follow the speaker's brain activity patterns

To visualize the relationship between SL and LL neural similarities, we first plot the total ROI × lag intersubject similarity matrices (Fig. 2). In agreement with previous studies [4,7,9,20], SL dynamics are markedly different from LL dynamics. LL similarities peak at lag 0 in most regions:

listener brain activity patterns are synchronized in processing the story's content. In contrast, significant SL similarities mainly occur at negative lags: listener activity patterns echo the speaker activity patterns with seconds-long lags.

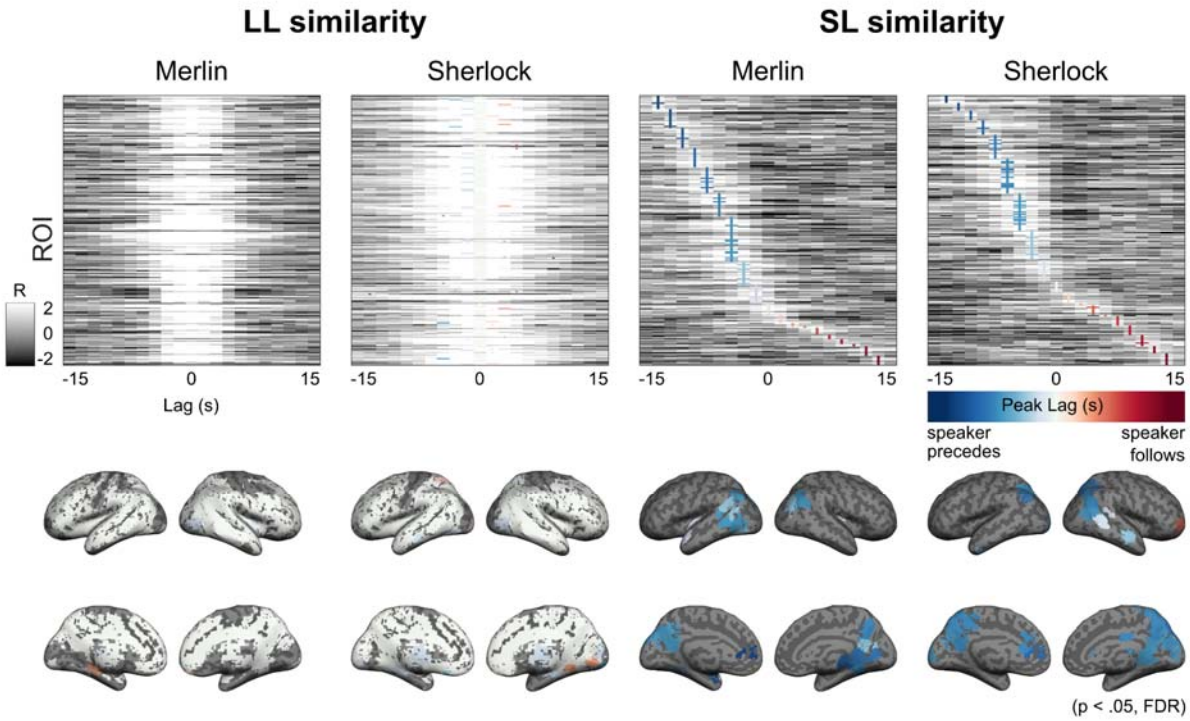

Figure 2. LL and SL neural pattern similarities. Full intersubject similarity matrices (upper) where each row shows the neural pattern similarities in each brain region at varying lags across columns. Brain regions are ordered by their peak lags. Intersubject pattern similarities for each TR were averaged across all TRs in the story. Lags with the peak correlation values are color-coded. Significant peak lags are marked with wide (horizontal) colored bars ($p < .05$, FDR correction). Nonsignificant peak lags are marked with narrow colored bars. The correlation values are normalized with Fisher's transformation and then z-scored across lags. ROIs with significant peak lags are plotted on the brain (lower). LL similarities peak at lag 0 in most brain regions, reflecting that the listeners are synchronized. On the other hand, SL similarities often peak at negative lags, indicating that the speaker precedes the listeners.

Herding effect: The more closely the listeners follow the speaker, the more tightly the listeners cluster together

The herding hypothesis predicts that the more closely the listeners follow the speaker, the more closely the listeners cluster together. We quantified the herding effect by computing the correlation between moment-by-moment LL coupling and lagged SL coupling throughout the narrative. Statistical significance for the herding effect was assessed using permutation procedures based on two surrogate datasets, one generated by replacing the speaker with a "pseudo-speaker" sampled from the listeners (Fig. S1), the other by applying unreasonable SL

lags (i.e., the speaker precedes the listeners for more than 15 seconds or the speaker does not precede the listeners). Only ROIs that passed both statistical tests are considered to show a significant herding effect. Namely, a stronger herding effect was found with the real speaker rather than pseudo-speakers and with reasonable rather than unreasonable SL lags. For comparison, we also computed the herding metrics based on SL coupling at 0 lag and found no significant effect (Fig. S2). We test the herding hypothesis with one-tailed tests. For the results of ad hoc two-tailed tests, see Fig. S3.

