Title:

A *studyforrest* extension, MEG recordings while watching the audio-visual movie "Forrest Gump"

Authors: Xingyu Liu[1], Yuxuan Dai[1], Hailun Xie[1], Zonglei Zhen[1,2]

Affiliation:[1]Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education, Faculty of Psychology, Beijing Normal University, Beijing, China. [2]State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China.

Correspondence address: Zonglei Zhen, Ph.D. (zhenzonglei@bnu.edu.cn), Faculty of Psychology, Beijing Normal University, Beijing, 100875, China.

Acknowledgements: This study was funded by the National Key R&D Program of China (Grant No. 2019YFA0709503), and the National Natural Science Foundation of China (Grant No. 31771251).

Conflict of Interest: None declared.

Abstract

Naturalistic stimuli, such as movies, are being increasingly used to map brain function because of their high ecological validity. The pioneering *studyforrest* and other naturalistic neuroimaging projects have provided free access to multiple movie-watching functional magnetic resonance imaging (fMRI) datasets to prompt the community for naturalistic experimental paradigms. However, sluggish blood-oxygenation-level-dependent fMRI signals are incapable of resolving neuronal activity with the temporal resolution at which it unfolds. Instead, magnetoencephalography (MEG) measures changes in the magnetic field produced by neuronal activity and is able to capture rich dynamics of the brain at the millisecond level while watching naturalistic movies. Herein, we present the first public prolonged MEG dataset collected from 11 participants while watching the 2 h long audio-visual movie "Forrest Gump". Minimally preprocessed data was also provided to facilitate the use. As a *studyforrest* extension, we envision that this dataset, together with fMRI data from the *studyforrest* project, will serve as a foundation for exploring the neural dynamics of various cognitive functions in real-world contexts.

Background & Summary

The mechanisms of human brain function in complex dynamic environments is the ultimate mystery that cognitive neuroscience aspires to quest. Most of the existing models on brain function have been obtained from tightly controlled experimental manipulations on carefully designed "artificial" stimuli. However, these simple stimuli are often irrelevant to ecological scenarios encountered in real-world environments, in terms of quantity, complexity, modality, and dynamics. To address this issue, naturalistic stimuli that encode a wealth of real-life content have become increasingly popular for understanding brain function in ecological contexts. Researchers have achieved significant advances in the areas of human memory, attention, language, emotions, and social cognition using naturalistic stimuli (for recent reviews, please refer to Sonkusare et al., 2019[1] and Jääskeläinen et al., 2021[2]). Simultaneously, emerging deep learning technologies that could afford multiple levels of representations for naturalistic stimuli are continuously expanding the application of naturalistic stimuli for exploring human brain function[3–6].

Notably, owing to its dynamics and multimodal content, movies have been successfully utilized as naturalistic stimuli to examine the mechanism by which the brain processes diverse psychological constructs and dynamic interactions. Functional magnetic resonance imaging (fMRI) is commonly employed to measure brain activity while watching a movie. In particular, the pioneering *studyforrest* and other naturalistic neuroimaging projects have released multiple fMRI datasets collected from participants who watched movie clips[7–11]. However, fMRI measures the relatively sluggish blood-oxygenation-level-dependent signal, therefore falling short of characterizing the complex neural dynamics underlying the cognitive processing of dynamic movies. In contrast, magnetoencephalography (MEG) measures the magnetic fields generated by neuronal activity on a millisecond time scale. Thus, it has great potential to pry open neural dynamics in processing naturalistic stimuli. Several studies have leveraged MEG to investigate brain activity for naturalistic movie stimuli in a short period ($\leqslant$20 min)[12–16]. So far, however, there is still a dearth of publicly accessible MEG recordings for naturalistic stimuli, especially prolonged MEG recordings for dynamic movies that are more likely to capture the temporal dynamics of regular functional brain states that occur in everyday life, and further contribute to unraveling human brain function in ecological contexts.

Herein, we present an MEG dataset obtained while watching the 2 h long audio-visual movie "Forrest Gump" (R. Zemeckis, Paramount Pictures, 1994). The recordings measure brain activation with a temporal resolution at the millisecond level, thus providing a timely and efficient extension to the *studyforrest* dataset. Specifically, MEG data were collected from 11 participants while they were watching the Chinese-dubbed movie "Forrest Gump" in eight consecutive runs, each lasting for roughly 15 min. High-resolution structural MRI was additionally acquired for all participants, thereby allowing the incorporation of the detailed anatomy of the brain and head in the source localization of MEG signals. Together with the raw data, preprocessed MEG and MRI data with standard pipelines were also provided to facilitate the use of the data. Considering MEG and fMRI are complementary to each other, synergy between our present MEG recordings and fMRI data from the *studyforrest* project will serve as a front to elaborate brain function in the wild. We believe the dataset is suitable for addressing many questions pertaining to the neural dynamics of various aspects, including perception, memory, language, and social cognition.

Methods

Participants

A total of 11 participants (mean±SD age: 22±1.7 years, 6 female) from the Beijing Normal University, Beijing, China, volunteered for this study. They completed both the MEG and MRI sessions. All participants were right-handed native Chinese speakers, with normal hearing and normal or corrected-to-normal vision. None of them had ever watched the film "Forrest Gump" before, except one who had watched some clips, however not the entire movie. Of the 10 participants, four had heard about the movie plot, while others did not. The study was approved by the Institutional Review Board of the Faculty of Psychology, Beijing Normal University. Written informed consent was obtained from all participants, prior to their participation. All participants provided additional consent for sharing their anonymized data for research purposes.

Procedures

Fig. 1 depicts the overall flow of data collection and preprocessing. Prior to data acquisition, all participants completed a questionnaire on their demographic

information and familiarity with the movie "Forrest Gump". The data acquisition consisted of two sessions for each participant, namely one MEG session to record their neural activities during movie watching and an MRI session with a T1-weighted (T1w) scan to measure the brain structure for the spatial localization of the MEG signal. The MRI scan immediately followed the MEG session for all participants, except for sub-07 and sub-11, who finished their MRI session a week later.

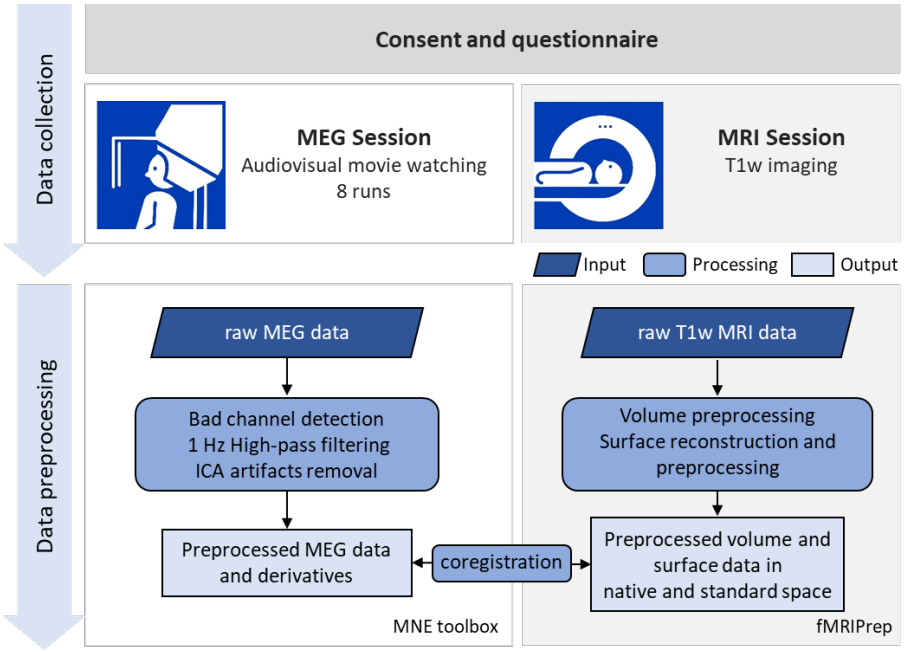


Fig. 1. Schematic of the data collection and preprocessing procedure. Data collection comprised one MEG session, followed by one MRI session. The neuromagnetic signals were recorded with a whole-scalp-covering MEG while the participants watched the audio-visual movie "Forrest Gump". An anatomical T1w imaging was acquired in the MRI session. The raw MEG data and MRI data were preprocessed with MNE and fMRIPrep toolbox, respectively. The MEG-MRI coregistration was performed on the preprocessed data.

