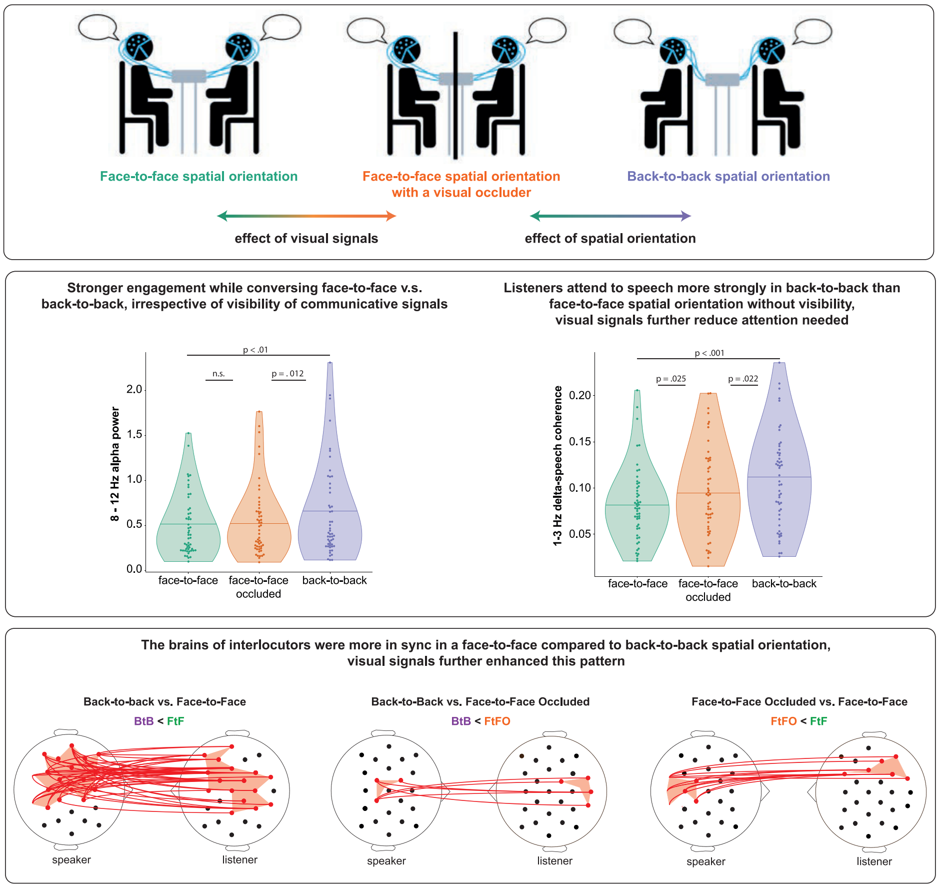
iScience

CellPress
OPEN ACCESS

Article

Face-to-face spatial orientation fine-tunes the brain for neurocognitive processing in conversation



Linda Drijvers,
Judith Holler

linda.drijvers@mpi.nl

Highlights

Listeners engage more strongly when conversing face to face than back to back

More attention to speech when conversing back to back than face to face

Inter-brain synchrony was stronger face to face than back to back

Face-to-face orientation induces a special social mode for neurocognitive processing

Drijvers & Holler, iScience *25*, 105413
November 18, 2022
https://doi.org/10.1016/j.isci.2022.105413

iScience

CellPress
OPEN ACCESS

Article

Face-to-face spatial orientation fine-tunes the brain for neurocognitive processing in conversation

Linda Drijvers[1,2,3,*] and Judith Holler[1,2]

SUMMARY

We here demonstrate that face-to-face spatial orientation induces a special 'social mode' for neurocognitive processing during conversation, even in the absence of visibility. Participants conversed face to face, face to face but visually occluded, and back to back to tease apart effects caused by seeing visual communicative signals and by spatial orientation. Using dual EEG, we found that (1) listeners' brains engaged more strongly while conversing face to face than back to back, irrespective of the visibility of communicative signals, (2) listeners attended to speech more strongly in a back-to-back compared to a face-to-face spatial orientation without visibility; visual signals further reduced the attention needed; (3) the brains of interlocutors were more in sync in a face-to-face compared to a back-to-back spatial orientation, even when they could not see each other; visual signals further enhanced this pattern. Communicating in face-to-face spatial orientation is thus sufficient to induce a special 'social mode' which fine-tunes the brain for neurocognitive processing in conversation.

INTRODUCTION

The primordial ecology of human communication involves orientating toward one another (Goffman, 1961). Such a face-to-face spatial orientation facilitates access to visual signals conveyed with the face, head, and hands, which are ubiquitous in human communication (Bavelas and Chovil, 2000; Goldin-Meadow, 2005; Holler and Levinson, 2019; Kendon, 1990; McNeill, 1992). But we also orient toward one another to initiate social interaction (Mondada, 2009), to signal willingness to respond to such initiations (Schegloff, 1998), to create the feeling of address (Nagels et al., 2015) and to increase proximity (Jung et al., 2016). It also allows us to hug, shake hands, attend to a shared space between us and collaboratively act on a third entity within this space. All this suggests a profound prosocial significance of orienting toward one another that goes beyond the transmission of visual signals. Indeed, this human trait may have deep evolutionary roots, as the enhanced proclivity for face-to-face orientation in bonobos compared to chimpanzees suggests (Allanic et al., 2021).

The Interaction Engine hypothesis postulates that social instincts and motivations accompanied by a suite of interactional abilities, such as turn-taking, visual signaling, and intention attribution, created the social niche that allowed human language to evolve (Levinson, 2006, 2019). We here argue that orienting toward one another for the purpose of interaction is part of this suite of predispositions. Moreover, we put forward the radical hypothesis that this social scaffolding of communication in the form of a face-to-face spatial orientation alone is sufficient for inducing a special 'social mode' of neurocognitive processing, rendering our brains particularly predisposed for the exchange of communicative information. That is, when perceiving language while immersed in a face-to-face spatial orientation compared to a back-to-back configuration, we expect the human brain to be particularly responsive to the neurocognitive processing of this linguistic information. Crucially, we expect that these effects are not dependent on the presence of visual communicative signals, but because interlocutors are seated and conversing with their faces and bodies spatially oriented toward each other (i.e., even without having visual access to the other).