Our results reveal a significant herding effect for both stories in the precuneus, posterior cingulate cortex, cuneus, superior and middle temporal gyrus, and superior/middle occipital gyrus (Fig. 3); many of these regions have been implicated in representing high-level events and narrative features [3,22–24]. See Fig. S4 for an exemplar ROI showing significant LL coupling but a nonsignificant herding effect. In addition, a similar herding effect is revealed with alternative SL coupling measurement, namely, averaged pattern similarity between the speaker and each listener (Fig. S5) instead of pattern similarly between the speaker and the averaged listener pattern (Fig. 3).

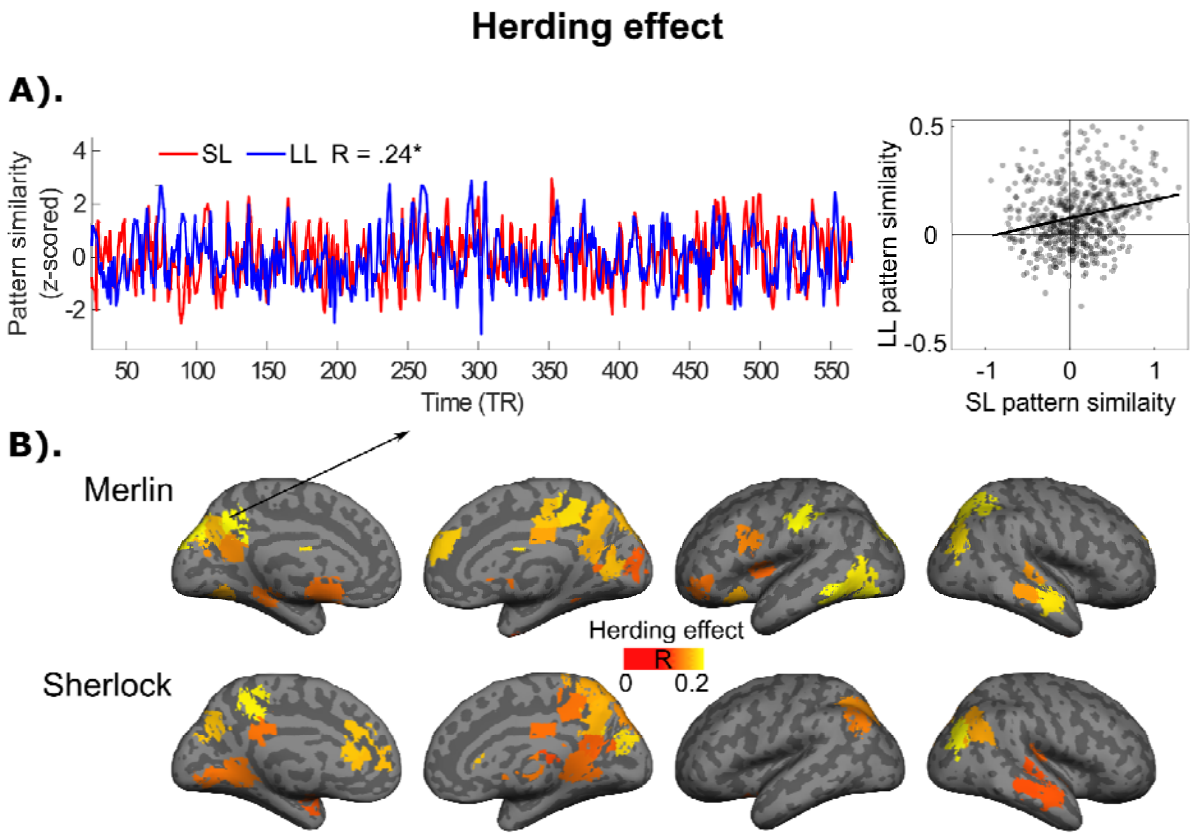

Figure 3. Cortical areas with a significant herding effect. A). A left precuneus ROI showing a significant correlation between SL and LL coupling over the course of "Merlin," namely, a significant herding effect. B). All ROIs with a significant herding effect. They are colored according to the amplitude of the correlation between SL and LL similarity ($p < .05$, FDR correction). SL and LL similarities were normalized with Fisher's transformation and z-scored across time before computing the herding effect.

More "herded" brain regions show stronger LL similarity at engaging moments of the story.