Stimulus material and presentation

The audio-visual stimuli were generated from the Chinese-dubbed "Forrest Gump" DVD, released in 2013 (ISBN: 978-7-7991-3934-0). The movie was split into eight segments, each of which lasted for approximately 15 min. The stimuli were initially obtained by concatenating all original VOB files from the DVD release into one MPEG-4 file, using FFmpeg (https://ffmpeg.org). The concatenated MPEG-4 file contained a video stream and a Chinese-dubbed audio stream, which was down-mixed from multi-channel to 2Channel stereo. The stimuli were then divided into eight

segments using Adobe Premier software (Adobe Premiere Pro CC 2017, Adobe, Inc., San Jose, CA, USA). Each segment conformed to the following specifications: video codec=avc1, display aspect ratio=4:3, resolution=1024×768 pixels, frame rate=25 FPS, color space=YUV, video bit depth=8 bits, audio codec=mp4a-40-2, audio sampling rate=48.0 kHz, and audio channels=2. Each successive segment began with a 4 s repetition of the end of the previous segment. It should be noted that the Chinese-dubbed "Forrest Gump" was slightly abridged than the German version. To align with the stimuli of the *studyforrest* dataset as much as possible, a short clip from the German-dubbed DVD released in 2011 (EAN: 4010884250916) was added to our stimuli. Table 1 summarizes the alignment of the stimuli sources from both the Chinese and German versions.

Segment	Frames	Duration	Start (cn)	End (cn)	Start (de)	End (de)
1	22499	15:00.07	63	22562	35	22534
2	22599	15:04.08	22463	32374	22438	32349
			36410	49098	36385	49073
3	22599	15:04.08	48999	57860	48974	57835
			58531	63717	58506	63692
			-	-	63692	64621
			63718	71341	64621	72244
4	22599	15:04.08	71242	85132	72146	86036
			88427	97136	89332	93902
					94464	98603
5	22599	15:04.08	97037	111719	98504	105793
					109959	117352
			115101	118317	120733	123949
			118797	123498	125347	130048
6	22599	15:04.08	123398	145602	129948	152152
			147736	148131	154286	154681
7	22599	15:04.08	148032	170631	154582	177181
8	17661	11:46.56	170532	188193	177082	194743

Table 1. Timing alignment for stimuli from the Chinese (cn) and German (de) version of "Forrest Gump". The start time and end time are different for the same movie clip in the two versions. Moreover, a clip with only background sound stream from the German version (frames from 63692 to 64621) was added into the third segment.

The visual stimuli were projected onto a screen in full-screen mode from a DLP projector with 1024×768-pixel resolution, using Psychophysics Toolbox Version 3[17] in MATLAB 2016 (The MathWorks, Inc., Natick, Massachusetts, USA). The participants watched the visual stimuli on a rear projection screen through mirror reflection (visual field angles=31.17°×23.69°; viewing distance=751 mm). The audio

stimuli were delivered to the participants using foam ear-tips connected to a loudspeaker via an air-conducting tube. The average delay of the peripheral devices was 33 ms and 15 ms for the visual and audio displays, respectively. The participants were instructed to watch the movie, without other tasks and keep still as best as possible.

MEG data acquisition

MEG data were recorded using a 275-channel whole-head axial gradiometer DSQ-3500 MEG system (CTF MEG, Canada) at the Institute of Biophysics, Chinese Academy of Sciences, Beijing, China. Three channels (i.e., MLF55, MRT23 and MRT16) were out of service due to failure of sensors. The neuromagnetic signals were recorded in continuous mode at a sampling rate of 600 Hz, without online digital band filters. A third-order synthetic gradiometer was employed to remove far-field noise. The precise timing of each frame was recorded. Upon presenting each frame to the participants, we recorded a trigger pulse lasting for five samples in the stimuli channel UPPT001, along with the MEG signals. The beginning of the movie was indicated with a value of 255. Owing to the limited bit-width of the stimuli channel, the frame number was unable to be marked with an accurate value > 20,000. The frame numbers were therefore marked as the ceiling of the timestamp of that frame divided by 10, resulting in step-like increasing marker values with a step width of 10 s.

At the beginning of each session, three HPI coils were attached to the participants' nasion (NAS), left preauricular (LPA), and right preauricular (RPA) points to continuously measure their head position during the MEG dewar. A customized wooden chin-rest supporter was introduced to prevent possible head movements. The MEG session consisted of eight runs, with each run playing one movie segment. Eight segments were played chronologically. The participants took a self-paced break between runs. Following the completion of the MEG scan, the participants underwent an anatomical T1w scan. The HPI coils were replaced with three customized MRI-compatible vitamin E caplets in the MRI scan to provide spatial reference for the spatial alignment between the MEG and MRI data.

MEG data preprocessing

MEG data processing was performed offline using the MNE-Python package[18]. The MEG preprocessing pipeline was conducted at the run level (Fig. 1). First, the bad channels were detected and marked. As a result, no bad channels were identified in all acquisitions except two in run-05 of sub-05. Second, a high-pass filter of 1 Hz was applied to remove possible slow drifts from the continuous MEG data (Fig. 2a). Finally, artifact removal was performed using an independent component analysis (ICA). The number of independent components (IC) was set to 20. Two raters (i.e., X.L. and Y.D.) manually identified the head movement, eye movement, eye blinks, and cardiac artifacts (Fig. 2b). On average, 3.21 ICs (SD: 0.85) were classified as artifacts. The denoised MEG data were eventually reconstructed from all the non-artifact components and residual components (Fig. 2c). Both the raw and preprocessed data were provided in the released dataset.

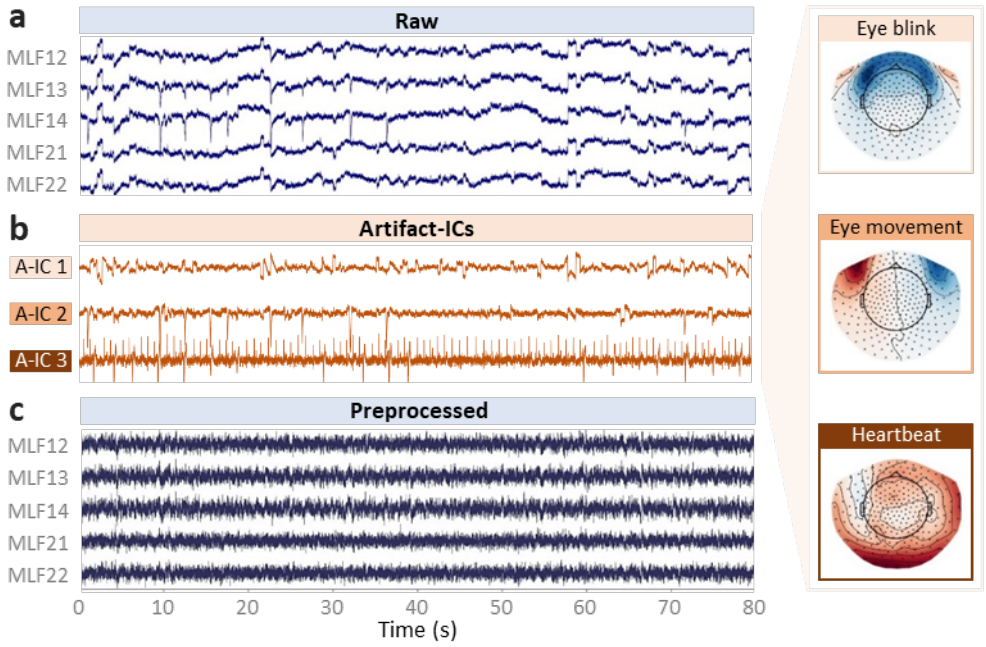


Fig. 2. Typical artifact-ICs and MEG signals from the pre- and post-preprocessing data. (a) MEG signals of example channels from the raw (i.e. pre-preprocessing) data. **(b)** Timeseries and scalp field distribution of three typical artifact-ICs (A-ICs), namely A-IC 1 for eye blink, A-IC 2 for horizontal eye movement, and A-IC 3 for heartbeat. **(c)** MEG signals of example channels from the preprocessed data. Data from the run-04 of sub-04 was used for this illustration.