We here directly test this hypothesis using recent advances in neuroscience. Based on interactive hyperscanning paradigms, studies are increasingly moving from a single to a two-person neuroscience perspective to investigate the neural correlates of human communication (Jiang et al., 2012; Nastase

[1] Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD, Nijmegen, the Netherlands

[2] Radboud University, Donders Institute for Brain, Cognition, and Behaviour, Montessorilaan 3, 6525 HR, Nijmegen, the Netherlands

[3] Lead contact

*Correspondence: linda.drijvers@mpi.nl

https://doi.org/10.1016/j.isci.2022.105413

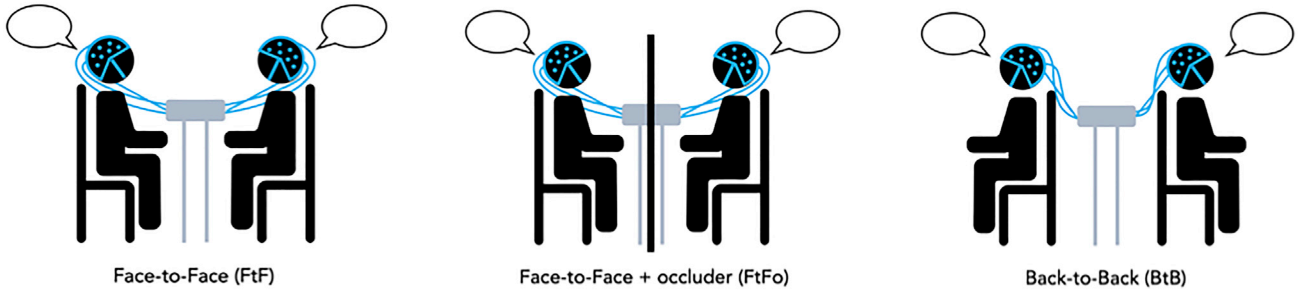


Figure 1. Experimental set-up and conditions

et al., 2020; Redcay and Schilbach, 2019; Schilbach et al., 2013; Stephens et al., 2010; Wheatley et al., 2019). Using such a dual-EEG setup, participants conversed while seated in a face-to-face spatial orientation (FtF, multimodal, i.e., with mutual visual access), a face-to-face spatial orientation with a visual occlusion between them (FtFO, vocal only) and a back-to-back spatial orientation (BtB, vocal only). We compared neurocognitive activity in these three conditions using participants' EEG for intra-brain and inter-brain analyses. We focus on the neurocognitive activity during the listening periods because our interest was in the processing of incoming speech information during conversation. Specifically, we compared the extent to which participants (1) (dis)engaged task-(ir)relevant brain regions in the task of conversing (as measured by modulations of intra-brain alpha power (Jensen and Mazaheri, 2010)), (2) tracked the speech at the phonological level (as measured by theta band tracking (Ding et al., 2016; Giraud and Poeppel, 2012; Kösem et al., 2018)), as well as at higher levels in terms of semantic processing, attention-demanding encoding, and prediction (as measured by delta band tracking (Arnal et al., 2015; Keitel et al., 2017; Kösem et al., 2018; Mai and Wang, 2019)), and (3) showed differences in alpha and beta band synchrony suggestive of differences in social coordination, shared attentional and representation states (e.g., Dikker et al., 2021; Dumas et al., 2010; Pérez et al., 2019, 2017; Tognoli et al., 2007). If we can demonstrate that intra-brain, brain-to-speech, and inter-brain activity is affected by spatial orientation irrespective of the presence of visual signals (i.e., by showing differences between the BtB and FtFO condition), then this would provide novel, first-time evidence for the hypothesis that spatial orientation in itself influences cognitive and neural processes in conversations on both the intra- and inter-person level.

Indeed, the findings showed that face-to-face spatial orientation in itself, irrespective of the presence of visual signals (i.e., being seated with their face and body spatially oriented toward each other, but without visual access), appears to induce a special 'social mode' for information processing, as evidenced by the back-to-back and face-to-face occluded conditions differing significantly from each other in cortical activity. The perception of visual signals further enhanced the effects when comparing the two face-to-face conditions in the brain-to-speech and inter-brain activity analyses, suggesting an independent effect of visual signals.

The present study breaks new ground by dissecting the effects of social interaction on neurocognitive processing into those due to dyadic visibility, as investigated by past research, and, critically, those due to face-to-face spatial orientation itself—a deeply fundamental aspect of human sociality but hitherto entirely uncharted terrain in terms of its potential effects on neurocognition.

RESULTS

Intra-brain alpha power is higher in back-to-back than in face-to-face spatial orientation

We first asked how intra-brain alpha power varied in the three different conditions (see Figure 1). Here, we focused on intra-brain alpha power in the periods where participants were listening to the speaker because our interest was in the processing of incoming speech information (an advantage of this also being that the EEG signal in these periods is less influenced by artifacts from speaking). Nonparametric cluster permutation tests revealed that alpha power was higher in the BtB condition than the FtFO condition ($p = 0.012$), higher in the BtB condition than in the FtF condition ($p < .01$), but did not differ between the FtFO and FtF condition (no clusters observed) (see Figures 2A and 2C).

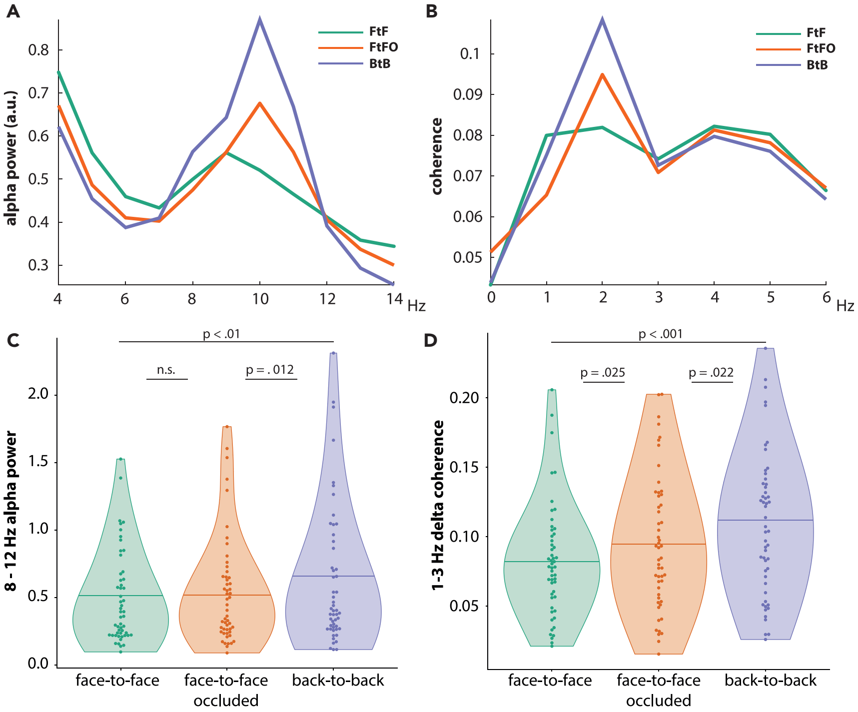


Figure 2. Intra-brain alpha power results and delta-speech tracking results

(A) Power spectrum per condition. Colored lines indicate different conditions. Alpha power was highest in the BtB condition (BTB > FtFO = FtF).