What drives the herding effect? We hypothesized that fluctuations in how engaging listeners find certain parts of the story may relate to how effectively the speaker "herds" the listeners. To behaviorally assess how engaging the spoken narrative was moment by moment, we collected continuous engagement scores from a separate group of participants. In agreement with a previous study [25], we found that engagement scores correlate with LL neural similarity; that is, higher LL similarity occurs at more engaging moments of a story. In a similar vein, higher neural synchronization has been reported for more memorable [26], surprising [27], and emotional moments during stories [28,29].

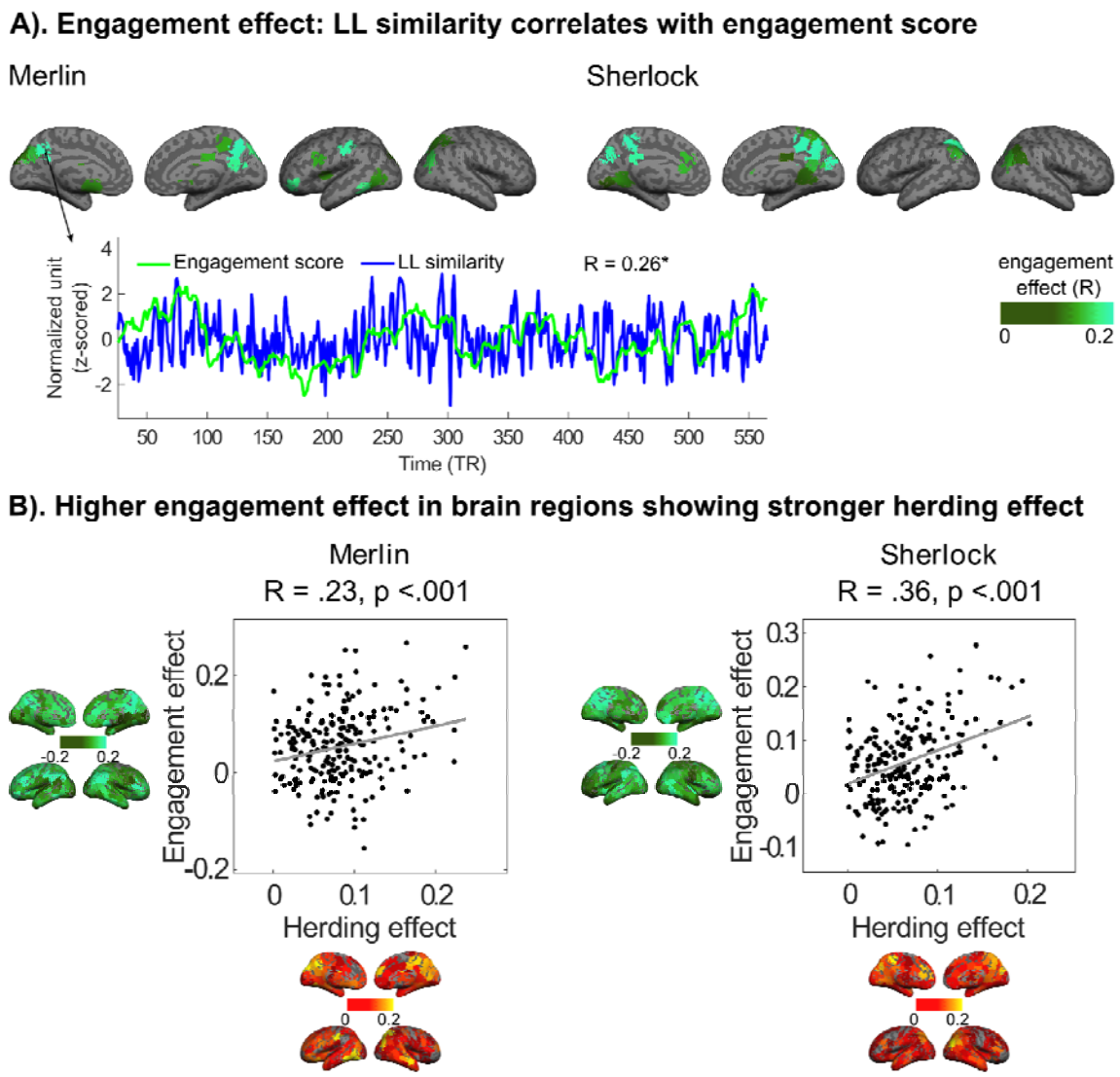

Figure 4. The engagement effect. A). Brain regions where the behavioral engagement score significantly correlated with LL neural similarity (p < .05, FDR correction). In these regions,

moments in the story with higher engagement ratings elicit higher LL neural similarity. The color bar indicates the magnitude of the correlation between the engagement score and LL similarity. B). Brain regions with a stronger herding effect tend to show a stronger engagement effect. This finding provides behavioral evidence that the herding effect reflects how engaging listeners find the narrative.