MRI data acquisition and preprocessing

High-resolution anatomical MRI was collected for each participant using a 3T SIEMENS Prisma^{fit} scanner (Siemens Healthcare GmbH, Erlangen, Germany), with a 20-channel headneck coil. All participants underwent a T1w scan with a 3-D magnetization-prepared rapid gradient-echo pulse sequence with identical parameters (TR=2530 ms, TE=1.26 ms, TI=1100 ms, flip-angle=7˚, 176 sagittal slices, slice thickness=1 mm, matrix size=256×256, and voxel size=1.0 ×1.0 mm), except that sub-01 was scanned with slightly different parameters (TR=2200 ms, TE=3.37 ms, TI=1100 ms, flip-angle=7˚, 192 sagittal slices, slice thickness=1 mm, matrix size=224×256, and voxel size=1.0 ×1.0 mm). Earplugs were used to attenuate the scanner noise. A foam pillow and extendable padded head clamps were applied to restrain the head motion.

The raw DICOM files of T1w images to NIFTI files using dcm2niix (https://github.com/rordenlab/dcm2niix). The T1w images were then minimally preprocessed using the anatomical preprocessing pipeline from fMRIPrep 20.2.1, with default settings[19]. In brief, the T1w data were skull-stripped and corrected for intensity nonuniformity with ANTs and N4ITK[20]. Brain surfaces were reconstructed using FreeSurfer[21]. Spatial normalization to both MNI152NLin6Asym and MNI152NLin2009cAsym was performed through nonlinear registration with ANTs, using the brain-extracted versions of both T1w volume and template.

MEG-MRI coregistration procedure

To reconstruct the source of MEG sensor signals, MEG data were co-registered with the high-resolution anatomical T1w MRI data for each participant. The NAS, LPA, and RPA points marked in both MEG and MRI sessions were used as fiducial points for the alignment of the MEG and MRI data. Specifically, following the generation of a high-resolution head surface using MNE make_scalp_surfaces based on FreeSurfer reconstruction, we performed MEG-MRI coregistration for each participant in the MNE COREG GUI[18]. First, the three fiducial points were manually pinned on the MRI-reconstructed head surface. an iterative algorithm (nearest-neighbor calculations) was then ran to align the MEG and MRI coordinates. The co-registration was refined by manual adjustment. The results showed that the averaged distances between the three fiducials in the coregistered MEG and MRI coordinate systems were 0.96 mm, 4.22 mm, and 4.90 mm for NAS, LPA, and RPA respectively.

Both the MRI-fiducials files and MEG-MRI coordinate transformation files were included in the released data.

Data Records

The dataset can be accessed at OpenNeuro (dataset accession number: ds003633, version 1.0.1, https://openneuro.org/datasets/ds003633/versions/1.0.1)[22]. The facial information was removed from the published dataset using pydeface (https://github.com/poldracklab/pydeface) to ensure anonymity. The data was organized according to the MEG-Brain Imaging Data Structure (MEG-BIDS)[23] using the MNE-BIDS toolbox[24] (Fig. 3). Besides dataset and participant description files, the data were sorted into different directories, including "sub-<participant_id>," "derivatives," and "code" directory for raw data from each participant, preprocessed data, and the code used for stimuli preparation and presentation, data preprocessing, respectively (Fig. 3a).

Fig. 3. The file structure of the dataset. (a) File structure of the project directory. **(b)** File structure of the raw data for each individual participant. **(c)** File organization of the derived (preprocessed) data.

Raw data

The raw data of each participant were stored separately in the "sub-<participant_id>" folders (Fig. 3b), consisting of two subfolders, namely "anat" and "meg". The T1w MRI data ("*T1w.nii.gz") and associated sidecar json files were located in the "anat" folders. The raw MEG data were provided as CTF ds files ("*_meg.ds") for each run, and located in the "meg" folder along with sidecar json files. In addition, "*_channel.tsv" files with MEG channel information, "*_events.tsv" files with the presentation timing of stimuli frames and "*coordsystem.json" files with coordinate system information of the MEG sensors were included in the "meg" folder.

In parallel with the "sub-<participant_id>" directories, a "sub-emptyroom" directory hosted empty-room MEG measurements, which recorded the environmental noise of the MEG system. The empty-room measurements lasted for 34 s and were acquired on each data acquisition day, except for the day 20190603.

Preprocessed data

All preprocessed data were deposited in the "preproc_meg-mne_mri-fmriprep" subdirectory under the "derivatives" (Fig. 3c). The preprocessed data of each participant were separately saved in the "sub-<participant_id>/ses-movie" directory, which contains two subfolders, namely "anat" and "meg". The "anat" folder comprised the preprocessed MRI volume, reconstructed surface, and other associations, including transformation files. The "meg" folder included preprocessed MEG recordings, including "*_meg.fif.gz", "*_ica.fif.gz" and "*_decomposition.tsv", and "*_trans.fif" for the preprocessed data, ICA decomposition, and MEG-MRI coordinate transformation, respectively. In addition, the FreeSurfer surface data, the high-resolution head surface ("freesurfer/sub-<participant_id>/bem/*"), and the MRI-fiducials ("freesurfer/sub-<participant_id>/bem/*fiducials.fif") were provided in "freesurfer/sourcedata" directory for MEG-MRI coregistration.

Technical Validation

We assessed the data quality of both the raw and preprocessed data using four measures as follows: head motion magnitude, stimuli-induced time-frequency characteristics, homotopic functional connectivity (FC), and inter-subject correlation (ISC).

Motion magnitude distribution

Head movements during MEG scans are one of the significant factors that degrade both sensor- and source-level analyses. Herein, we calculated the motion magnitude for each sample as the Euclidian distance between the current and the initial head position while the movie segment began playing. The head motion across all runs and all participants were summarized for each of the three fiducials (NAS, LPA, and RPA) to provide an overview of the head movement of the dataset. As show in Fig. 4, motion magnitude of 95% of the samples had head motions lower than 3.43 mm, 4.11 mm, and 3.87 mm for NAS, LPA, and RPA, respectively. Furthermore, 50% of the samples had head motions smaller than 0.99 mm, 1.10 mm, and 1.46 mm for NAS, LPA, and RPA, respectively. These findings indicated low head motion magnitude on average. The head motion magnitude of each participant could be found in Supplementary Fig. 1.

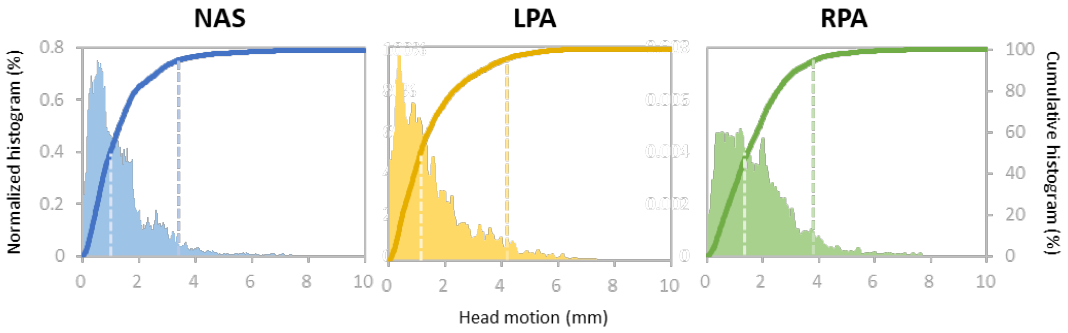


Fig. 4. The ensemble distribution of head motion magnitude across all runs and all participants. The density and accumulative histogram of motion magnitude of all samples from all acquisitions for three fiducials (NAS, LPA, and RPA) have been plotted. The dashed lines indicate the 50% and 95% of the cumulative density. Left Y-axis: normalized histogram; Right Y-axis: cumulative histogram.

