Distributed representations of prediction error signals across the cortical hierarchy are synergistic

Frank Gelens[a,b,k], Juho Äijälä[b,k], Louis Roberts[b,h], Misako Komatsu[c], Cem Uran[d,e], Michael A. Jensen[f], Kai J. Miller[f], Robin A.A. Ince[g], Max Garagnani[h,i], Martin Vinck[d,e,l], Andres Canales-Johnson[b,j,l]

[a]*Department of Psychology, University of Amsterdam, Nieuwe Achtergracht 129-B, 1018 WT, Amsterdam, The Netherlands*
[b]*Department of Psychology, University of Cambridge, CB2 3EB Cambridge, United Kingdom*
[c]*Laboratory for Adaptive Intelligence, RIKEN Brain Science Institute, Saitama 351-0198, Japan*
[d]*Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation with Max Planck Society, 60528 Frankfurt am Main, Germany*
[e]*Donders Centre for Neuroscience, Department of Neuroinformatics, Radboud University Nijmegen, 6525 Nijmegen, Netherlands*
[f]*Department of Neurosurgery, Mayo Clinic, Rochester, Minnesota 55905*
[g]*Institute of Neuroscience and Psychology, University of Glasgow, Scotland G12 8QB, United Kingdom*
[h]*Department of Computing, Goldsmiths, University of London, SE14 6NW London, United Kingdom*
[i]*Brain Language Lab, Freie Universität Berlin, 14195 Berlin, Germany*
[j]*Neuropsychology and Cognitive Neurosciences Research Center, Faculty of Health Sciences, Universidad Católica del Maule, 3460000 Talca, Chile*
[k]*These authors contributed equally*
[l]*Correspondence to: Andres Canales-Johnson (afc37@cam.ac.uk), Martin Vinck (martin.vinck@esi-frankfurt.de)*

Abstract

An important question concerning inter-areal communication in the cortex is whether these interactions are synergistic, i.e. brain signals can either share common information (redundancy) or they can encode complementary information that is only available when both signals are considered together (synergy). Here, we dissociated cortical interactions sharing common information from those encoding complementary information during prediction error processing. To this end, we computed co-information, an information-theoretical measure that distinguishes redundant from synergistic information among brain signals. We analyzed auditory and frontal electrocorticography (ECoG) signals in five common awake marmosets performing two distinct auditory oddball tasks and investigated to what extent event-related potentials (ERP) and broadband (BB) dynamics encoded redundant and synergistic information during auditory prediction error processing. In both tasks, we observed multiple patterns of synergy across the entire cortical hierarchy with distinct dynamics. The information conveyed by ERPs and BB signals was highly synergistic even at lower stages of the hierarchy in the auditory cortex, as well as between auditory and frontal regions. Using a brain-constrained neural network, we simulated the spatio-temporal patterns of synergy and redundancy observed in the experimental results and further demonstrated that the emergence of synergy between auditory and frontal regions requires the presence of strong, long-distance, feedback and feedforward connections. These results indicate that the distributed representations of prediction error signals across the cortical hierarchy can be highly synergistic.

INTRODUCTION

The traditional modular view of brain function is increasingly challenged by the finding that information about external stimuli and internal variables is distributed across brain areas (de Schotten and Forkel, 2022; Urai et al., 2022; Shenoy and Kao, 2021; Breakspear, 2017; Panzeri et al., 2022). When information in a complex system is carried by multiple nodes, this could imply that there is a large degree of redundancy in the information carried by the different nodes. That is, the whole is actually less than the sum of the parts. An alternative possibility, however, is that information is carried in a synergistic manner, i.e. the different nodes might carry extra information about task variables when they are combined. This can occur when the relationship between the nodes encodes the stimulus in a way that is not apparent when observing each node's activity alone – in other words, the whole is greater than the sum of the parts (Luppi et al., 2022).

Both recent large-scale spiking and electrocorticographic (ECoG) recordings support the notion that information about task variables is widely distributed rather than highly localized (Urai et al., 2022; Steinmetz et al., 2019; Parras et al., 2017; Saleem et al., 2018; Voitov and Mrsic-Flogel, 2022). For example, in the visual domain, widespread neuronal patterns across nearly every brain region are non-selectively activated before movement onset during a visual choice task (Steinmetz et al., 2019). Similarly, distributed and reciprocally interconnected areas of the cortex maintain high-dimensional representations of working memory (Voitov and Mrsic-Flogel, 2022). In the case of multisensory integration, sound-evoked activity and its associated motor correlate can be dissociated from spiking activity in the primary visual cortex (V1) (Lohuis et al., 2022; Bimbard et al., 2023). A last example, and the one used in the current study, is the case of communication of prediction error

(PE) signals. Hierarchical predictive coding theory has been proposed as a general mechanism of processing in the brain (Rao and Ballard, 1999). The communication of prediction error (PE) signals using spikes and local field potentials (LFPs) recorded from subcortical and cortical regions reveal a large-scale hierarchy of PE potentials (Parras et al., 2017).