Among regions showing a significant herding effect, a significant engagement effect was found for both stories in the precuneus, posterior cingulate cortex, cuneus, and superior/middle occipital gyrus (Fig. 4A). More importantly, the engagement effect is larger in areas showing a stronger herding effect (Fig. 4B). This finding provides behavioral evidence that the herding effect reflects how engaged the listeners are with the content of the story.

Discussion

This study examined the multi-brain neural dynamics underlying storytelling. We first verified that the audience echoed the speaker's neural activation patterns with a temporal lag (Fig. 2) [4,7,9,12,17,20]. As predicted by the herding hypothesis (Fig. 1), the more closely the listeners' brain activity matched that of the speaker, the more closely the listeners clustered together (Fig. 3). We argue that this herding effect is an index of effective communication during storytelling, indicating that, metaphorically, the speaker guides the listeners' neural activity, especially in higher-order brain areas. We also show that LL neural similarity increases at more engaging moments of the story (Fig. 4A). This engagement effect is stronger in the more "herded" brain regions (Fig. 4B), supporting the hypothesis that the herding effect reflects effective storytelling—that is, when the storyteller most successfully conveys their thoughts to the listeners.

A significant herding effect was found in several high-order brain areas in the DMN, including the precuneus, middle/posterior cingulate cortex, lateral parietal cortex, and right anterolateral temporal cortex (Fig. 3). The posterior medial regions in particular have been shown to encode paragraph-level narrative structure [30]. These regions are thought to host content-specific, supramodal event representations [3,22,24,31–33], linking the production and comprehension of spoken narratives [4,9,34]. The current results add a dynamical perspective to this body of work, suggesting that the speaker's own neural trajectory through high-level event features may guide the listeners' upcoming event representations with varying effectiveness over the course of a narrative. This dynamic convergence and divergence of idiosyncratic internal representations with the unfolding narrative would be particularly interesting for further investigation.

If the herding effect can be interpreted as an index of effective speech, what factors may impact how closely the speaker resonates with the audience? In keeping with recent theoretical work positioning the DMN as a high-level interface between external events with prior knowledge (Yeshurun et al., 2021), we speculate that individual differences in how closely listeners follow the speaker may in part reflect differences in the way the speaker's narrative aligns with each listener's internal state and idiosyncratic memories. Prior work, for example, has suggested that brain-to-brain coupling may vary as a function of social closeness [35,36] or whether the speaker and listener share similar beliefs (e.g. similar political orientation; [37]. It is worth noting, however, that there may be different kinds of effective speech. For example, a speaker may seek to (mis)direct listeners toward a different understanding than their own; listeners that are unfamiliar

versus experts with a particular topic may experience the same speech very differently [10,14]; and a speaker may attempt to "meet certain listeners where they are" rather than wrangling all listeners similarly.

The methodology we introduce takes into account both LL and SL couplings in a moment-by-moment manner. In our stories, most of the time, the listeners converge to trail the speaker and when they lose track, they disperse in different directions (low LL and low SL; time points in the lower-left quadrant of the scatterplot in Fig. 3A). However, there are moments where LL is high despite low SL (time points in the upper-left quadrant of the scatterplot in Fig. 3A), or LL is low despite relatively high SL (e.g. time points in the lower-right quadrant of the scatterplot in Fig. 3A). We speculate that in the former case, the audience might share the same misunderstanding, or the speaker might not undergo the experience from the same perspective as the listeners [38] while in the latter case, the listeners might only form a loose group around the speaker due to ambiguous speech or heterogeneous apprehension [32].

We hope that, with more diverse speakers and larger audiences, future work will be able to more extensively sample the less successful moments of communication across narratives. Our framework for measuring dynamic, multi-brain coupling can highlight when and how a speaker and the audience become misaligned. By relating these moments back to the speaker's delivery, narrative content, and the personal backgrounds of both the speaker and individual listeners, we will be better positioned to identify why communication breaks down and develop solutions for accommodating different learning styles.

Methods

fMRI datasets

This study relied on two openly available auditory story-listening datasets from the "Narratives" collection (OpenNeuro: https://openneuro.org/datasets/ds002245; [39]), including "Sherlock" and "Merlin" (18 participants, 11 females). The speaker data reported in the original study [9] was also included. All participants reported fluency in English and were 18–40 years of age. The criteria for participant exclusion have been described in Zadbood et al. [9]. All participants provided informed, written consent, and the experimental protocol was approved by the institutional review board of Princeton University.

fMRI preprocessing

fMRI data were preprocessed using FSL (https://fsl.fmrib.ox.ac.uk/), including slice time correction, volume registration, and high-pass filtering (140 s cutoff). All data were aligned to standard 3 × 3 × 4 mm Montreal Neurological Institute space (MNI152). A gray matter mask was applied. The first 25 and last 20 volumes of fMRI data were discarded to remove large signal fluctuations at the beginning and end of the time course to account for signal stabilization and stimulus onset/offset prior to computing intersubject dissimilarities [40]. The global mean responses were subtracted before pattern similarity analyses [41,42].