Time-frequency characterization of brain activity

Herein, we validated whether MEG recordings could successfully detect the change in stimuli-induced brain activity. Because the movie stimuli do not have explicit condition structures as in conventional design, we selected two exemplar movie clips, within which the audio or visual features showed pronounced changes to examine if the expected change in MEG signals could be detected at the related sensors. In one clip (Seg 3: frame 15864±125), the audio features changed

significantly (a vocal voice developed from background music) whereas the visual features were stable. In contrast, the other clip (Seg 1: frame 21768±125) comprised stable audio features, whereas the visual features changed from landscape to human figures. Validation was performed according to the following procedure[25–27]: First, time-frequency analysis with Morlet wavelets was conducted for each sensor in the occipital and temporal lobes. The baseline was set to 1000 ms before the change points of the audio or visual features, and the baseline mean was subtracted for each channel. Second, the time-frequency representations were averaged across the participants. Finally, the time-frequency representations were averaged across the sensors from the occipital and temporal lobes, respectively. As shown in Fig. 5, the time-frequency representations from the occipital sensors, and not the temporal sensors, were locked with changes in visual features. The opposite pattern was observed for the audio feature changes in the stimuli. The results demonstrated that current MEG data could accurately detect stimulus-induced brain activity.

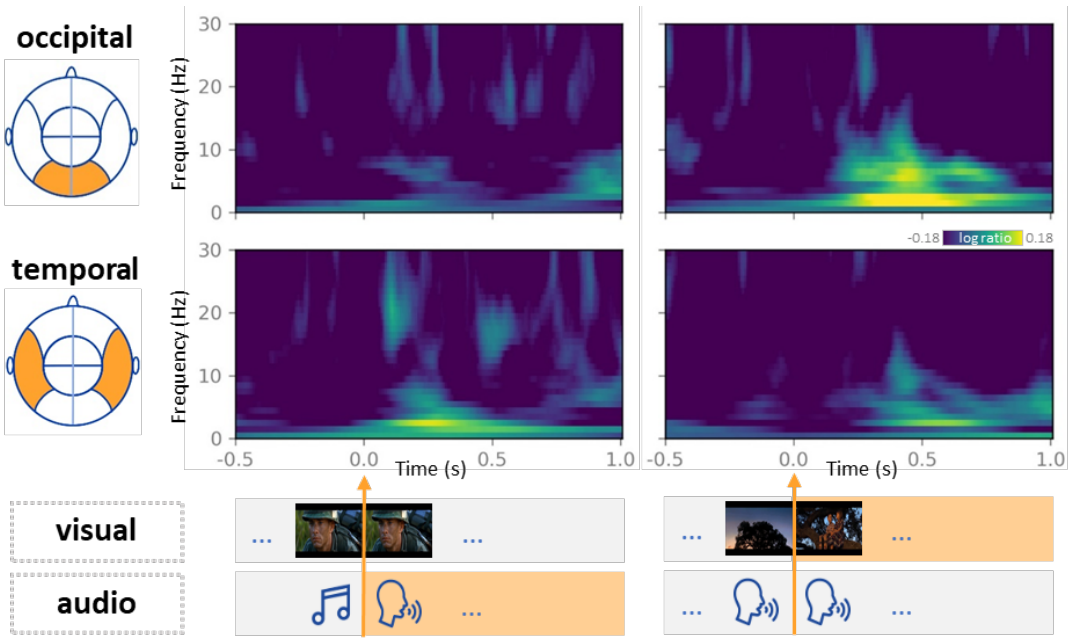


Fig. 5. Time-frequency characterization of the brain activity for stimuli. MEG signals from two movie clips with pronounced changes in either audio or visual features of the stimuli have been examined. The occipital sensors (top) display significant signal changes with a change in the visual features of the stimuli. The temporal sensors (bottom) display significant changes with a change in the audio features.

Homotopic functional connectivity

A basic principle of the brain's functional architecture is that FC between inter-hemispheric homologs (i.e., homotopic regions) is particularly stronger than other interhemispheric (i.e., heterotopic) FC[28,29]. Herein, we tested if MEG data for dynamic movies at the sensor level could reveal strong homotopic FC. First, we calculated the absolute envelope amplitude of the MEG signal from each sensor via the Hilbert transform and down-sampled to 1 Hz. Second, the sensor-level homotopic FC was calculated using Pearson's correlation between the envelope amplitude of each homotopic sensor pair. For comparison, the heterotopic FC was also calculated for each sensor as the average correlation between all heterotopic sensor pairs. Finally, the homotopic and heterotopic FC values were averaged across all runs and all participants. The homotopic sensors expectedly revealed stronger FC than the heterotopic sensor pairs (Fig. 6a). Moreover, the high homotopic FC primarily appeared at the sensors located in the occipital and temporal cortices, thereby indicating strong couplings driven by the movie stimuli (Fig. 6b).

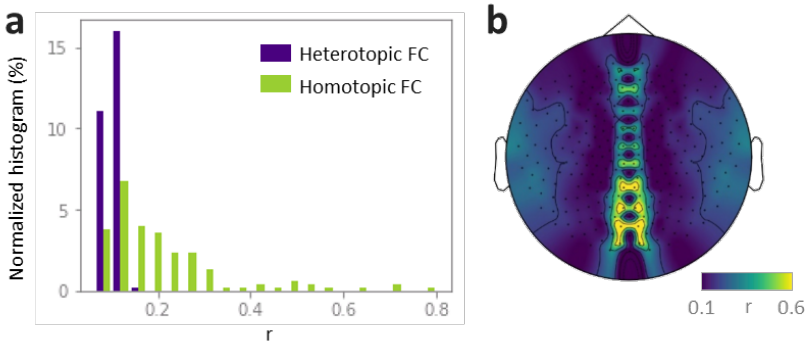


Fig. 6. The sensor-level homotopic functional connectivity (FC) is stronger than the heterotopic FC. (a) The histogram of the sensor-level homotopic FC and heterotopic FC pooled across all runs and participants. **(b)** The topographic map of the sensor-level homotopic FCs averaged across all acquisitions. Identical homotopic FC values are displayed for the corresponding homotopic sensors from the two hemispheres. Sensors in the central axis with no corresponding homotopic sensors were not included in the analysis.

Inter-subject correlation

ISC analysis uses the brain responses of a subject to naturalistic stimuli as a model to predict the brain responses of other subjects[30]. Numerous studies have

demonstrated that visual and auditory cortices display significant ISC while watching audio-visual movies. We validated if a high ISC could be detected in our MEG data. For simplicity, the ISC analysis was conducted at the sensor-level. The MEG recordings captured neural oscillations at different frequency bands[31]. Therefore, the ISC was calculated in five bands (delta: 1-4 Hz, theta: 4-8 Hz, alpha: 8-13 Hz, beta: 13-30 Hz, and gamma: 30-100 Hz). First, the MEG signal was filtered for each band. Second, the absolute envelope amplitude of each band was calculated via the Hilbert transform and down-sampled to 1 Hz. Third, for each frequency band, a leave-one-participant-out ISC was calculated for the left participant as the temporal correlation between the envelope amplitude from the participant and the average of other participants. Finally, the mean ISC was calculated by averaging the ISC across all participants. As shown in Fig. 7, sensors with higher ISC were located near the visual and audio cortices, which reportedly displayed high ISC during movie watching in previous studies[14,32–34]. In particular, the high ISC predominantly occurred in the delta, theta, and alpha bands, consistent with previous studies[14,33,34]. Together, the dataset demonstrated good validity in detecting ISC. In addition, a high ISC was observed in the orbital frontal area in the raw data, likely caused by eye blink or movement considering that high ISC was not observed in the preprocessed data, in which the artifacts have been removed.