A major question is whether such distributed signals exhibit a high degree of redundancy (i.e. shared information) or a high degree of synergy (i.e. extra information) about their corresponding task variables. Electrophysiological studies have shown that synergy and redundancy have functional relevance (Nigam et al., 2019; Ince et al., 2017; Park et al., 2018; Ince et al., 2016; Giordano et al., 2017; Luppi et al., 2022; Varley et al., 2023). For instance, laminar recordings in V1 suggest that synergistic interactions can efficiently decode visual stimuli better than redundant interactions, even in the presence of noise and overlapping receptive fields (Nigam et al., 2019). In contrast, the information processing of olfactory stimuli exhibits higher levels of redundant information across olfactory regions (Olivares et al., 2022). Here we investigate this question by using co-Information (co-I), an information theoretical metric capable of decomposing neural signals into what is informationally redundant and what is informationally synergistic between stimuli (McGill, 1954; Ince et al., 2017). Redundant information quantifies the shared information between signals, suggesting a common processing of the stimuli. Synergistic information quantifies something different: whether there is extra information only available when signals are combined, indicating that the information about the variable is in the actual relationship between the signals. Using ECoG recordings, we investigated synergistic and redundant interactions in five common marmosets performing two types of auditory tasks. This allowed us to determine the processing of communication of prediction error information across the brain during a range of auditory deviancy effects. Finally, we applied the same oddball stimulation task used in the experiments to a brain-constrained neurocomputational model of the relevant cortical areas. We computed synergy and redundancy in the simulated responses while manipulating the network's connectivity structure to unravel a potential/candidate mechanism responsible for generating the synergistic interactions observed *in vivo*.

RESULTS

Mutual Information reveals prediction error effects within cortical areas

To characterize the distribution of PE across multiple cortical areas, we quantified PE in each electrode of the five marmosets by contrasting deviant and standard tones (Figure 2). For each electrode, we computed Mutual Information (MI) to quantify the relationship between tone category (standard vs deviant) with their corresponding ECoG signal across trials. Within the framework of information theory, MI is a statistical quantity that measures the strength of the dependence (linear or non-linear) between two random variables. It can be also seen as the effect size, quantified in bits, for a statistical test of independence (Ince et al., 2017). Thus, for each electrode and time point, we considered ECoG signals corresponding to standard and deviant trials and utilized MI to quantify the effect size of their difference.

We have recently proposed that a suitable candidate for broadcasting unpredicted information across the cortex is the transient, aperiodic activity reflected at the level of the evoked-related potentials (ERP) and broadband power (Vinck et al., 2023). A well-studied ERP marker of auditory PE is the mismatch negativity (MMN), an ERP that peaks around 150–250 ms after the onset of an infrequent acoustic stimulus (Parras et al., 2017; Blenkmann et al., 2019; Komatsu et al., 2015; Canales-Johnson et al., 2021). A second neural marker of auditory PE is the broadband response (BB), an increase in spectral power spanning a wide range of frequencies usually above 100 Hz (Canales-Johnson et al., 2021; Jiang et al., 2022). Whereas ERPs reflect a mixture of local potentials and volume conducted potentials from distant sites, BB is an electrophysiological marker of underlying averaged spiking activity generated by the thousands of neurons that are in the immediate vicinity of the recording electrodes (Miller, 2019; Lachaux et al., 2012). MI was computed separately for the two neural markers of prediction error (i.e. ERP and BB signals).

Electrodes showing significant differences in MI over time (see METHODS) are depicted in Figure 2. In the Roving oddball task, ERP signals showed PE effects across multiple cortical regions not necessarily restricted to canonical auditory areas (Figure 2B). In the case of the BB signal, MI analyses revealed PE effects located predominantly in the auditory cortex of the three marmosets, as well as in a few electrodes located in the frontal cortex of marmoset Kr and Go (Figure 2A). These results agree with previous studies in different sensory modalities (Miller, 2019) showing that broadband responses are spatially localized. In the case of the Local/Global Task, although the dataset for marmoset Nr and Ji contained ECoG recording only from the temporal and frontal cortices, the overall PE effects in the ERP signals were observed in a higher number of electrodes than in the BB signal (Figure 2 and Figure S9)

Co-Information reveals redundant and synergistic cortical interactions

To investigate how auditory PE signals are integrated within and between the cortical hierarchy, we quantified redundant and synergistic cortical interactions using an information-theoretic metric known as co-Information (co-I) (Ince et al., 2017). Co-I quantifies the type of information that interacting signals encode about a stimuli variable: positive co-I indicates redundant interactions between signals; and negative co-I accounts for synergistic interactions (Figure 1D). Redundancy implies that the signals convey the same information about PE, indicating a shared encoding of PE information across time or space from trial to trial. On the other hand, synergy implies that signals from different time points or areas convey extra information about PE only when considered together, indicating that the relationship itself contains information about PE that is not available from either of the signals alone (Figure 1D).