(B) Coherence between the speaker's speech and the listener's brain signal. Coherence only differed in the delta band (1–3 Hz, BTB > FtFO > FtF).

(C) Violin plots of averaged spectral alpha power per condition. Single dots represent single participant data points. Line in violin represents the median.

(D) Violin plots of coherence between the speaker's speech and the listener's brain signal in the delta band. Single dots represent single participant data points. Line in violin represents the median.

Brain-speech tracking in the delta band is stronger in back-to-back than in face-to-face spatial orientation

We then investigated whether brain-speech tracking in the delta (1–3 Hz) and theta (4–7 Hz) band differed in the three conditions. To this end, we again focused on the periods where participants were listening to the other speaker. Brain-speech tracking was stronger in the delta band in the BtB condition than in the FtFO condition ($p = 0.022$), stronger in the BtB than the FtF condition ($p < 0.001$), and stronger in the FtFO condition than in the FtF condition ($p = 0.025$) (see Figures 2B and 2D). We found no differences in the theta band (all $p > 0.05$).

Inter-brain coupling in the beta band is stronger in face-to-face than in back-to-back spatial orientation

Finally, we investigated the consistency of phase alignment of the oscillatory activity of each pair in both the alpha and beta band (using imaginary coherence, following Dikker et al. (2021). This measure minimizes the effect of instantaneous co-fluctuations of the brain signal. Cluster statistics revealed significantly stronger inter-brain coupling in the beta band in the FtF condition than in the BtB condition ($p <. 001$, Figure 3A), stronger inter-brain coupling in the FtFO condition than in the BtB condition ($p = 0.012$ Figure 3B), and stronger inter-brain coupling in the FtF than in the FtFO condition ($p = <0.01$ Figure 3C). We observed no significant differences in inter-brain coupling in the alpha band (all $p > 0.05$).

DISCUSSION

We used a dual EEG to test whether communicating in a face-to-face spatial orientation induced a special 'social mode' for neurocognitive processing during conversation. To this end, we asked participants to

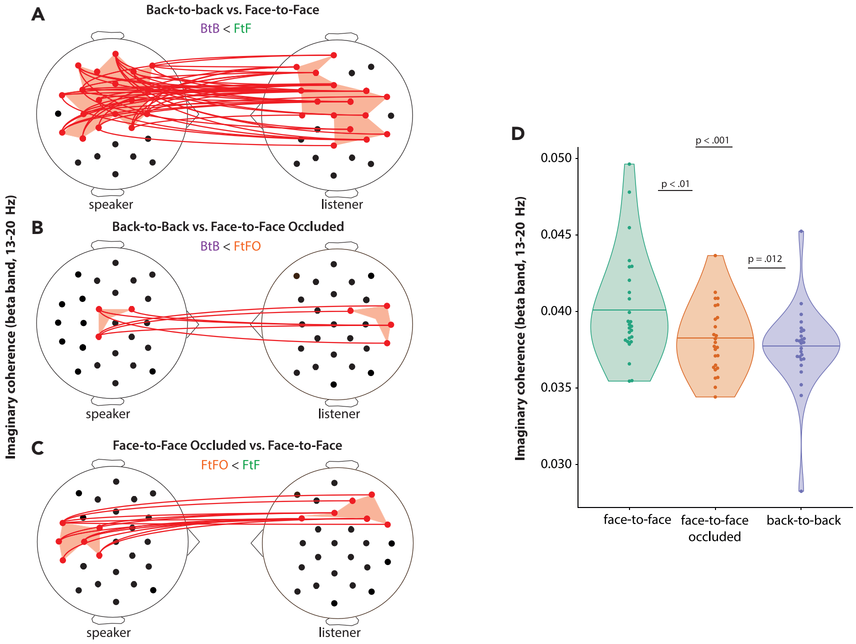


Figure 3. Inter-brain neural synchrony in the beta band between the speaker (right) and listener (left)

Red fields and lines represent statistically significant clusters and electrode pairs between the listener and speaker.

(A) Inter-brain synchrony is stronger in the FtF than the BtB condition.

(B) Inter-brain synchrony is stronger in the FtFO condition than the BtB condition.

(C) Inter-brain synchrony is stronger in the FtF condition than the FtFO condition.

(D) Violin plots of inter-brain neural synchrony per condition. Single dots represent single dyad data points. Line in violin represents the median.

converse in a face-to-face spatial orientation (multimodally, i.e., with mutual, visual access), a face-to-face spatial orientation with a visual occlusion (vocal only), and in a back-to-back spatial orientation (vocal only). This manipulation ensured that we could capture any effects on intra- and inter-brain neural activity caused by spatial orientation as well as effects caused by dyadic visibility, a fundamental distinction that has so far been missing from the literature (see e.g., Jiang et al., 2012; Rolison et al., 2020; Yun, 2013).

The results advance our understanding of human communication by demonstrating that, first and foremost, task-relevant brain regions engage more strongly when listeners converse in a face-to-face spatial orientation than in a back-to-back spatial orientation, irrespective of dyadic visibility. Second, the results show that a listener's brain tracks the speaker's speech more strongly in a back-to-back spatial orientation than in a face-to-face spatial orientation when visually occluded; the presence of visual bodily signals further enhanced this difference. Third, the data show that inter-brain synchronization is stronger when participants are conversing in a face-to-face spatial orientation than in a back-to-back spatial orientation, irrespective of dyadic visibility; this pattern is enhanced further when they can see each other's visual bodily signals. We discuss these results, and their implications, in more detail below.