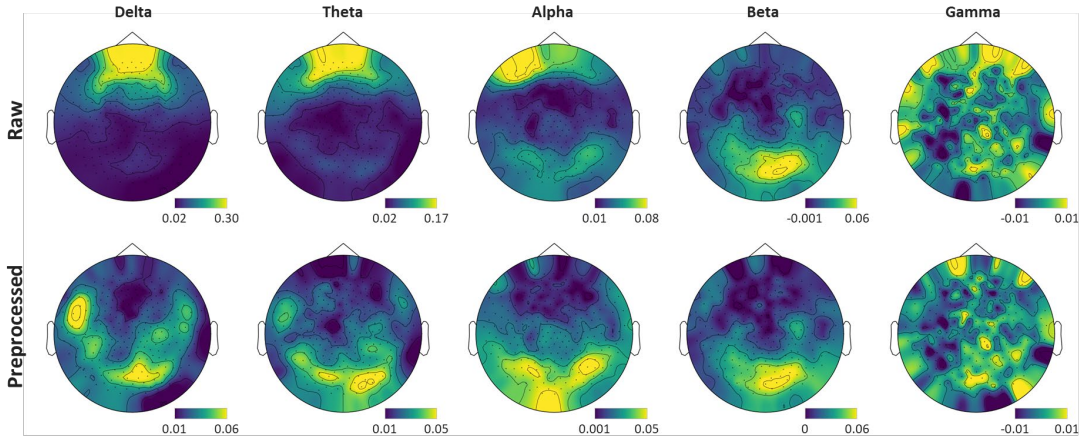


Fig. 7. Topographic maps of ISC in different frequency bands derived from both the raw and preprocessed MEG data. A high ISC occurs near the visual and audio cortices in the delta, theta and alpha bands for both the raw (top) and preprocessed (bottom) MEG data. A high orbital frontal ISC is only observed in the raw data, likely caused by artifacts.

Usage Notes

We presented the first public MEG dataset for a full-length movie. The MEG signals were recorded while the participants watched the 2 h long Chinese-dubbed audio-visual movie "Forrest Gump". The dataset provided a versatile resource for studying information processing in real-life contexts. First, MEG data could be independently used to study the neural dynamics of sensory processing and higher-level cognitive functions under real-life conditions. Second, as a *studyforrest* extension, the dataset could be integrated with publicly available fMRI data from the *studyforrest* project. The fusion of fMRI and MEG may shed new light on the relationship between spatially localized networks observed in fMRI and the MEG-derived temporal dynamics. Moreover, our massive MEG recordings enable the direct training of DNNs with neural activity patterns. In contrast to the DNNs that were usually trained with stimuli without referring to any neural representation, this kind of brain-constrained DNNs would act more like the human brain and generalize well across many tasks[35,36].

Despite the importance of the aforementioned dataset as an extension of the *studyforrest* dataset in studying the spatiotemporal dynamics underlying cognitive processing in real-life contexts, the limitations should be acknowledged. First, the participants in our MEG data did not overlap with that in the *studyforrest* project. Therefore, the fMRI-MEG fusion can be only performed at the group level (i.e., across participants) instead of the individual level (i.e., within participants). This may make it unable to study the individual differences in the coupling between spatially localized networks and temporal dynamics. Second, the dubbed languages used in our dataset and the *studyforrest* project were radically different, thus limiting the application of the data in examining spatiotemporal dynamics of brain activity underlying auditory and language. In addition, caution should be taken with timing differences between the stimuli in the MEG and fMRI data. Considering the lower sensitivity of fMRI signals to the exact timing than MEG signals, we recommend the use of MEG stimuli in fusing fMRI and MEG data.

Code Availability

All custom codes for data preprocessing and technical validation are available at https://github.com/BNUCNL/MEG_Gump. Preprocessing was performed using

MNE-BIDS (https://mne.tools/stable/index.html), MNE (https://mne.tools/stable/install/mne_python.html), fMRIPrep (https://fmriprep.org/en/stable/), pydeface (https://github.com/poldracklab/pydeface), and dcm2niix (https://github.com/rordenlab/dcm2niix).

Author contributions

X.L. conceived, performed the study, and wrote the manuscript. Y.D. performed the study. H.X. contributed to the data collection. Z.Z. conceived, supervised the study and wrote the manuscript.

References

1. Sonkusare, S., Breakspear, M. & Guo, C. Naturalistic Stimuli in Neuroscience: Critically Acclaimed. *Trends in Cognitive Sciences* **23**, 699–714 (2019).
2. Jääskeläinen, I. P., Sams, M., Glerean, E. & Ahveninen, J. Movies and narratives as naturalistic stimuli in neuroimaging. *NeuroImage* **224**, 117445 (2021).
3. Wen, H. *et al.* Neural Encoding and Decoding with Deep Learning for Dynamic Natural Vision. *Cerebral Cortex* **28**, 4136–4160 (2018).
4. Shen, G., Horikawa, T., Majima, K. & Kamitani, Y. Deep image reconstruction from human brain activity. *PLOS Computational Biology* **15**, e1006633 (2019).
5. Güçlü, U. & van Gerven, M. A. J. Increasingly complex representations of natural movies across the dorsal stream are shared between subjects. *NeuroImage* **145**, 329–336 (2017).
6. Cichy, R. M. *et al.* The Algonauts Project 2021 Challenge: How the Human Brain Makes Sense of a World in Motion. *arXiv:2104.13714 [cs, q-bio]* (2021).
7. Hanke, M. *et al.* A high-resolution 7-Tesla fMRI dataset from complex natural stimulation with an audio movie. *Scientific Data* **1**, sdata20143 (2014).

8. Hanke, M. *et al.* A studyforrest extension, simultaneous fMRI and eye gaze recordings during prolonged natural stimulation. *Scientific Data* **3**, 160092 (2016).

9. Liu, X., Zhen, Z., Yang, A., Bai, H. & Liu, J. A manually denoised audio-visual movie watching fMRI dataset for the studyforrest project. *Sci Data* **6**, 295 (2019).

10. Aliko, S., Huang, J., Gheorghiu, F., Meliss, S. & Skipper, J. I. A naturalistic neuroimaging database for understanding the brain using ecological stimuli. *Scientific Data* **7**, 347 (2020).

11. Visconti di Oleggio Castello, M., Chauhan, V., Jiahui, G. & Gobbini, M. I. An fMRI dataset in response to "The Grand Budapest Hotel", a socially-rich, naturalistic movie. *Scientific Data* **7**, 383 (2020).

12. Chang, W.-T. *et al.* Combined MEG and EEG show reliable patterns of electromagnetic brain activity during natural viewing. *Neuroimage* **114**, 49–56 (2015).

13. Betti, V. *et al.* Natural Scenes Viewing Alters the Dynamics of Functional Connectivity in the Human Brain. *Neuron* **79**, 782–797 (2013).

14. Lankinen, K., Saari, J., Hari, R. & Koskinen, M. Intersubject consistency of cortical MEG signals during movie viewing. *NeuroImage* **92**, 217–224 (2014).

15. Lankinen, K. *et al.* Consistency and similarity of MEG- and fMRI-signal time courses during movie viewing. *NeuroImage* **173**, 361–369 (2018).

16. Nunes, A. S. *et al.* Neuromagnetic activation and oscillatory dynamics of stimulus-locked processing during naturalistic viewing. *NeuroImage* **216**, 116414 (2020).

17. Kleiner, M. *et al.* What's New in Psychtoolbox-3? *Perception* **36**, 1–16 (2007).

18. Gramfort, A. MEG and EEG data analysis with MNE-Python. *Front. Neurosci.* **7**, (2013).

19. Esteban, O. *et al.* fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods* **16**, 111–116 (2019).

20. Tustison, N. J. *et al.* N4ITK: Improved N3 Bias Correction. *IEEE Trans. Med. Imaging* **29**, 1310–1320 (2010).

21. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical Surface-Based Analysis. *NeuroImage* 16 (1999).

22. Liu, X., Dai, Y., Xie, H. & Zhen, Z. ForrestGump-MEG. *OpenNeuro* (2021) doi:10.18112/openneuro.ds003633.v1.0.1.

23. Niso, G. *et al.* MEG-BIDS, the brain imaging data structure extended to magnetoencephalography. *Sci Data* **5**, 180110 (2018).

24. Appelhoff, S. *et al.* MNE-BIDS: Organizing electrophysiological data into the BIDS format and facilitating their analysis. *JOSS* **4**, 1896 (2019).

25. Uno, T. *et al.* Dissociated Roles of the Inferior Frontal Gyrus and Superior Temporal Sulcus in Audiovisual Processing: Top-Down and Bottom-Up Mismatch Detection. *PLoS ONE* **10**, e0122580 (2015).

26. Chikara, R. K. & Ko, L.-W. Modulation of the Visual to Auditory Human Inhibitory Brain Network: An EEG Dipole Source Localization Study. *Brain Sciences* **9**, 216 (2019).