To quantify the dynamics of redundancy and synergy temporally and spatially, we computed the co-I within and between

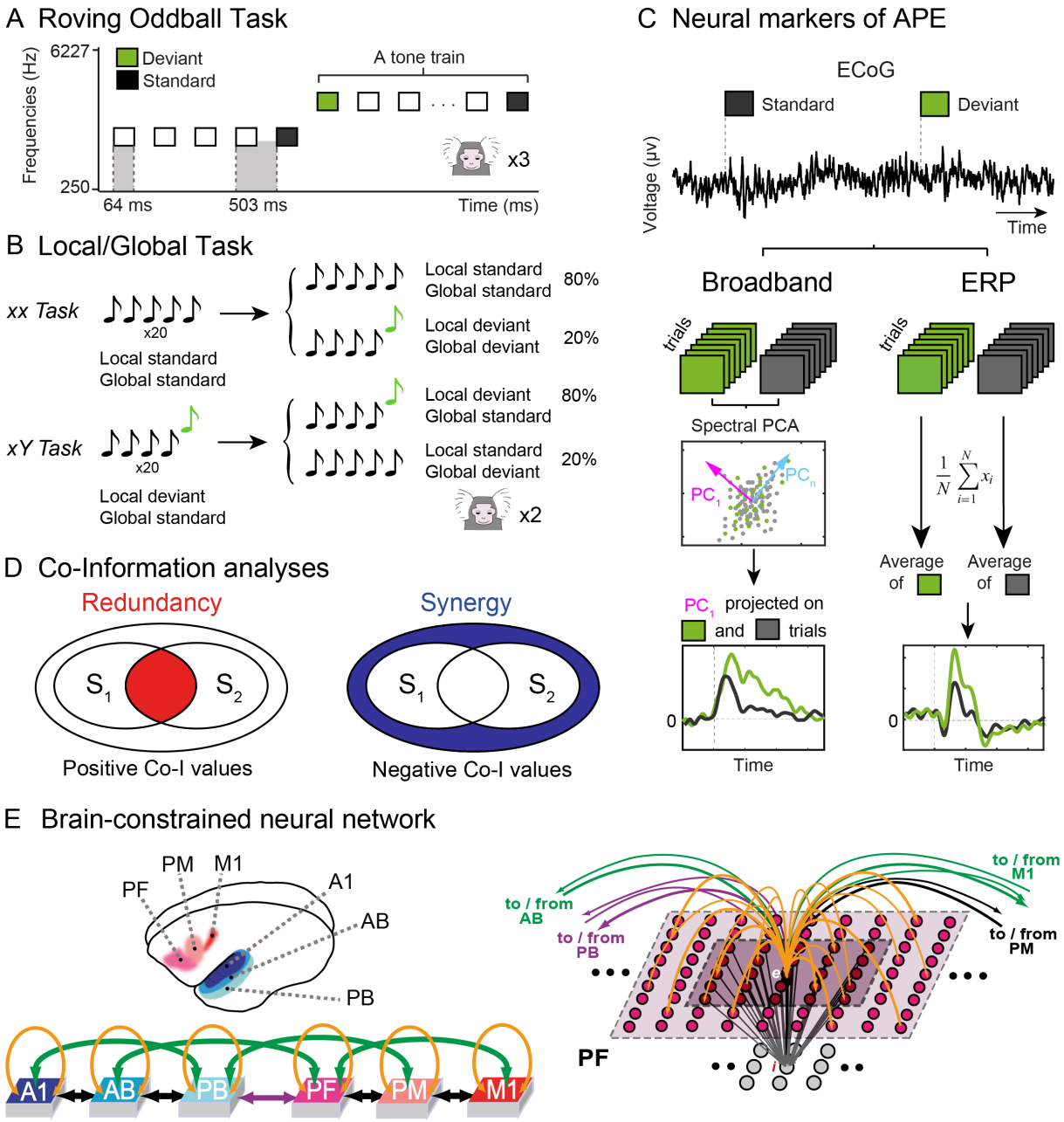

Figure 1: Experimental design, information-theory analyses, and modelling. **(A)** Using a Roving oddball Task, 20 different single tones were presented in the trains of 3, 5, or 11 identical stimuli. Any two subsequent trains consisted of different tones. This way, while the adjacent standard (depicted in black) and deviant (depicted in green) tones deviated in frequency due to the transition between the trains, the two expectancy conditions were physically matched, as the first and the last tones of the same train were treated as deviant and standard tones in the analysis of the adjacent stimuli pairs. This task was performed by 3 marmosets (Fr, Kr, and Go). **(B)** Local/Global Task. On each trial, five tones of 50-ms-duration each were presented with a fixed stimulus onset asynchrony of 150 ms between sounds. The first 4 tones were identical, either low-pitched (tone A) or high-pitched (tone B), but the fifth tone could be either the same (AAAAA or BBBBB, jointly denoted by xx) or different (AAAAB or BBBBA, jointly denoted by xY). Each block started with 20 frequent series of sounds to establish global regularity before delivering the first infrequent global deviant stimulus. This task was performed by 2 different marmosets (Ji and Nr). **(C)** Neural markers of auditory prediction error. Deviant (green) and standard (black) epochs are used to compute the broadband and ERP responses. Broadband is computed by