First, we demonstrated that intra-brain alpha power was higher when interlocutors were conversing in a back-to-back spatial orientation than in a face-to-face spatial orientation, but did not differ when comparing the face-to-face spatial orientation with a visual occlusion to the face-to-face spatial orientation where visual bodily signals were visible. We interpret the observed differences in alpha power in light of the 'functional inhibition hypothesis', which postulates that higher or increased alpha power reflects a stronger need for inhibition of task-irrelevant brain regions (Jensen and Mazaheri, 2010). This way, alpha power modulations can act as a gating mechanism to route information to task-relevant brain regions. Building on this, we therefore propose that the observed alpha power increases act as a gating mechanism to route information to task-relevant brain regions involved in processing information in conversational interaction. However, the exact specification of which brain regions are involved in this is an issue best addressed with methods that have a better spatial resolution (e.g., by using source reconstruction in a high-density dual-EEG paradigm, by combining MEG/MRI, or in an fMRI paradigm.). Following the functional inhibition hypothesis, we speculate that as social coordination and establishing shared engagement is more effortful in the back-to-back condition, a stronger need for inhibition of task-irrelevant brain regions and routing is required. The higher alpha power in the back-to-back condition might thus demonstrate more effortful neurocognitive processing during linguistic social interaction. Crucially, the fact that alpha power did not differ between FtF and FtFO, suggests that these results cannot be explained by the visibility of visual bodily signals, and that a face-to-face spatial orientation in itself causes the human brain to be particularly engaged in neurocognitive processing during conversation.

Second, we observed stronger delta-speech tracking in the BtB condition, followed by the FtFO, and finally, the FtF condition, but observed no differences between the three conditions when comparing theta-speech tracking. Recent work has suggested that delta- and theta-band tracking have distinct roles at different hierarchical levels during speech processing. Specifically, theta-band tracking has been shown to be modulated by the phonological content of a speech signal (e.g., Ding et al., 2016), whereas increases in delta-band tracking have been observed when higher-level (semantic) information is deficient, indicating more attention-demanding and, consequently, more effortful phonological encoding (Mai and Wang, 2019). Similarly, delta-speech tracking has been associated with attentional selection (Ding and Simon, 2014) and forming temporal predictions about the speech (Arnal et al., 2015), suggesting that delta oscillations could be associated with higher-order cognitive processes, such as attention and semantic comprehension (Keitel et al., 2017).

We interpret our results in line with these studies. Specifically, we here argue that the stepwise differences in delta-speech tracking between the conditions suggest that spatial orientation and dyadic visibility have orthogonal effects on the ease of speech processing: Listeners might need to attend to the speech content more strongly when the semantic and pragmatic information conveyed by visual bodily signals is lacking. Critically, this effect leads to even more effortful speech processing when interlocutors are not seated in a face-to-face spatial orientation; this effect is independent from the availability of visual signals because the back-to-back condition differed from both face-to-face orientation conditions. Importantly, the lack of differences between the three conditions in the theta band suggests that these results cannot be explained by differences in the acoustic features or phonological content of the speech signal. Thus, face-to-face spatial orientation by itself, even in the absence of visibility, appears to induce a brain state beneficial for processing linguistic information.

Finally, we observed stronger inter-brain synchrony in the beta band in the FtF condition than in the FtFO or BtB condition, and stronger inter-brain synchrony in the beta band in the FtFO condition than in the BtB condition. These results advance the existing hyperscanning literature by demonstrating that both dyadic visibility and spatial orientation affect the strength of inter-brain synchrony during social interaction (see Dai et al., 2018; Jiang et al., 2012; Rolison et al., 2020; Vestner et al., 2019; Yun, 2013). Recent work has suggested that increases in beta-band imaginary coherence between interlocutors reflects the establishment of shared perceptuomotor and cognitive representations during human interaction (Dikker et al., 2021).

We here extend these results by proposing that both spatial orientation and dyadic visibility independently affect an interlocutor's ability to establish shared attentional states. We speculate that the semantic and pragmatic information conveyed by visual communicative signals might facilitate joint attention toward coordinating predictive processes during language processing (Dikker et al., 2021; Pérez et al., 2017, 2019; Sebanz et al., 2006; Sebanz and Knoblich, 2021). Crucially, we propose that irrespective of dyadic visibility, interacting in a face-to-face spatial orientation in itself induces a special 'social mode' during conversation that facilitates the formation of shared attentional states during social interaction.

In sum, the current work is the first to demonstrate that both spatial orientation (the awareness of which may be based on belief alone or result also from perceptual acoustic cues) and dyadic visibility affect intra- and inter-brain neural activity during natural, human social interactions. Using an ecologically valid dual-EEG paradigm, we established that listeners' brains engage more strongly while communicating in a face-to-face spatial orientation, irrespective of the presence of visual bodily signals. Second, we provide novel evidence showing that listeners need to attend to the speech of the speaker more strongly when they are not in a face-to-face spatial orientation, even when comparing back-to-back and visually occluded face-to-face spatial orientation. This pattern was further enhanced when interlocutors could also see each other. Importantly, our work also provides first evidence that when two people communicate, their brains are more in sync when communicating face-to-face than back-to-back. That is, a face-to-face spatial orientation in and of itself, irrespective of the visibility of visual bodily signals, appears to be sufficient to induce a special 'social mode' of neurocognitive processing during human social interaction.

Our results thus suggest that in addition to social-interactional abilities such as turn-taking, intention attribution, and visual signaling, face-to-face spatial orientation may be another fundamental feature of human sociality that forms part of the putative Interaction Engine (Levinson, 2006, 2019). One important avenue for future work is to investigate how ontogenetic (Augusti et al., 2010) and phylogenetic (Allanic et al., 2021) factors feed into the development of this apparent human predisposition and its cultural variation (e.g., Brown, 2010).

Limitations of the study

The current research offers several potential follow-up options to further elucidate how spatial orientation affects intra- and inter-brain neural activity during human interaction. Currently, this study is limited in that it used a structured conversation paradigm. The cued alternation between speakers allowed us to more easily separate brain activity related to the speaker and the listener role. The present study lays the foundations for future work based on free conversation to more closely mirror natural, human social interaction. Also, it must be noted that the three conditions we used might have differed in how the speech was directionally perceived, which might have had an influence on listeners' acoustic filtering. Although we did not find any differences in theta-band tracking or in any of the pretests, future research might further control for this by, for example, by running additional pretests that explicitly ask participants to group sound fragments per condition, or letting participants directly compare sound fragments from the three different conditions. However, please note that whether it is just belief or a convergence of belief and acoustic cues that shaped the participants' awareness of their mutual spatial orientation is not critical to the conclusions of our study (i.e., awareness about being in a face-to-face spatial orientation inducing a special social mode for information processing). The main aim was to test whether face-to-face spatial orientation (i.e. the occluded condition) would be sufficient for inducing a change in cognitive processing. What *is* critical in conjunction with this is the fact that what we measured reflects cognitive processing on a level *other than acoustic processing*. As laid out above, the differences we found between the conditions were reflected in oscillatory patterns associated with information processing on semantic, pragmatic and social levels, and not reflected in those typically associated with differences in speech acoustics. Thus, although acoustic differences between conditions may have contributed to causing the awareness of being positioned in different spatial orientations with respect to one another, these potential acoustic differences cannot explain the differences in brain activity that we found between conditions.