27. Ferraro, S. *et al.* Stereotactic electroencephalography in humans reveals multisensory signal in early visual and auditory cortices. *Cortex* **126**, 253–264 (2020).

28. Lowe, M. J., Mock, B. J. & Sorenson, J. A. Functional Connectivity in Single and Multislice Echoplanar Imaging Using Resting-State Fluctuations. *NeuroImage* **7**, 119–132 (1998).

29. Mancuso, L. *et al.* The homotopic connectivity of the functional brain: a meta-analytic approach. *Sci Rep* **9**, 3346 (2019).

30. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *science* **303**, 1634–1640 (2004).

31. Hari, R. & Salmelin, R. Magnetoencephalography: From SQUIDs to neuroscience. *NeuroImage* **61**, 386–396 (2012).

32. Hasson, U. *et al.* Neurocinematics: The neuroscience of film. *Projections* **2**, 1–26 (2008).

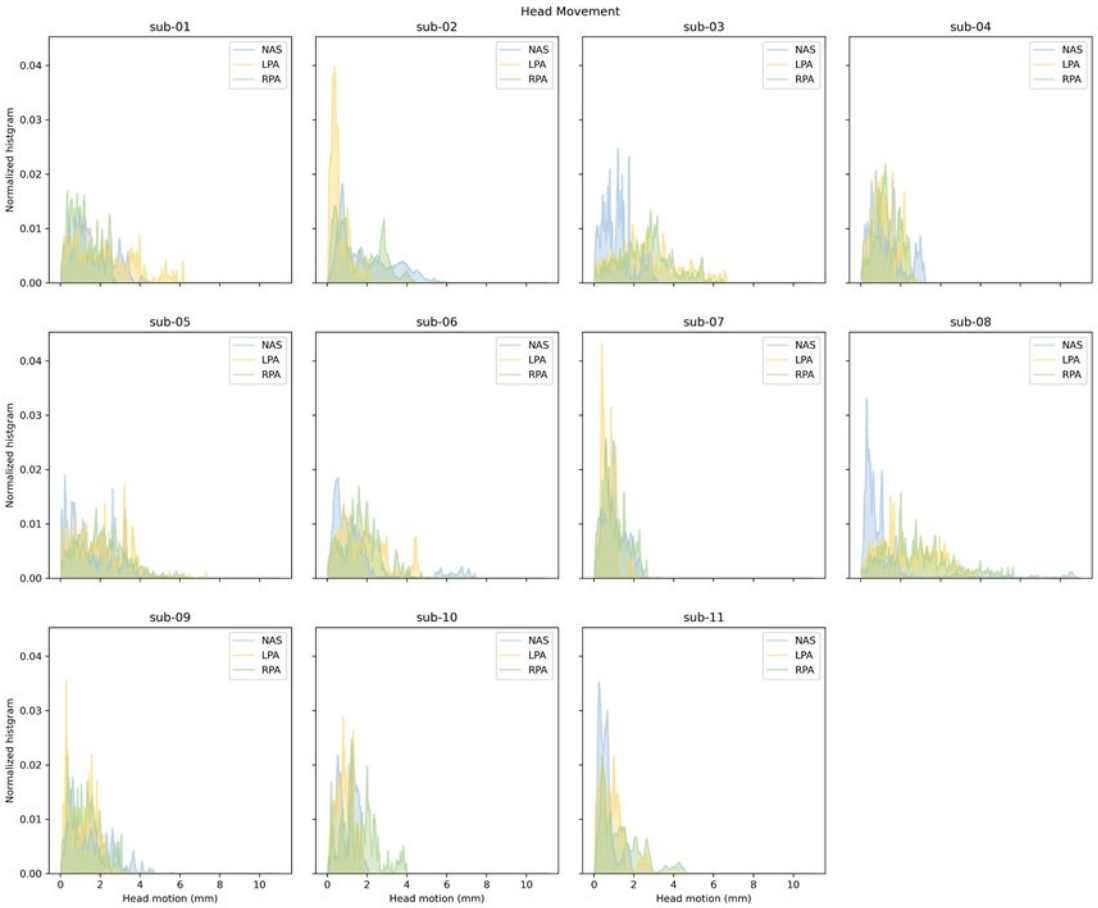
33. Thiede, A., Glerean, E., Kujala, T. & Parkkonen, L. Atypical MEG inter-subject correlation during listening to continuous natural speech in dyslexia. *NeuroImage* **216**, 116799 (2020).

34. Puschmann, S., Regev, M., Baillet, S. & Zatorre, R. J. MEG Intersubject Phase Locking of Stimulus-Driven Activity during Naturalistic Speech Listening Correlates with Musical Training. *J. Neurosci.* **41**, 2713–2722 (2021).

35. Fong, R. C., Scheirer, W. J. & Cox, D. D. Using human brain activity to guide machine learning. *Sci Rep* **8**, 5397 (2018).

36. Spampinato, C. *et al.* Deep Learning Human Mind for Automated Visual Classification. *arXiv:1609.00344 [cs]* (2019).

Supplementary materials



Supplementary Figure 1. Head motion magnitude from each individual participant. The density histogram of motion magnitude calculated for three fiducials (NAS, LPA, RPA) were plotted for all samples, across all runs for each participant.