extracting by reconstructing the time series of standard and deviants with the first spectral principal component (SPCA) of the ECoG signal; ERPs are computed by averaging the raw voltage values for standard and deviant trials (see Methods). **(D)** Schematic representation of redundancy and synergy analyses computed using co-Information. Each inner oval (A1 and A2) represents the mutual information between the corresponding ECoG signals and the stimuli category (standard or deviant). The overlap between A1 and A2 represents the redundant information about the stimuli (red; left panel). The outer circle around A1 and A2 represents the synergistic information about the stimuli (blue; right panel). **(E)** Brain areas modelled, network architecture, and its connectivity. *Top left*: Cortical areas modelled. Three cortices in the left temporal lobe (primary auditory: A1, auditory belt: AB, and parabelt: PB) involved in auditory processing, and three in the frontal lobe (prefrontal: PF; premotor: PM; primary motor: M1) directly linked to them. *Bottom left*: Network architecture. All the (sparse and random) connections are based on marmoset neuroanatomy (see Methods). *Right*: Schematic of links to/from a single excitatory cell '**e**'. Each model area consists of two layers of excitatory (upper) and inhibitory (lower) graded-response leaky integrator cells with neuronal fatigue. Dense links between these layers (grey arrows) implement mutual inhibition between **e** and its neighbors. Panel E, right and bottom left is adapted from Garagnani and Pulvermüller (2013), their Figure 1E.