ROI masks

We used 238 functional ROIs defined independently by Shen and colleagues [43] based on whole-brain parcellation of resting-state fMRI data. ROIs with less than 10 voxels based on the coverage of our BOLD acquisition were excluded from further analyses.

SL and LL neural similarities

We computed intersubject pattern correlations in each ROI at each time point of the story, i.e. TR by TR spatial pattern similarities (Fig. 1), using the leave-one-participant-out method [40]. For LL similarity, we computed the correlation between the activation pattern from one listener and the averaged pattern of the other 17 listeners. Similarly, SL similarity was computed between the speaker and the average pattern of 17 listeners, excluding each listener in turn. Note that quantifying SL coupling in this way entails that SL coupling can be high while LL coupling is low; i.e. listeners may be widely dispersed but roughly centered on the speaker. We also recomputed SL coupling by first computing the similarities between the speaker and each individual listener and then averaging these similarities. This analysis yielded qualitatively similar results (Fig. S5). According to the literature, the speaker and listener activation patterns are not necessarily temporally synchronized [4,6,7,9,20]. Therefore, we also computed the neural similarities at varying lags.

Pearson correlation was used to estimate pattern similarity. Time-lagged neural similarities were computed by circularly shifting the time series such that the non-overlapping edge of the shifted time series was concatenated to the beginning or end. The resulting correlation values were normalized with Fisher's z transformation before further statistical analyses.

We statistically evaluated the SL and LL neural similarities separately before examining the herding effect (Fig. 2). We generated surrogates with the same mean and autocorrelation as the original time series by time-shifting and time-reversing the functional data prior to computing the intersubject similarities. We computed the correlation between the original seed and time-shifted/-reversed target time series. All possible time shifts were used to generate the null distribution. The resulting correlation values were compiled into null distributions after averaging across time points and participants. One-tailed z-tests were applied to compare neural similarities within the window of lag -10 to +10 TRs against this null distribution. We corrected for multiple comparisons across lags and ROIs by controlling the false discovery rate (FDR) at $q < .05$ [44].

Computing the herding metric

We defined the herding metric as the correlation between LL neural similarity at lag 0 and SL neural similarity at lags within the window of -10 to -1 TRs (i.e. speaker precedes the listeners for 1.5 to 15 seconds). A significant herding effect indicates that the listeners are more synchronized when they echo the speaker's activation pattern. Two statistical tests were applied.

First, to verify that only the speaker showed the herding effect, we replaced the actual speaker with each of the 18 listeners to serve as the pseudo-speaker (Fig. S1). LL similarity was computed among the remaining 17 listeners, excluding the pseudo-speaker, using the leave-one-out method, i.e. correlation between the activation pattern from one listener and the averaged pattern of all the other 16 listeners. SL similarity was computed between the real speaker and the average pattern of the 17 listeners, Pseudo-SL similarity was computed using the same method as the SL similarity, except that the real speaker was replaced by the pseudo-speaker. We computed the herding metric with the real and pseudo-SL similarity and compared the real and pseudo-herding effects using a two-sample one-tailed t-test ($N = 18$). We corrected for multiple comparisons across lags and ROIs by controlling the FDR at $q < .05$. Only the ROI x SL lag combinations that passed this test were included for the second statistical test.

Second, the speaker must precede the listeners to "herd" them. Therefore, we tested the real herding effect against correlation values between LL at lag 0 and SL at all the possible lags outside of the chosen lag window (-10~-1 TR) using one-tailed z-tests. We circularly shifted the original time series to obtain a time-lagged time series. The number of possible lags equals the number of time points. The FDR method was used to control for multiple comparisons (ROI x SL lag; $q < .05$). Only ROIs that passed both statistical tests are considered to show a significant herding effect.

To quantify the amplitude of the herding effect, we extracted the peak LL-SL correlation value within the -10 to -1 TR SL lag window. We required that the peak value be larger than the absolute value of any negative peak and excluded any peaks occurring at the edges of the window.