Nevertheless, future research may try to disentangle whether a participant's belief of being in a certain spatial orientation is sufficient to induce the inter-brain synchrony results, or whether acoustic cues about mutual spatial orientation alone may be sufficient for inducing the 'social mode' of information processing. In relation with this, a promising future avenue would be to further investigate the effect of participants' prior belief about versus sensory evidence for spatial orientation by making participants unaware of the spatial orientation they are in with respect to one another (e.g. by using an occluder, but letting one participant face the occluder, and the other one not), and to zoom in on the neural sources that might underlie these effects.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Participants
- METHOD DETAILS
  - Task

  - ○ Procedure
  - ○ Dual-EEG data acquisition
- ● QUANTIFICATION AND STATISTICAL ANALYSIS
  - ○ EEG data analysis
  - ○ Statistical analyses

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.105413.

ACKNOWLEDGMENTS

JH, LD, and this research were funded by the European Research Council (CoG grant #773079). LD was partially funded by a Minerva Fast Track Fellowship awarded by the Max Planck Society. We would like to thank Johan Weustink and the Max Planck Institute for Psycholinguistics, for technical assistance and the use of their laboratory architecture, and Lina van Otterdijk, Mareike Geiger, Aran van Hout, Iris van der Wulp and Hanne van Uden for help with data collection. We thank Guillaume Dumas for helpful methodological discussions in early stages of this project.

AUTHOR CONTRIBUTIONS

L.D.: Conceptualization, methodology, software, final analysis, investigation, data curation, writing – original draft, writing – review and editing, visualization, project administration. J.H.: Conceptualization, visualization, writing – original draft, writing – review and editing, supervision, project administration, funding acquisition.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse and equitable conduct of research.

Received: June 20, 2022
Revised: July 20, 2022
Accepted: October 18, 2022
Published: November 18, 2022

REFERENCES

Allanic, M., Hayashi, M., Furuichi, T., and Matsuzawa, T. (2021). Body site and body orientation preferences during social grooming: a comparison between wild and captive chimpanzees and bonobos. Folia Primatol. *92*, 79–90. https://doi.org/10.1159/000512901.

Arnal, L.H., Doelling, K.B., and Poeppel, D. (2015). Delta-beta coupled oscillations underlie temporal prediction accuracy. Cereb. Cortex *25*, 3077–3085. https://doi.org/10.1093/cercor/bhu103.

Augusti, E.-M., Melinder, A., and Gredebäck, G. (2010). Look who's talking: pre-verbal infants' perception of face-to-face and back-to-back social interactions. Front. Psychol. *1*, 161. https://doi.org/10.3389/fpsyg.2010.00161.

Barraza, P., Dumas, G., Liu, H., Blanco-Gomez, G., van den Heuvel, M.I., Baart, M., and Pérez, A. (2019). Implementing EEG hyperscanning setups. MethodsX *6*, 428–436. https://doi.org/10.1016/j.mex.2019.02.021.

Bastos, A.M., and Schoffelen, J.-M. (2016). A tutorial Review of functional connectivity analysis methods and their interpretational pitfalls. Front. Syst. Neurosci. *9*, 175. https://doi.org/10.3389/fnsys.2015.00175.

Bavelas, J.B., and Chovil, N. (2000). Visible acts of meaning: an integrated message model of language in face-to-face dialogue. J. Lang. Soc. Psychol. *19*, 163–194. https://doi.org/10.1177/0261927X00019002001.

Brown, P. (2010). Questions and their responses in Tzeltal. J. Pragmat. *42*, 2627–2648. https://doi.org/10.1016/j.pragma.2010.04.003.

Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., and Ghazanfar, A.A. (2009). The natural statistics of audiovisual speech. PLoS Comput. Biol. *5*, e1000436. https://doi.org/10.1371/journal.pcbi.1000436.

Dai, B., Chen, C., Long, Y., Zheng, L., Zhao, H., Bai, X., Liu, W., Zhang, Y., Liu, L., Guo, T., et al. (2018). Neural mechanisms for selectively tuning in to the target speaker in a naturalistic noisy situation. Nat. Commun. *9*, 2405. https://doi.org/10.1038/s41467-018-04819-z.

De Clercq, W., Vergult, A., Vanrumste, B., Van Paesschen, W., and Van Huffel, S. (2006). Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. IEEE Trans. Biomed. Eng. *53*, 2583–2587. https://doi.org/10.1109/TBME.2006.879459.

De Vos, M., Riès, S., Vanderperren, K., Vanrumste, B., Alario, F.X., Van Huffel, S., and Burle, B. (2010). Removal of muscle artifacts from EEG recordings of spoken language production. Neuroinformatics *8*, 135–150. https://doi.org/10.1007/s12021-010-9071-0.

Dikker, S., Michalareas, G., Oostrik, M., Serafimaki, A., Kahraman, H.M., Struiksma, M.E., and Poeppel, D. (2021). Crowdsourcing neuroscience: inter-brain coupling during face-to-face interactions outside the laboratory. Neuroimage *227*, 117436. https://doi.org/10.1016/j.neuroimage.2020.117436.

Ding, N., Melloni, L., Zhang, H., Tian, X., and Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. Nat. Neurosci. *19*, 158–164. https://doi.org/10.1038/nn.4186.

Ding, N., and Simon, J.Z. (2014). Cortical entrainment to continuous speech: functional roles and interpretations. Front. Hum. Neurosci. *8*, 311. https://doi.org/10.3389/fnhum.2014.00311.

Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., and Garnero, L. (2010). Inter-brain synchronization during social interaction. PLoS One *5*, e12166. https://doi.org/10.1371/journal.pone.0012166.

Giraud, A.-L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. Nat. Neurosci. *15*, 511–517. https://doi.org/10.1038/nn.3063.

Goffman, E. (1961). Encounters: Two Studies in the Sociology of Interaction (Ravenio Books).

Goldin-Meadow, S. (2005). Hearing Gesture: How Our Hands Help Us Think (Harvard University Press).

Holler, J., and Levinson, S.C. (2019). Multimodal language processing in human communication. Trends Cogn. Sci. *23*, 639–652. https://doi.org/10.1016/j.tics.2019.05.006.

Jensen, O., and Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. Front. Hum. Neurosci. *4*, 186. https://doi.org/10.3389/fnhum.2010.00186.