Data collection

Consent and questionnaire

MEG Session
Audiovisual movie watching
8 runs

MRI Session
T1w imaging

Input Processing Output

Data preprocessing

raw MEG data

Bad channel detection
1 Hz High-pass filtering
ICA artifacts removal

Preprocessed MEG data and derivatives

MNE toolbox

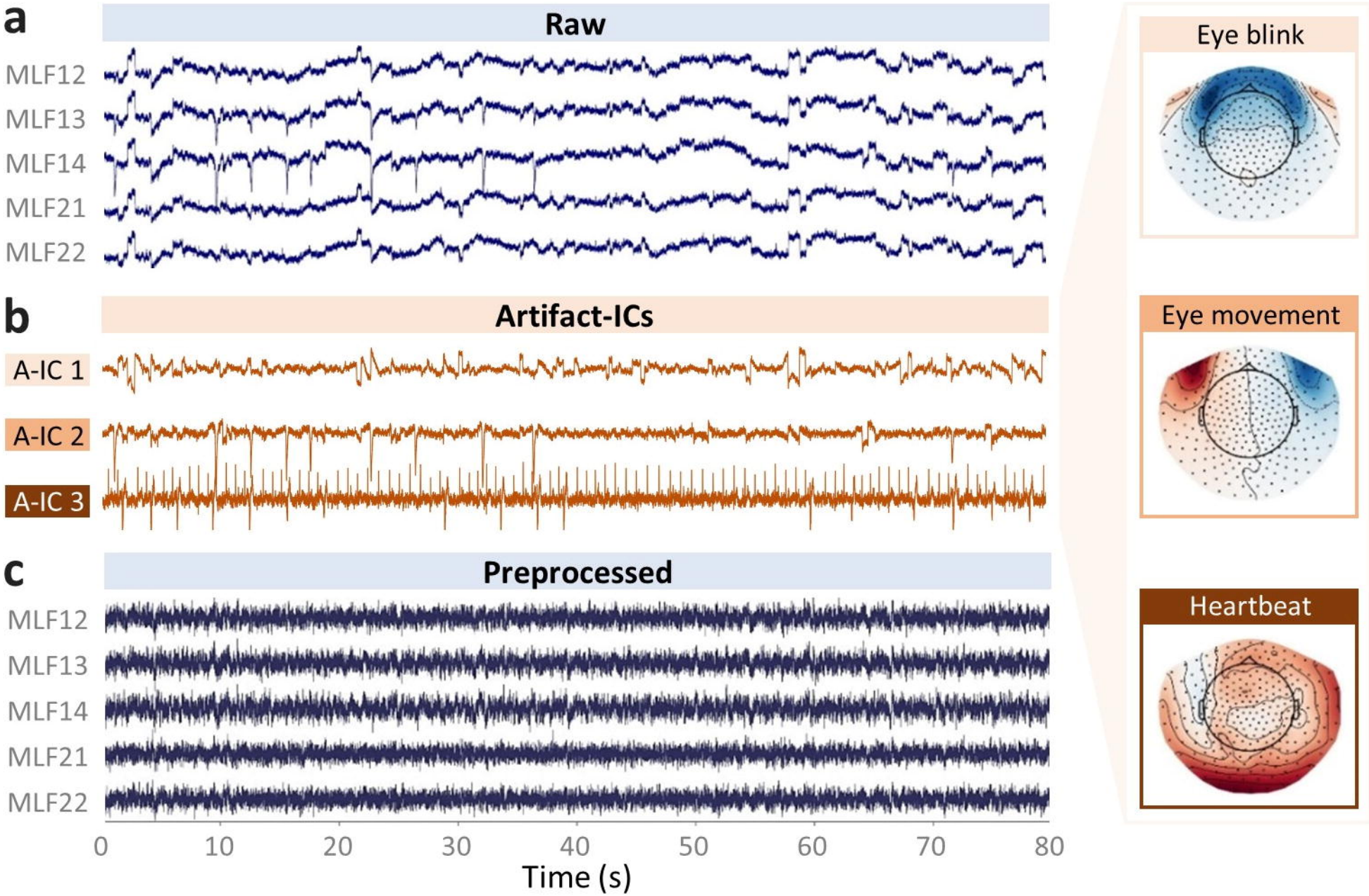
coregistration

raw T1w MRI data

Volume preprocessing
Surface reconstruction and preprocessing

Preprocessed volume and surface data in native and standard space

fMRIPrep


a
Raw
MLF12
MLF13
MLF14
MLF21
MLF22
b
Artifact-ICs
A-IC 1
A-IC 2
A-IC 3
c
Preprocessed
MLF12
MLF13
MLF14
MLF21
MLF22
0
10
20
30
40
50
60
70
80
Time (s)
Eye blink
Eye movement
Heartbeat

a

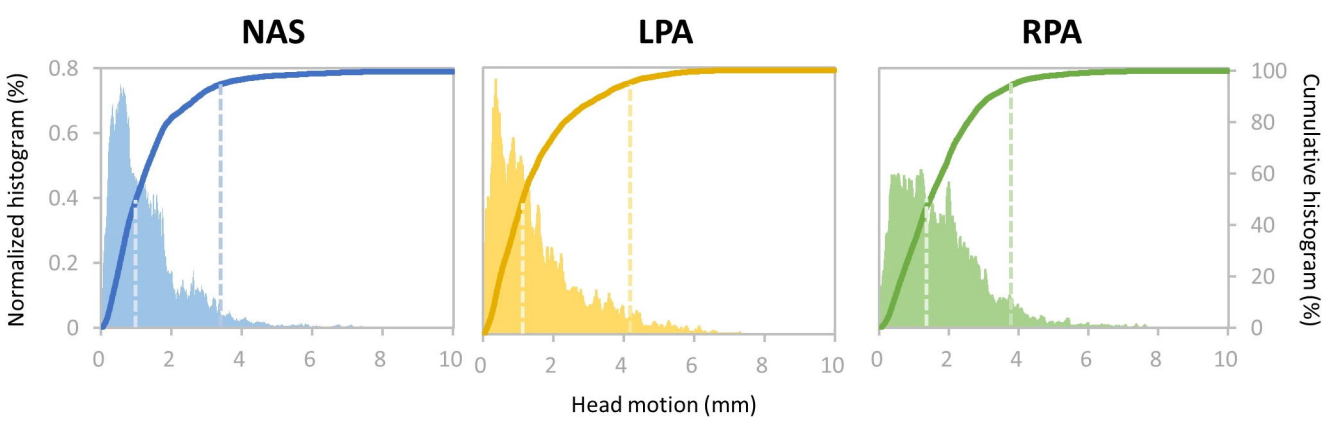
```
./
├── dataset_description.json
├── participants.json
├── participants.tsv
├── README
├── code/
├── sub-01/
├── sub-xx/
├── sub-emptyroom/
└── derivatives/
```

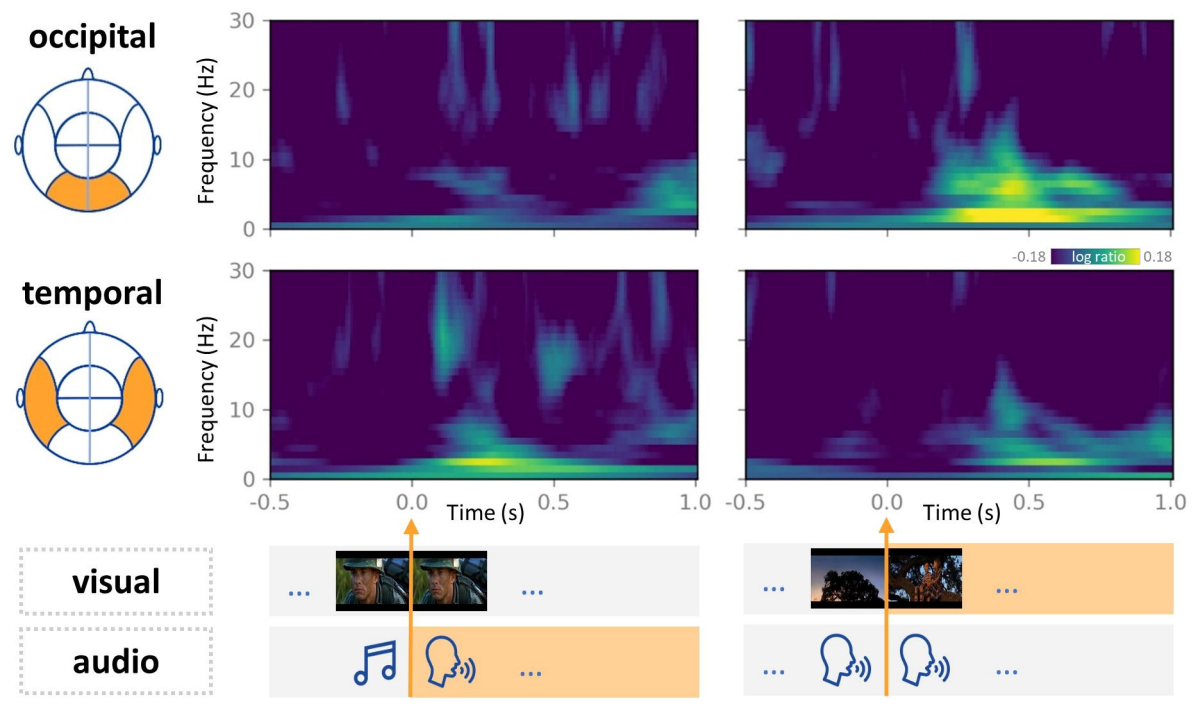
b

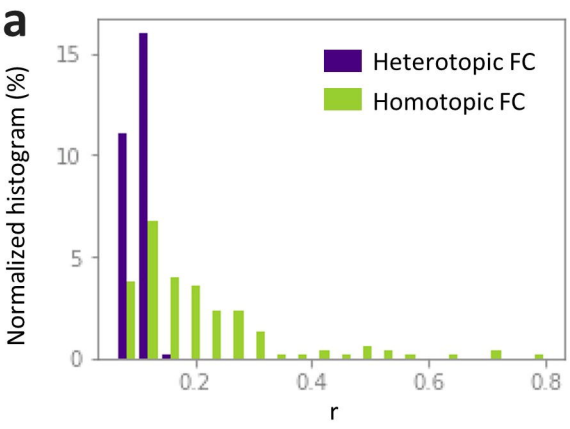
```
./
├── ...
├── sub-01/
│   └── ses-movie/
│       ├── anat/
│       │   ├── sub-01_ses-movie_T1w.json
│       │   └── sub-01_ses-movie_T1w.nii.gz
│       ├── meg/
│       │   ├── sub-01_ses-movie_coordsystem.json
│       │   ├── sub-01_ses-movie_task-movie_run-01_channels.tsv
│       │   ├── sub-01_ses-movie_task-movie_run-01_events.tsv
│       │   ├── sub-01_ses-movie_task-movie_run-01_meg.ds
│       │   ├── sub-01_ses-movie_task-movie_run-01_meg.json
│       │   └── ...
│       └── sub-01_ses-movie_scans.tsv
├── sub-xx/
└── sub-emptyroom/
    ├── ses-20190429/
    │   ├── meg/
    │   │   ├── sub-emptyroom_ses-20190429_task-noise_channels.tsv
    │   │   ├── sub-emptyroom_ses-20190429_task-noise_meg.ds
    │   │   ├── sub-emptyroom_ses-20190429_task-noise_meg.json
    │   │   └── ...
    │   └── sub-emptyroom_ses-20190429_scans.tsv
    └── ses-xxxxxxxx/
```