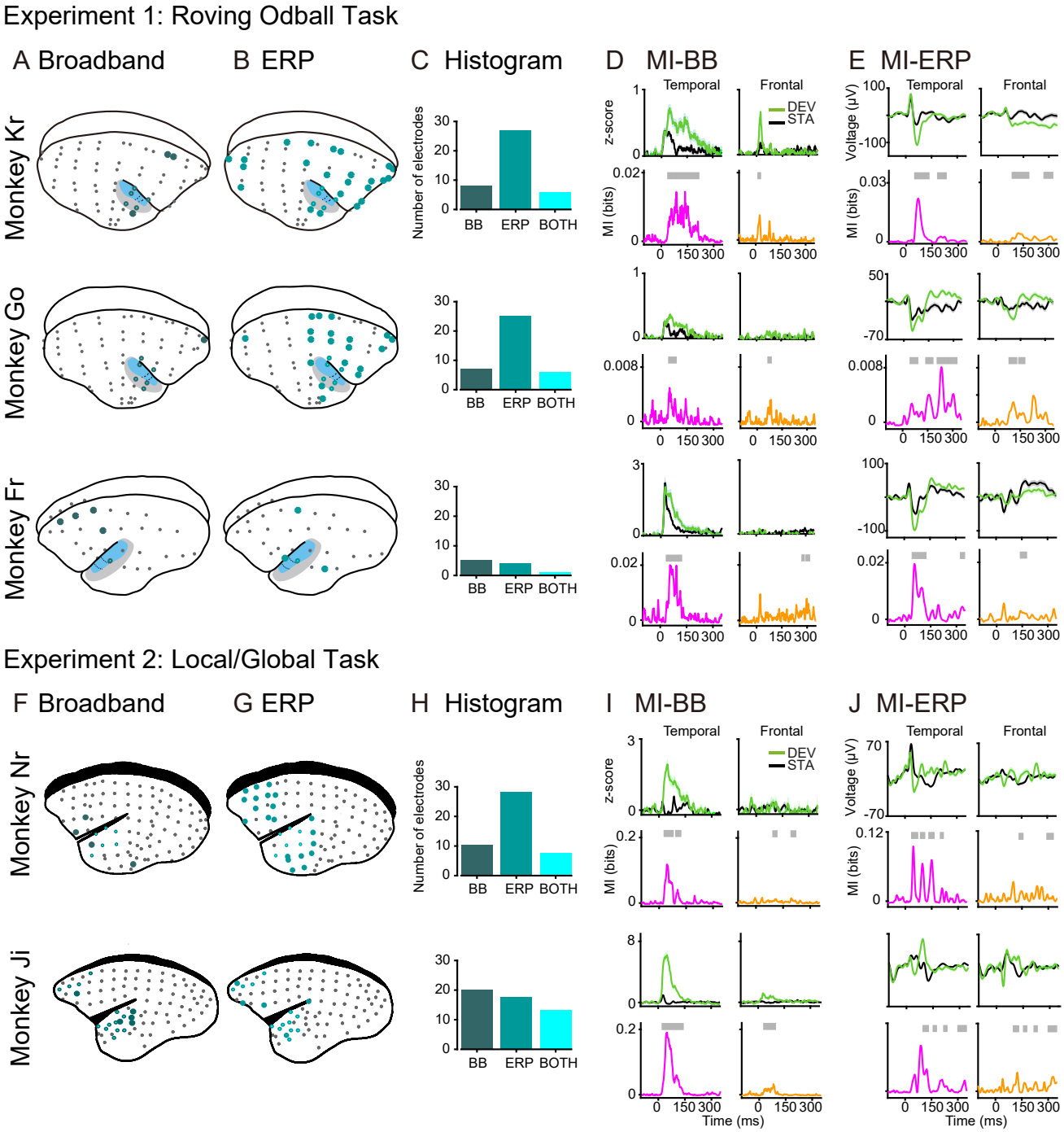

Figure 2: Broadband and ERP markers of PE across the monkey brain. Electrode locations for marmoset Kr (64 electrodes), Go (64 electrodes), and Fr (32 electrodes) in Experiment 1; and Nr (96 electrodes in EcoG-array, 39 used for analyses) and Ji (96 electrodes in EcoG-array, 27 used for analyses) in Experiment 2. Electrodes showing significant PE effect after computing MI between standard and deviant trials for the **(A, F)** Broadband (dark green circles) and **(B, G)** ERP (light green circles) markers of auditory prediction error. Electrodes showing significant MI for both markers are depicted in cyan. **(C, H)** Histogram of electrodes showing significant MI between tones for BB (left), ERP (middle), and both markers (right) for each animal. **(D, I)** Electrodes with the highest MI in the temporal and frontal cortex showing the BB signal for deviant and standard tones. Deviant tone (green) and standard tone (black), and the corresponding MI values in bits (effect size of the difference) for the temporal (pink trace) and frontal (orange trace) electrodes. Significant time points after a permutation test are shown as grey bars over the MI plots. **(E, J)** Electrodes with the highest MI in the temporal and frontal cortex showing the ERP signal for deviant and standard tones. Error bars represent standard error of the mean (S.E.M) across trials.

cortical areas (see METHODS). We analyzed ERP and BB markers of PE separately, focusing our contrasts on the electrodes that showed significant MI effects in the analyses described in Figure 2C.