Behavioral engagement

Engagement ratings

Behavioral assessments of dynamic engagement were acquired in another group of participants recruited via Amazon Mechanical Turk. Participants with less than 20 unique rating scores (i.e. effectively flat ratings across the story) were excluded. 33 raters were included for "Merlin" (15 females). A separate sample of 34 raters was included for "Sherlock" (14 females). All participants reported fluency in English and were 25–71 years of age. All participants provided informed, written consent, and the experimental protocol was approved by the institutional review board of Princeton University.

The participants were instructed to indicate "how engaging the current event is" while listening to the stories by moving a slider continuously. We presented the stories and collected the data using the web-based tool DANTE (Dimensional Annotation Tool for Emotions) (https://github.com/phuselab/DANTE) [45]. The rating scores were acquired with a resolution of 0.04 seconds and then downsampled to 1.5 seconds (= 1 TR).

The engagement scores were z-scored across time, detrended, and averaged across raters.

Correlation between engagement and LL neural similarity

To quantify the relationship between time-point-by-time-point engagement ratings and LL neural similarity, we computed the Pearson correlation between the engagement scores and the LL similarity over time within ROIs showing a significant herding effect (one-tailed). We corrected for multiple comparisons across ROIs by controlling the FDR at $q < .05$.

Correlation between engagement effect and herding effect across ROIs

To quantify the relationship between engagement ratings and group-level herding, we computed the Pearson correlation between the herding effect and the engagement effect across all ROIs ($p < .05$). Note that since the number of ROIs is fixed, a significant p-value associated with this correlation does not indicate generalization to other regions (and does not support population-level inference).

Data and code availability.

This study relied on openly available spoken story datasets from the "Narratives" collection (OpenNeuro: https://openneuro.org/datasets/ds002245) [39].

Acknowledgments

This study is supported by the National Institute of Mental Health (R01-MH112357 and DP1-HD091948).

Author contributions

C.H.C.C. designed research; C.H.C.C. performed research; C.H.C.C. contributed new reagents/analytic tools; C.H.C.C. analyzed data; and C.H.C.C., S.A.N., and U.H. wrote the paper.

Declaration of interests

The authors declare no competing interests.

References

1. Hasson, U., Ghazanfar, A.A., Galantucci, B., Garrod, S., and Keysers, C. (2012). Brain-to-brain coupling: A mechanism for creating and sharing a social world. Trends Cogn. Sci. *16*, 114–121. 10.1016/j.tics.2011.12.007.

2. Pickering, M.J., and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. Behav. Brain Sci. *27*, 169–190. 10.1017/s0140525x04000056.

3. Chen, J., Leong, Y.C., Norman, K., and Hasson, U. (2016). Shared experience, shared memory: a common structure for brain activity during naturalistic recall 10.1101/035931.

4. Liu, L., Li, H., Ren, Z., Zhou, Q., Zhang, Y., Lu, C., Qiu, J., Chen, H., and Ding, G. (2022). The “two-brain” approach reveals the active role of task-deactivated default mode network in speech comprehension. Cereb. Cortex. 10.1093/cercor/bhab521.

5. Nguyen, M., Chang, A., Micciche, E., Meshulam, M., Nastase, S.A., and Hasson, U. (2022). Teacher–student neural coupling during teaching and learning. Soc. Cogn. Affect. Neurosci. *17*, 367–376. 10.1093/scan/nsab103.

6. Silbert, L.J., Honey, C.J., Simony, E., Poeppel, D., and Hasson, U. (2014). Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. Proc. Natl. Acad. Sci. U. S. A. *111*, E4687–E4696. 10.1073/pnas.1323812111.

7. Stephens, G.J., Silbert, L.J., and Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. Proc. Natl. Acad. Sci. U. S. A. *107*, 14425–14430. 10.1073/pnas.1008662107.

8. Zada, Z., Goldstein, A., Michelmann, S., Simony, E., Price, A., Hasenfratz, L., Barham, E., Zadbood, A., Doyle, W., Friedman, D., et al. A shared linguistic space for transmitting our thoughts from brain to brain in natural conversations.

9. Zadbood, A., Chen, J., Leong, Y.C., Norman, K.A., and Hasson, U. (2017). How we transmit memories to other brains: Constructing shared neural representations via communication. Cereb. Cortex *27*, 4988–5000. 10.1093/cercor/bhx202.

10. Chen, J., Qian, P., Gao, X., Li, B., Zhang, Y., and Zhang, D. (2023). Inter-brain coupling reflects disciplinary differences in real-world classroom learning. Npj Sci. Learn. *8*, 11. 10.1038/s41539-023-00162-1.

11. Cohen, S.S., Madsen, J., Touchan, G., Robles, D., Lima, S.F.A., Henin, S., and Parra, L.C. (2018). Neural engagement with online educational videos predicts learning performance for individual students. Neurobiol. Learn. Mem. *155*, 60–64. 10.1016/j.nlm.2018.06.011.