Jiang, J., Dai, B., Peng, D., Zhu, C., Liu, L., and Lu, C. (2012). Neural synchronization during face-to-face communication. J. Neurosci. *32*, 16064–16069. https://doi.org/10.1523/JNEUROSCI.2926-12.2012.

Jung, E., Takahashi, K., Watanabe, K., de la Rosa, S., Butz, M.V., Bülthoff, H.H., and Meilinger, T. (2016). The influence of human body orientation on distance judgments. Front. Psychol. *7*, 217. https://doi.org/10.3389/fpsyg.2016.00217.

Keitel, A., Ince, R.A.A., Gross, J., and Kayser, C. (2017). Auditory cortical delta-entrainment interacts with oscillatory power in multiple fronto-parietal networks. Neuroimage *147*, 32–42. https://doi.org/10.1016/j.neuroimage.2016.11.062.

Kendon, A. (1990). Conducting Interaction: Patterns of Behavior in Focused Encounters (CUP Archive).

Kösem, A., Bosker, H.R., Takashima, A., Meyer, A., Jensen, O., and Hagoort, P. (2018). Neural entrainment determines the words we hear. Curr. Biol. *28*, 2867–2875.e3. https://doi.org/10.1016/j.cub.2018.07.023.

Levinson, S.C. (2019). Interactional foundations of language: the interaction engine hypothesis. In Human Language: From Genes and Brain to Behaviour, P. Hagoort, ed. (MIT Press), pp. 189–200.

Levinson, S.C. (2006). On the human "interaction engine. In Roots of Human Sociality, N.J. Enfield and S.C. Levinson, eds. (Routledge), pp. 39–69. https://doi.org/10.4324/9781003135517-3.

Mai, G., and Wang, W.S.-Y. (2019). Delta and theta neural entrainment during phonological and semantic processing in speech perception. Preprint at bioRxiv. https://doi.org/10.1101/556837.

Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods *164*, 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

McNeill, D. (1992). Hand and Mind: What Gestures Reveal about Thought (University of Chicago press).

Mondada, L. (2009). Emergent focused interactions in public places: a systematic analysis of the multimodal achievement of a common interactional space. J. Pragmat. *41*, 1977–1997. https://doi.org/10.1016/j.pragma.2008.09.019.

Nagels, A., Kircher, T., Steines, M., and Straube, B. (2015). Feeling addressed! the role of body orientation and co-speech gesture in social communication. Hum. Brain Mapp. *36*, 1925–1936. https://doi.org/10.1002/hbm.22746.

Nastase, S.A., Goldstein, A., and Hasson, U. (2020). Keep it real: rethinking the primacy of experimental control in cognitive neuroscience. Neuroimage *222*, 117254. https://doi.org/10.1016/j.neuroimage.2020.117254.

Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S., and Hallett, M. (2004). Identifying true brain interaction from EEG data using the imaginary part of coherency. Clin. Neurophysiol. *115*, 2292–2307. https://doi.org/10.1016/j.clinph.2004.04.029.

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. *2011*, 156869. https://doi.org/10.1155/2011/156869.

Pérez, A., Carreiras, M., and Duñabeitia, J.A. (2017). Brain-to-brain entrainment: EEG interbrain synchronization while speaking and listening. Sci. Rep. *7*, 4190. https://doi.org/10.1038/s41598-017-04464-4.

Pérez, A., Dumas, G., Karadag, M., and Duñabeitia, J.A. (2019). Differential brain-to-brain entrainment while speaking and listening in native and foreign languages. Cortex *111*, 303–315. https://doi.org/10.1016/j.cortex.2018.11.026.

Redcay, E., and Schilbach, L. (2019). Using second-person neuroscience to elucidate the mechanisms of social interaction. Nat. Rev. Neurosci. *20*, 495–505. https://doi.org/10.1038/s41583-019-0179-4.

Riès, S.K., Xie, K., Haaland, K.Y., Dronkers, N.F., and Knight, R.T. (2013). Role of the lateral prefrontal cortex in speech monitoring. Front. Hum. Neurosci. *7*, 703. https://doi.org/10.3389/fnhum.2013.00703.

Rolison, M.J., Naples, A.J., Rutherford, H.J.V., and McPartland, J.C. (2020). The presence of another person influences oscillatory cortical dynamics during dual brain EEG recording. Front. Psychiatr. *11*, 246. https://doi.org/10.3389/fpsyt.2020.00246.

Schegloff, E.A. (1998). Body torque. Soc. Res. *65*, 535–596.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., and Vogeley, K. (2013). Toward a second-person neuroscience. Behav. Brain Sci. *36*, 393–414. https://doi.org/10.1017/S0140525X12000660.

Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: bodies and minds moving together. Trends Cogn. Sci. *10*, 70–76. https://doi.org/10.1016/j.tics.2005.12.009.

Sebanz, N., and Knoblich, G. (2021). Progress in joint-action research. Curr. Dir. Psychol. Sci. *30*, 138–143. https://doi.org/10.1177/0963721420984425.

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

Tognoli, E., Lagarde, J., DeGuzman, G.C., and Kelso, J.A.S. (2007). The phi complex as a neuromarker of human social coordination. Proc. Natl. Acad. Sci. USA *104*, 8190–8195. https://doi.org/10.1073/pnas.0611453104.

Vestner, T., Tipper, S.P., Hartley, T., Over, H., and Rueschemeyer, S.-A. (2019). Bound together: social binding leads to faster processing, spatial distortion, and enhanced memory of interacting partners. J. Exp. Psychol. Gen. *148*, 1251–1268. https://doi.org/10.1037/xge0000545.

Wheatley, T., Boncz, A., Toni, I., and Stolk, A. (2019). Beyond the isolated brain: the promise and challenge of interacting minds. Neuron *103*, 186–188. https://doi.org/10.1016/j.neuron.2019.05.009.

Yun, K. (2013). On the same wavelength: face-to-face communication increases interpersonal neural synchronization. J. Neurosci. *33*, 5081–5082. https://doi.org/10.1523/JNEUROSCI.0063-13.2013.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Fieldtrip v20200409	Oostenveld et al. (2011)	https://www.fieldtriptoolbox.org/download/
MATLAB v2018b	MathWorks	https://www.mathworks.com/products/matlab.html
Circular Statistics Toolbox	Philipp Berens	https://jp.mathworks.com/matlabcentral/fileexchange/10676-circul
Data used in this paper	This paper	https://osf.io/m9rvy/
Code used in this paper	This paper	https://osf.io/m9rvy/
Other		
Human subjects (52)	This paper	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Linda Drijvers (linda.drijvers@mpi.nl).