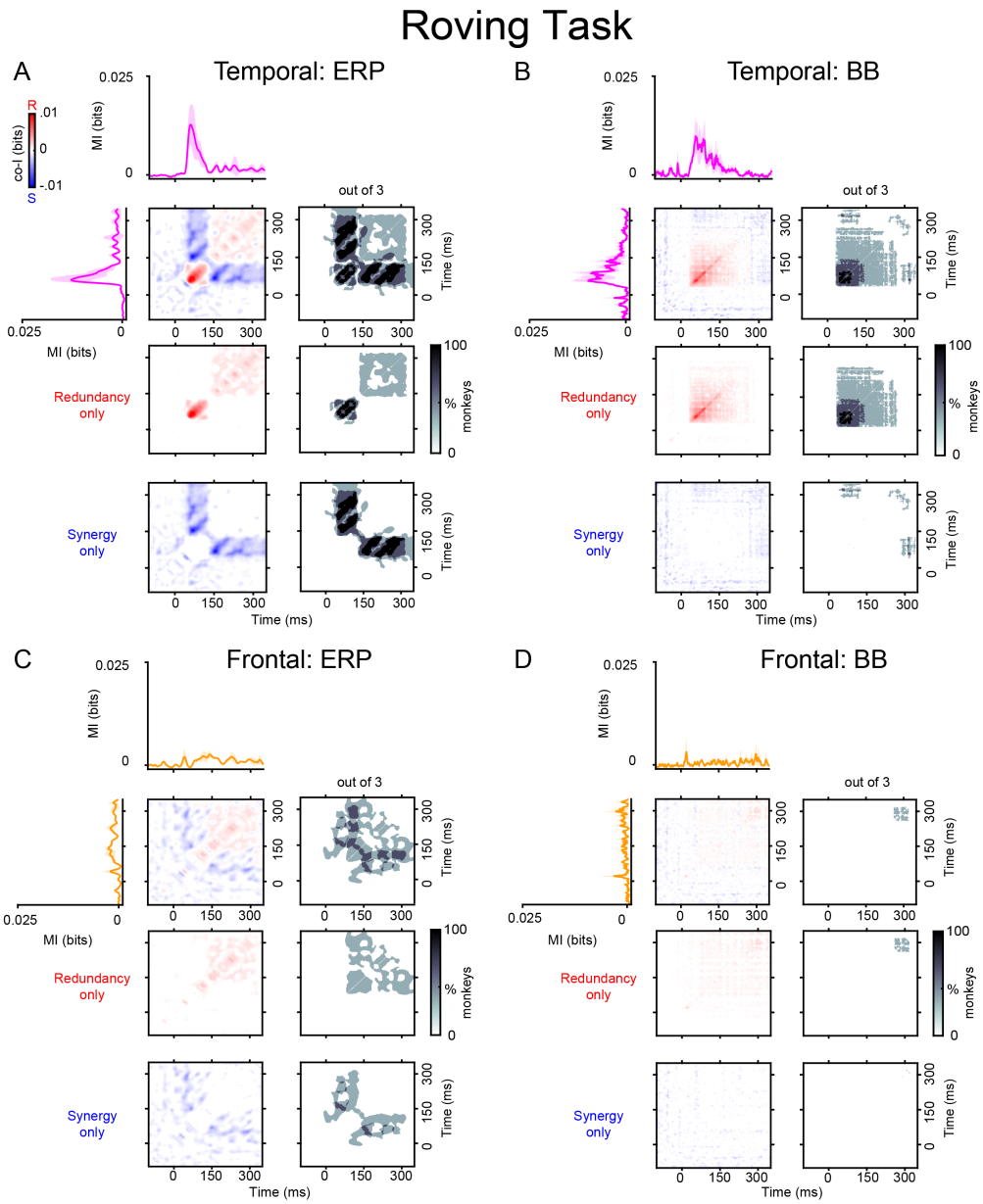

Figure 3: Temporal synergy and redundancy within ERP and BB signals in the auditory and frontal electrodes with the highest MI for the Roving Task (Experiment 1). Co-information revealed synergistic and redundant temporal patterns within ERP (Panel **A**) and BB (Panel **B**) signals in the auditory cortex, and within ERP (**C**) and BB (**D**) signals in the frontal cortex. MI (solid traces) between standard and deviant trials for auditory (pink color) and frontal (orange color) electrodes averaged across the three monkeys. Error bars represent standard error of the mean (S.E.M) across electrodes. Temporal co-I was computed within the corresponding signal (ERP, BB) across time points between -100 to 350 ms after tone presentation. The average of the corresponding electrodes across monkeys is shown for the complete co-I chart (red and blue plots); for positive co-I values (redundancy only; red panel); and for negative co-I values (synergy only; blue plot). The grey-scale plots show the proportion of monkeys showing significant co-I differences in the single electrodes analysis depicted in Figure S1.

Experiment 1: Roving Oddball Task

Temporal synergy and redundancy

The finding that multiple recording sites encode information about PE raises the question of whether these signals convey the same or complementary PE information over time within a cortical region. Thus, we first characterized synergistic and redundant temporal interactions within ERP and BB signals. In the Roving Oddball Task, co-I analyses revealed widespread temporal clusters of synergistic information (in blue) and redundant information (in red) across the three monkeys in the auditory cortex (Figure 3A,B), and frontal cortex (Figure 3C,D). The ERP signal in the auditory (Figure 3A) and frontal (Figure 3C) electrodes showed characteristic off-diagonal synergistic patterns, resulting from the interaction between early and late time points within the same ERP signal (e.g. Figure 3A,C; grey clusters between ~140-300 ms after tone presentation), and revealed by the single electrode contrast depicted in Figure S1.