12. Davidesco, I., Laurent, E., Valk, H., West, T., Dikker, S., Milne, C., and Poeppel, D. (2019). Brain-to-brain synchrony between students and teachers predicts learning outcomes. bioRxiv, 644047.

13. Liu, J., Zhang, R., Geng, B., Zhang, T., Yuan, D., Otani, S., and Li, X. (2019). Interplay between prior knowledge and communication mode on teaching effectiveness: Interpersonal neural synchronization as a neural marker. NeuroImage *193*, 93–102. 10.1016/j.neuroimage.2019.03.004.

14. Meshulam, M., Hasenfratz, L., Hillman, H., Liu, Y.-F.F., Nguyen, M., Norman, K.A., and Hasson, U. (2021). Neural alignment predicts learning outcomes in students taking an introduction to computer science course. Nat. Commun. *12*, 1922. 10.1038/s41467-021-22202-3.

15. Pan, Y., Dikker, S., Goldstein, P., Zhu, Y., Yang, C., and Hu, Y. (2020). Instructor-learner brain coupling discriminates between instructional approaches and predicts learning. NeuroImage *211*. 10.1016/j.neuroimage.2020.116657.

16. Zhang, L., Xu, X., Li, Z., Chen, L., and Feng, L. (2022). Interpersonal neural synchronization predicting learning outcomes from teaching-learning interaction: A meta-analysis. Front. Psychol. *13*, 1–11. 10.3389/fpsyg.2022.835147.

17. Zheng, L., Chen, C., Liu, W., Long, Y., Zhao, H., Bai, X., Zhang, Z., Han, Z., Liu, L., Guo, T., et al. (2018). Enhancement of teaching outcome through neural prediction of the students' knowledge state. Hum. Brain Mapp. *39*, 3046–3057. 10.1002/hbm.24059.

18. Zhu, Y., Leong, V., Hou, Y., Zhang, D., Pan, Y., and Hu, Y. (2022). Instructor–learner neural synchronization during elaborated feedback predicts learning transfer. J. Educ. Psychol. *114*, 1427–1441. 10.1037/edu0000707.

19. Finn, E.S., Glerean, E., Khojandi, A.Y., Nielson, D., Molfese, P.J., Handwerker, D.A., and Bandettini, P.A. (2020). Idiosynchrony: From shared responses to individual differences during naturalistic neuroimaging. NeuroImage *215*. 10.1016/j.neuroimage.2020.116828.

20. Dikker, S., Silbert, L.J., Hasson, U., and Zevin, J.D. (2014). On the same wavelength: Predictable language enhances speaker-listener brain-to-brain synchrony in posterior superior temporal gyrus. J. Neurosci. *34*, 6267–6272. 10.1523/JNEUROSCI.3796-13.2014.

21. Zada, Z., Goldstein, A., Michelmann, S., Simony, E., Price, A., Hasenfratz, L., Barham, E., Zadbood, A., Doyle, W., Friedman, D., et al. (2023). A shared linguistic space for transmitting our thoughts from brain to brain in natural conversations. bioRxiv. https://doi.org/10.1101/2023.06.27.546708.

22. Baldassano, C., Chen, J., Zadbood, A., Pillow, J.W., Hasson, U., Norman, K.A., Hasson, U., and Norman, K.A. (2017). Discovering event structure in continuous narrative perception and memory. Neuron *95*, 709-721.e5. 10.1016/j.neuron.2017.06.041.

23. Baldassano, C., Hasson, U., and Norman, K.A. (2018). Representation of real-world event schemas during narrative perception. J. Neurosci. *38*, 9689–9699. 10.1523/JNEUROSCI.0251-18.2018.

24. Chang, C.H.C., Lazaridi, C., Yeshurun, Y., Norman, K.A., and Hasson, U. (2021). Relating the past with the present: Information integration and segregation during ongoing narrative processing. J. Cogn. Neurosci. *33*, 1106–1128. 10.1162/jocn_a_01707.

25. Song, H., Finn, E.S., and Rosenberg, M.D. (2021). Neural signatures of attentional engagement during narratives and its consequences for event memory. Proc. Natl. Acad. Sci. U. S. A. *118*. 10.1073/pnas.2021905118.

26. Simony, E., Honey, C.J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., and Hasson, U. (2016). Dynamic reconfiguration of the default mode network during narrative comprehension. Nat. Commun. *7*, 12141. 10.1038/ncomms12141.

27. Brandman, T., Malach, R., and Simony, E. (2021). The surprising role of the default mode network in naturalistic perception. Commun. Biol. *4*, 1–18. 10.1038/s42003-020-01602-z.