Materials availability

This study did not generate new unique reagents or materials.

Data and code availability

- Data are available via https://osf.io/m9rvy/.
- Code is available via https://osf.io/m9rvy/.
- Any additional information required to reanalyze the data reported in this study is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Participants

52 native Dutch speakers (26 dyads; 32 females; mean age = 22.46, SD = 2.69, range = 18–30 years) participated in the experiment. All participants were right-handed, reported no neurological, language, hearing or motor disorders, and reported normal or corrected-to-normal vision. All 26 dyads knew each other prior to the experiment (mean *n* of months: 38.82, SD = 33.14, range = 1–144 months) and were of equivalent age (no more than 5 years different in age). We included dyads that knew each other to help the ease with which participants engaged in the conversation, as they had to talk to each other for 30 min. Including only participants that were familiar with each other would ensure that any neural patterns that we found were not due to our participants not being used to seeing, hearing or talking to the other speaker, or any other conversational difficulties (e.g., running out of things to talk to in a later block in the experiment) caused by them being strangers. All participants were recruited via the Max Planck Institute for Psycholinguistics database and were paid for their participation. The study was conducted in accordance with the Declaration of Helsinki and approved by the local ethics committee. All participants gave informed written consent at the start of the experiment.

METHOD DETAILS

Task

We asked participants to engage in a structured conversation and alternate between the roles of speaker and listener while they had to either plan a festival, a dinner, or a holiday together. Participants were asked to communicate with each other as naturally as possible, and were not restricted in their verbal or bodily behaviour. The three conversational topics (i.e., planning a festival, a dinner or a holiday) were each assigned to a 10-min block, resulting in three 10-min blocks in total. At the start of each block, a short sound played to indicate the start of the block. Then, after 60 s, another sound would play to indicate that the participants had to change roles from speaker to listener or from listener to speaker. The end of each block was indicated by another sound cue.

The participants were seated in three different spatial orientations during the experiment: face-to-face (FtF), face-to-face with an occluding screen between the interlocutors (FtFO), or back-to-back (BtB) (see Figure 3). They were seated in a 7.78 m^2 (2.79 × 2.79m) electrically shielded room with grey walls, all of equal size. Reverberance and noise was reduced by padding the walls with specialized panels from Ecophon (type Super G Plus A, 40 mm, 200 mm, see supplemental information for acoustic tests of these wall panels). The participants were asked to not turn around in any of the conditions, and keep facing forward. Which participant could speak first, as well as the order of conversational topics and the order of the spatial orientations was counter-balanced over dyads. The participant's speech was recorded by a Sennheiser ME 64 microphone that was placed between the participants, and we recorded videos of the conversations using two Canon Legria HFG30 cameras that were placed on a table next to the participants. Three pretests were conducted to ensure that the three different spatial orientations that we used did not differ in terms of sound quality, sound level or other auditory characteristics.

Pretest 1

In the first pre-test, we asked two volunteers that did not participate in the main experiment to talk to each other in the three different spatial configurations while we measured potential sound level differences between the conditions from both the speaker and listener's perspective, by placing a microphone approximately 20 cm below the speaker's mouth, attached to their collar, as well as by placing two microphones on the shoulders of the listener. The rationale behind this pretest was that we reasoned that speakers might raise their voice more when they are unable to see their conversational partner, or that there might be differences in sound level from the listener's perspective when comparing the three different spatial orientations. For example, when talking in a back-to-back spatial orientation, the walls that oppose the speaker could, despite their sound absorbing qualities (see Figure S1, Table S1) reinforce the perceived sound level for the listener sitting behind them.

From the speaker's perspective, the sound level difference was less than 1 dB(A) between the different conditions (BtB: 60.9 dB(A), FtFO: 60.4 dB(A), FtF: 60.1 dB(A)). Similar results were observed from the listener's perspective (BtB: 58.3 dB(A), FtFO: 58.9 dB(A), FtF: 59.1 dB(A)). We therefore concluded that there were no significant differences between the three different spatial orientations.

Pretest 2

Although we did not observe any differences in sound level in Pretest 1, we used the recorded audio of the conversation that was held in Pretest 1 to investigate whether participants could reliably identify in which spatial configuration the conversation was held by just listening to the audio of the conversation. The rationale behind this pretest was that there could potentially be other auditory characteristics that differed between the spatial orientations. For this purpose, we selected 30 10-s fragments of the conversation from Pretest 1 (10 per condition), and played these fragments to 5 participants that did not participate in the main experiment or other pretest. We then asked these participants to indicate whether the sound fragments they heard took place in a face-to-face spatial orientation, a face-to-face orientation with an occluding screen, or a back-to-back spatial orientation. On average, participants were able to correctly identify 11.8 out of 30 items ($SD = 3.7$), which is at chance level. We therefore concluded that participants were not able to identify which spatial orientation they were listening to, and consequently that there were no audible differences between the three spatial orientations for participants in the three conditions.

Pretest 3

Finally, to ensure that there were no other noticeable differences between the three conditions, we asked another 5 participants to listen to the same sound fragments that we used in Pretest 2. We instructed these participants that we were trying out different sound recording materials for a new experiment, and asked them to note down whether there was anything they noticed in terms of sound quality in the different recordings. None of their answers indicated that any auditory characteristics were different in the three different spatial orientations.

Behavioral questionnaires

After the main experiment, we asked the participants to complete a few behavioral questionnaires. First, participants were asked to fill out a general questionnaire about their demographics. Second, participants were asked to indicate how long they have known their conversational partner, and how close their relationship was. These questionnaires were followed by the Autism Quotient questionnaire, and the Empathy Quotient questionnaire. These questionnaires have not been further analysed for this research.

Procedure

Upon arrival, participants were informed about the experimental procedure, completed the informed consent forms, and were fitted with an EEG cap. The participants were then seated approximately 1.5 m apart from each other in one of the three spatial orientations. The experimenter then explained the rest of the procedure of the experiment, when the participants would have to change speaker role, and when the experimenter would come in again to change the spatial orientation the participants were seated in. The participants were aware of their mutual spatial orientation. The experimenter asked the participants were asked to sit as relaxed as possible, and not change position during the experiment.

The experiment would start with a sound cue, which was immediately followed by another sound cue that indicated that the first speaker could start talking about one of the three activities that the participants needed to plan together (a festival, dinner or holiday). After 60 s, the next sound would indicate that the participants had to change roles. The sound cues that indicated a change of speaker were presented using Presentation (version 20.2, Neurobehavioural Systems Inc.). Participants spoke about one topic for 10 min, and after 10 min the experimenter would come in again to change the spatial configuration the participants were seated in, and to instruct them on the next topic that should be discussed. After the experiment, participants filled out the behavioral questionnaires.