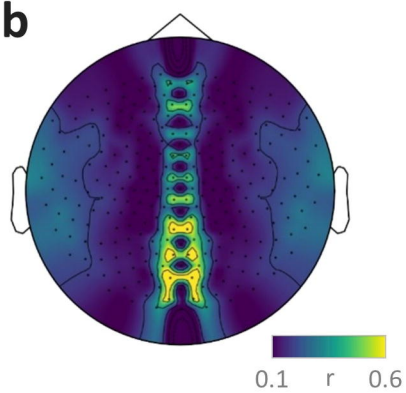
c

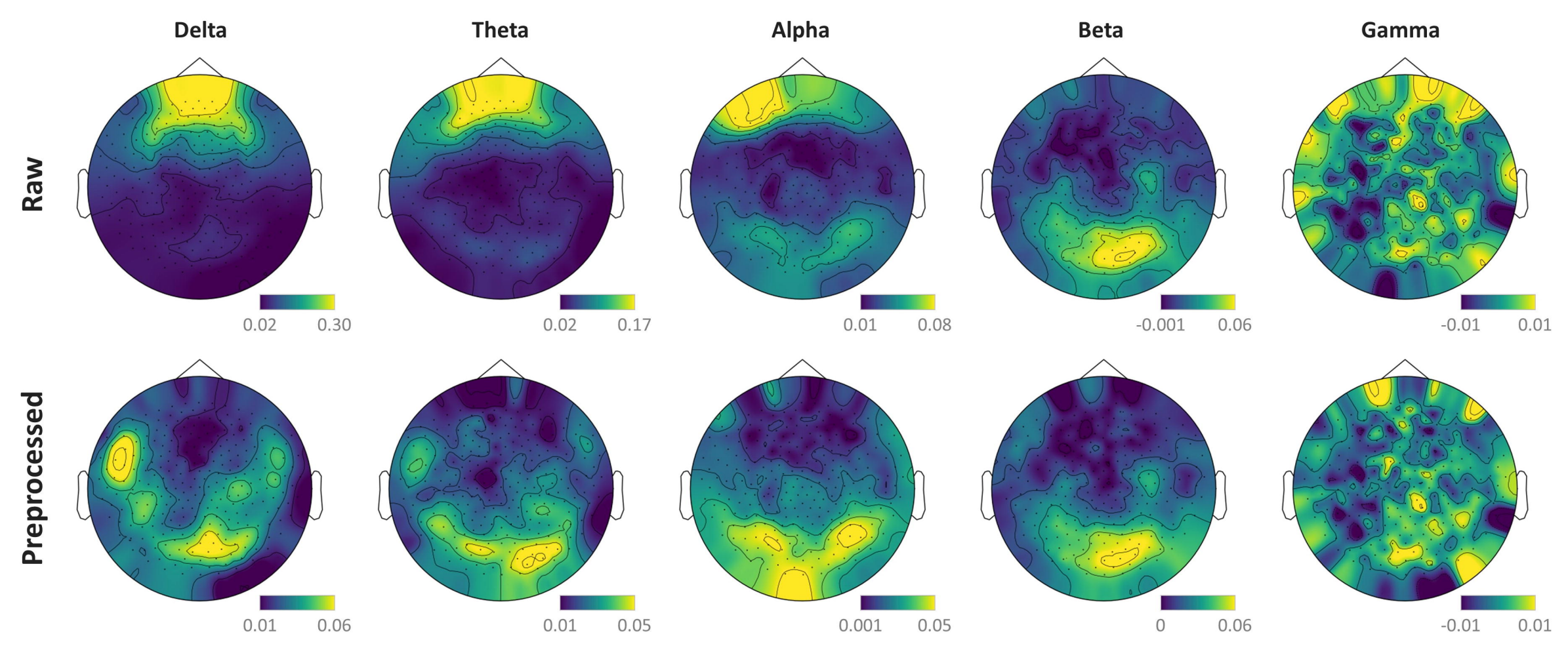
```
./
├── ...
└── derivatives/
    └── preproc_meg-mne_mri-fmriprep/
        ├── dataset_description.json
        ├── desc-aparcaseg_dseg.tsv
        ├── desc-aseg_dseg.tsv
        ├── participants.json
        ├── participants.tsv
        ├── sourcedata/
        │   └── freesurfer/
        │       ├── sub-01/
        │       │   ├── bem/
        │       │   │   ├── brain.surf
        │       │   │   ├── inner_skull.surf
        │       │   │   ├── outer_skin.surf
        │       │   │   ├── outer_skull.surf
        │       │   │   ├── sub-01-fiducials.fif
        │       │   │   ├── sub-01-head-dense.fif
        │       │   │   ├── sub-01-head.fif
        │       │   │   ├── sub-01-head-medium.fif
        │       │   │   ├── sub-01-head-sparse.fif
        │       │   │   └── ...
        │       │   ├── mri/
        │       │   ├── surf/
        │       │   └── ...
        │       └── sub-xx/
        ├── sub-01/
        │   └── ses-movie/
        │       ├── anat/
        │       │   ├── sub-01_ses-movie_desc-preproc_T1w.nii.gz
        │       │   ├── sub-01_ses-movie_hemi-L_inflated.surf.gii
        │       │   ├── sub-01_ses-movie_hemi-L_midthickness.surf.gii
        │       │   ├── sub-01_ses-movie_hemi-L_pial.surf.gii
        │       │   ├── sub-01_ses-movie_hemi-L_smoothwm.surf.gii
        │       │   ├── sub-01_ses-movie_hemi-R_inflated.surf.gii
        │       │   ├── sub-01_ses-movie_hemi-R_midthickness.surf.gii
        │       │   ├── sub-01_ses-movie_hemi-R_pial.surf.gii
        │       │   ├── sub-01_ses-movie_hemi-R_smoothwm.surf.gii
        │       │   ├── sub-01_ses-movie_space-MNI152NLin2009cAsym_desc-preproc_T1w.nii.gz
        │       │   ├── sub-01_ses-movie_space-MNI152NLin6Asym_desc-preproc_T1w.nii.gz
        │       │   └── ...
        │       ├── meg/
        │       │   ├── sub-01_ses-movie_coordsystem.json
        │       │   ├── sub-01_ses-movie_task-movie_run-01_channels.tsv
        │       │   ├── sub-01_ses-movie_task-movie_run-01_decomposition.tsv
        │       │   ├── sub-01_ses-movie_task-movie_run-01_events.tsv
        │       │   ├── sub-01_ses-movie_task-movie_run-01_ica.fif.gz
        │       │   ├── sub-01_ses-movie_task-movie_run-01_meg.fif
        │       │   ├── sub-01_ses-movie_task-movie_run-01_meg.json
        │       │   ├── ...
        │       │   └── sub-01_ses-movie_task-movie_trans.fif
        │       └── sub-01_ses-movie_scans.tsv
        └── sub-xx/
```


NAS
LPA
RPA
Normalized histogram (%)
Cumulative histogram (%)
Head motion (mm)
0.8
0.6
0.4
0.2
0
100
80
60
40
20
0
0
2
4
6
8
10


occipital
temporal
Frequency (Hz)
30
20
10
0
-0.18
log ratio
0.18
-0.5
0.0
0.5
1.0
Time (s)
visual
audio


a
Normalized histogram (%)
15
10
5
0
0.2
0.4
0.6
0.8
r
Heterotopic FC
Homotopic FC


b
0.1
r
0.6


Delta
Theta
Alpha
Beta
Gamma
Raw
Preprocessed
0.02
0.30
0.02
0.17
0.01
0.08
-0.001
0.06
-0.01
0.01
0.01
0.06
0.01
0.05
0.001
0.05
0
0.06
-0.01
0.01