28. Nummenmaa, L., Saarim??ki, H., Glerean, E., Gotsopoulos, A., Jääskeläinen, I.P., Hari, R., and Sams, M. (2014). Emotional speech synchronizes brains across listeners and engages large-scale dynamic brain networks. NeuroImage *102*, 498–509. 10.1016/j.neuroimage.2014.07.063.

29. Smirnov, D., Saarimäki, H., Glerean, E., Hari, R., Sams, M., and Nummenmaa, L. (2019). Emotions amplify speaker–listener neural alignment. Hum. Brain Mapp. *40*, 4777–4788. 10.1002/hbm.24736.

30. Lerner, Y., Honey, C.J., Silbert, L.J., and Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. J. Neurosci. *31*, 2906–2915. 10.1523/JNEUROSCI.3684-10.2011.

31. Honey, C.J., Thesen, T., Donner, T.H., Silbert, L.J., Carlson, C.E., Devinsky, O., Doyle, W.K., Rubin, N., Heeger, D.J., and Hasson, U. (2012). Slow cortical dynamics and the accumulation of information over long timescales. Neuron *76*, 423–434. 10.1016/j.neuron.2012.08.011.

32. Nguyen, M., Vanderwal, T., and Hasson, U. (2019). Shared understanding of narratives is correlated with shared neural responses. NeuroImage *184*, 161–170. 10.1016/J.NEUROIMAGE.2018.09.010.

33. Yeshurun, Y., Nguyen, M., and Hasson, U. (2017). Amplification of local changes along the timescale processing hierarchy. Proc. Natl. Acad. Sci., 201701652. 10.1073/pnas.1701652114.

34. Chen, J., Leong, Y.C., Honey, C.J., Yong, C.H., Norman, K.A., and Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. Nat. Neurosci. *20*, 115–125. 10.1038/nn.4450.

35. Bevilacqua, D., Davidesco, I., Wan, L., Chaloner, K., Rowland, J., Ding, M., Poeppel, D., and Dikker, S. (2019). Brain-to-brain synchrony and learning outcomes vary by student–teacher dynamics: Evidence from a real-world classroom electroencephalography study. J. Cogn. Neurosci. *31*, 401–411. 10.1162/jocn_a_01274.

36. Dikker, S., Wan, L., Davidesco, I., Kaggen, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., Van Bavel, J.J., Ding, M., et al. (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. Curr. Biol. *27*, 1375–1380. 10.1016/j.cub.2017.04.002.

37. Katabi, N., Simon, H., Yakim, S., Ravreby, I., Ohad, T., and Yeshurun, Y. (2023). Deeper than you think: Partisanship-dependent brain responses in early sensory and motor brain regions. J. Neurosci. *43*, 1027–1037. 10.1523/JNEUROSCI.0895-22.2022.

38. Sun, B., Xiao, W., Feng, X., Shao, Y., Zhang, W., and Li, W. (2020). Behavioral and brain synchronization differences between expert and novice teachers when collaborating with students. Brain Cogn. *139*, 105513. 10.1016/j.bandc.2019.105513.

39. Nastase, S.A., Liu, Y.-F., Hillman, H., Zadbood, A., Hasenfratz, L., Keshavarzian, N., Chen, J., Honey, C.J., Yeshurun, Y., Regev, M., et al. (2021). The "Narratives" fMRI dataset for evaluating models of naturalistic language comprehension. Sci. Data *8*, 250. 10.1038/s41597-021-01033-3.

40. Nastase, S.A., Gazzola, V., Hasson, U., and Keysers, C. (2019). Measuring shared responses across subjects using intersubject correlation. Soc. Cogn. Affect. Neurosci. *14*, 667–685. 10.1093/scan/nsz037.

41. Garrido, L., Vaziri-Pashkam, M., Nakayama, K., and Wilmer, J. (2013). The consequences of subtracting the mean pattern in fMRI multivariate correlation analyses. Front. Neurosci. *7*, 174. 10.3389/fnins.2013.00174.

42. Murphy, K., Birn, R.M., Handwerker, D.A., Jones, T.B., and Bandettini, P.A. (2008). The impact of global signal regression on resting state correlations: Are anti-correlated networks introduced? NeuroImage *44*, 893–905. 10.1016/j.neuroimage.2008.09.036.

43. Shen, X., Tokoglu, F., Papademetris, X., and Constable, R.T.T. (2013). Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. NeuroImage *82*, 403–415. 10.1016/j.neuroimage.2013.05.081.

44. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Methodol. *57*. 10.2307/2346101.

45. Boccignone, G., Conte, D., Cuculo, V., and Lanzarotti, R. (2017). AMHUSE: A multimodal dataset for humour sensing. In Proceedings of the 19th ACM International Conference on Multimodal Interaction, pp. 438–453. 10.1145/3136755.3136806.