Dual-EEG data acquisition

We followed the recommended dual-EEG setup and procedures described in (Barraza et al., 2019) and recorded the EEG signals of the two participants simultaneously from 32 AG-AgCl electrodes, of which 27 were mounted in a cap (actiCap) according to the 10–20 standard system. One electrode was placed on the right mastoid for re-referencing, and 4 were used for bipolar horizontal and vertical electrooculograms (EOG). The ground electrode was placed on the nasion, and electrode impedance was kept below 5 kOhm at the start of the experiment. The EEG data was acquired at a 500 Hz sampling rate. Each participant had their own ground and reference electrode, and impedance measurements were done separately for each participant.

QUANTIFICATION AND STATISTICAL ANALYSIS

EEG data analysis

All data were analysed using the FieldTrip toolbox (Oostenveld et al., 2011) running in a MATLAB environment (version 2018b, MathWorks, Natick, MA). We re-referenced the EEG data online to the right mastoid and filtered the data with a high-pass filter at 0.5 Hz and a low-pass filter at 35 Hz. The datasets were then split into two separate recordings, corresponding to each participant. Subsequently, we segmented the raw data into 1-s epochs. Artifacts were removed by using a semi-automatic rejection routine. We used a Blind Source Separation method based on Canonical Correlation Analysis (BSS-CCA) to separate cortical sources from electromyographic responses (De Vos et al., 2010). This method has been successfully applied to remove any artifacts in the EEG signal that are caused by speech artifacts and attenuates muscle contamination on EEG recordings (see DeClercq et al., 2006; Riès et al., 2013). After this procedure, we inspected all epochs on an individual basis to remove artifacts that were not identified using the abovementioned rejection procedures. On average, we excluded 7.46% of epochs and 3.75 sources per participant. We recorded a silent period per block to use as a baseline in our analyses.

Spectral analysis

To investigate intra-brain low-frequency power spectra, we focused on the periods in which a participant was listening to the other participant. For these listening periods, we calculated spectral power in the alpha band (8–12 Hz) based on Fast Fourier Transformation after application of a single Hanning taper. These spectra were calculated per participant and averaged per condition.

Brain-speech envelope synchronization

We estimated the brain-speech envelope tracking between a listener's EEG signal and the speaker's speech amplitude envelope. To achieve this, the wideband speech envelope was computed by band-pass filtering the auditory signal with a 3rd order forward and reverse Butterworth filter into ten bands (100–10000 Hz, equidistant on the cochlear frequency map; see Chandrasekaran et al., 2009; Keitel et al., 2017). We used the magnitude of the Hilbert transform to obtain band-limited envelopes, which were then averaged to obtain the wide-band speech envelope. The speech envelope was then resampled to 500 Hz to match the sampling rate of the EEG signals. We then calculated the coherence between the listener's EEG signal and the speaker's speech envelope in both the delta (1–3 Hz) and theta (4–7 Hz) bands. In this context, this measure indicates the phase locking consistency of the speech signal and the neuronal oscillatory activity across epochs. Both the speech envelope and EEG data were transformed to the frequency domain using a Fast Fourier Transformation with a Hanning window, restricted to a frequency range from 1–8 Hz.

Inter-brain synchronisation

We followed a similar approach to Dikker et al. (2021) and used imaginary coherence (Nolte et al., 2004) to calculate the synchronisation between the brain signals of the two participants. Instantaneously synchronized EEG signals without any relevant neurocognitive interpretation, which is often manifested as signals with a 0 or 180 degree phase difference across electrodes or participants, often occur especially in situations where the data might be heavily contaminated with experimental noise. In the current experiment, the two participants were seated in a small room that included a lot of audio and video electronic equipment in close vicinity to the measurements. As described in Dikker et al.(2021) using imaginary coherence is especially suited for the analysis of dual-EEG data that might be heavily contaminated by such environmental noise, and has shown to reliably detect brain-to-brain interactions during movement in EEG data (Nolte et al., 2004).

In more detail, imaginary coherence is calculated by computing the power and cross-spectral density between two series of complex spectral coefficients. Then, coherency is computed and expressed as a complex number with a real and imaginary part. The imaginary part represents how much of the magnitude is driven by lagged interactions of the two signals, whereas the real part represents how much of the magnitude is driven by instantaneous interactions (see for a detailed description Dikker et al., 2021; Nolte et al., 2004). Imaginary coherence captures phase differences between signal fluctuations in the range of tens of milliseconds, and is insensitive to volume conduction (Bastos and Schoffelen, 2016). We computed imaginary coherence for all between-participant electrode pairs in the alpha (8–12 Hz) and low beta (13–20 Hz) bands, and averaged these values in each condition.

Statistical analyses

Intra-brain spectral analysis & brain-speech envelope tracking

Nonparametric cluster-based permutation tests were performed across subjects to statistically assess differences in alpha power (averaged over 8–12 Hz) and brain-speech envelope coherence (averaged over 1-3Hz (delta) and 4–7 Hz (theta)) between the different conditions (BtB vs FtF, BtB vs FtFO, FtFO vs FtF) (Maris and Oostenveld, 2007). For both types of analyses, we calculated the difference between every data point of the two conditions that were compared. All adjacent data points that exceeded a pre-set threshold of 5% were then grouped into clusters. In each of these clusters, we then summed the difference values to calculate the cluster-level statistics. A Monte-Carlo permutation distribution was created by randomly assigning a participant's average value to one of the two conditions and subsequently selecting the largest cluster-level statistic for every permutation (n = 5000). Cluster-level statistics were then calculated for the measured data and compared against the permutation distribution. Clusters that fell in the highest or lowest 2.5th percentile were considered significant.

Inter-brain synchronisation

Similar to the spectral and brain-speech envelope analyses, we used nonparametric cluster-based permutation tests across dyads to statistically assess differences in inter-brain synchronisation in the alpha band and beta band between the different conditions. To cluster across the two different brains, we followed a similar approach as Dumas et al. (2010), and designed a custom neighbourhood structure that was appropriate to cluster over electrode pairs between two participants. Following their approach, pairs of electrodes between two EEG caps were formed by one electrode on the cap of participant A, and one on the cap of participant B. Two of such pairs were considered neighbours when two electrodes were adjacent to each other on the cap of participant A, connected to two adjacent electrodes on the cap of participant B, or when one electrode on the cap of participant A connected to two adjacent electrodes on the cap of participant B, or the other way around.