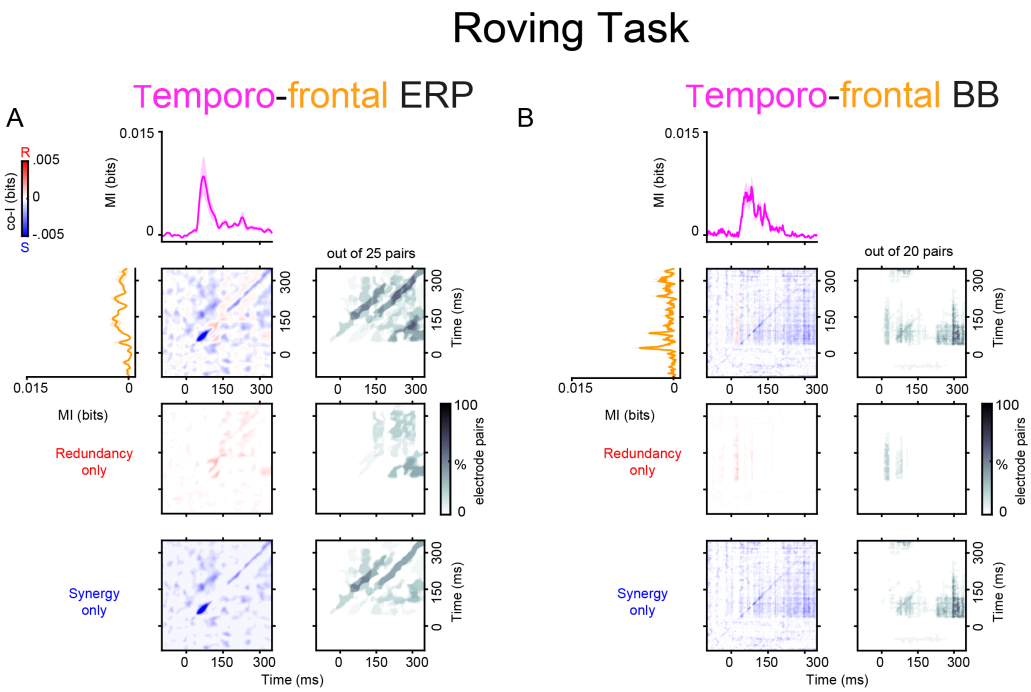

Figure 4: Spatio-temporal synergy and redundancy between auditory and frontal electrodes in the Roving Task (Experiment 1). Co-information revealed synergistic and redundant spatio-temporal patterns between auditory and frontal electrodes in the ERP (Panel **A**) and BB (Panel **B**) signals for the Roving task. MI (solid traces) between standard and deviant trials for temporal (pink color) and frontal (orange color) electrodes. Error bars represent standard error of the mean (S.E.M) across electrodes. Co-I was computed between each pair of electrodes and across time points between -100 to 350 ms after tone presentation. The average of the temporo-frontal pairs across the three monkeys is shown for the complete co-I chart (red and blue plots); for the positive co-I values (redundancy only; red plot); and for the negative co-I values (synergy only; blue plot). The proportion of electrode pairs showing significant co-I differences is shown in the corresponding grey-scale plots. The average co-I charts for the individual monkeys are shown in Figures S3 for the ERP signals, and in Figure S6 for the BB signals.

We observed significant temporal redundancy in the auditory (Figure 3B) and frontal (Figure 3D) BB signals. For auditory BB signals, the dynamics of the redundant patterns were observed along the diagonal of the co-I chart, they were sustained over time and observed between time points around the early MI peaks (i.e., during the transient period when the effect sizes are larger between tones) (Figure 3B; grey clusters ~120-280 ms after tone presentation). In the frontal electrodes, we observed significant clusters of sustained redundant interactions around later time points (Figure 3D; grey cluster around 300 ms after tone presentation).

Spatio-temporal synergy and redundancy

The finding that multiple recording sites encode information about PE raises the question of whether these regions are dynamically interacting and whether these inter-areal interactions are redundant or synergistic. To test this possibility, we characterized the redundancy and synergy between auditory and frontal electrodes. Spatio-temporal co-I was computed between the auditory and frontal electrodes over time (Figure 4) and averaged across monkeys separately in the Roving Oddball Task (i.e. ERP and BB signals). The dynamics of spatio-temporal synergy in the ERP and BB signals showed complex and heterogenous patterns between early time points of the auditory electrodes and later time points in the frontal electrodes (Figure 4). For example, while the ERP signals encoded both diagonal (Figure 4A; grey clusters ~100-350 ms after tone presentation) and off-diagonal synergistic patterns (Figure 4A; grey clusters ~150-350 ms after tone presentation), the BB signals mainly showed off-diagonal synergy between temporal and frontal electrodes (Figure 4B; grey clusters ~220-350 ms after tone presentation). In Figure 4A, the diagonal stripes suggest the possibility of oscillatory dynamics, where the representation in frontal regions between 50-300 ms is enhanced by knowledge of the activity of temporal regions ~50 ms earlier (the upper diagonal line). Note that 50 ms peak-to-peak timescale corresponds to a frequency of ~10 Hz, i.e. the alpha range. In Figure 4B the off-diagonal block suggests that the frontal representation of the stimulus between 20-120 ms initiates a state change: later temporal activity (200 ms+) enhances the readout of the stimulus class, even though there is no representation of PE in the BB signal of the temporal area at that time.

Experiment 2: Local/Global Task
Temporal synergy and redundancy

Although PE processing has been widely studied using the Roving Oddball Task (Canales-Johnson et al., 2021), the contribution of stimulus-specific adaptation (SSA) to the amplitude

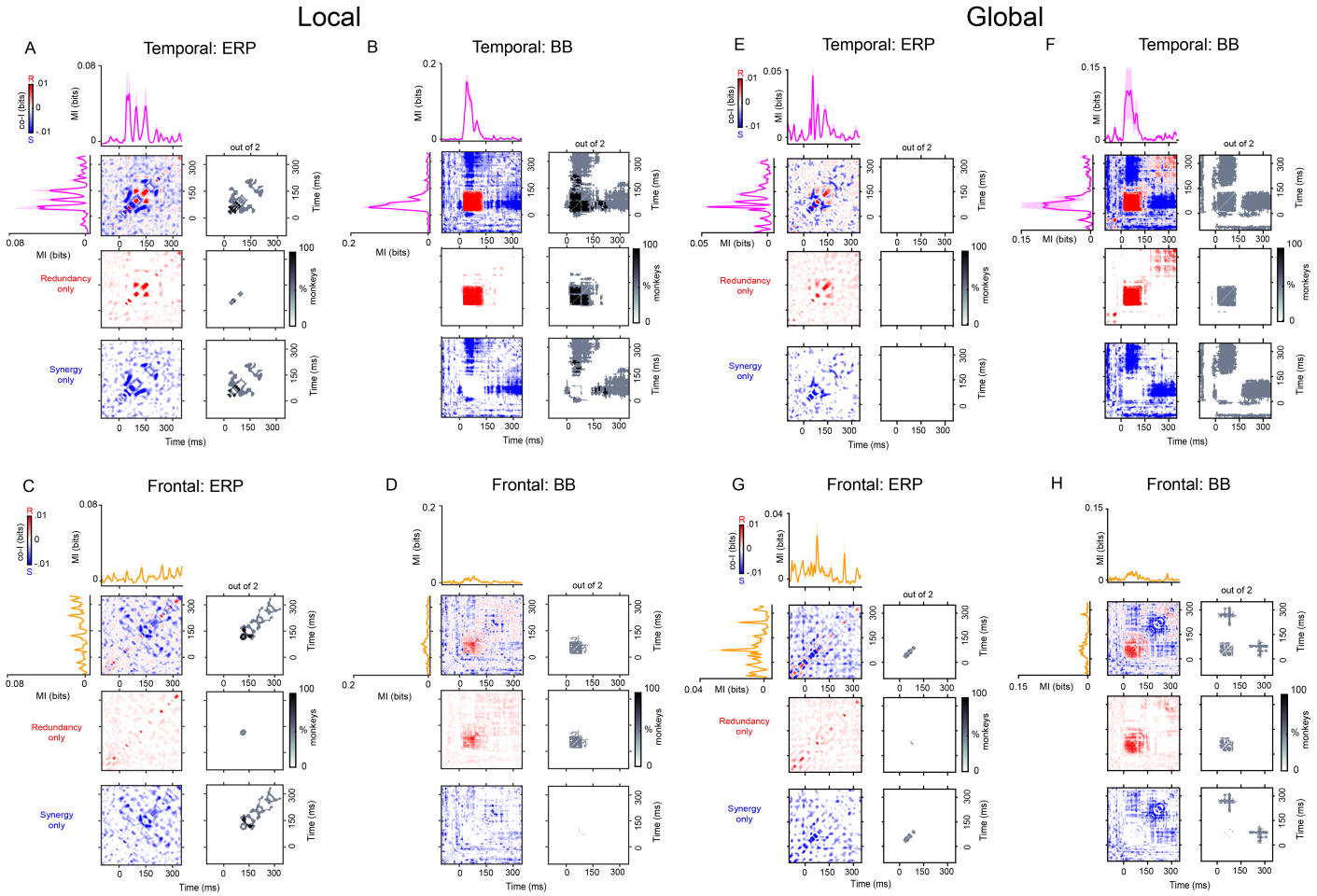

Figure 5: Temporal synergy and redundancy within ERP and BB signals in the auditory and frontal electrodes with the highest MI for the Local/Global Task (Experiment 2). In the Local and Global contrasts, co-information revealed synergistic and redundant temporal patterns within ERP (Panels **A, E**) and BB (Panels **B, F**) signals in the auditory cortex, and within ERP (Panels **C, G**) and BB (Panels **D, H**) signals in the frontal cortex. MI (solid traces) between standard and deviant trials for auditory (pink color) and frontal (orange color) electrodes averaged across the three monkeys. Error bars represent standard error of the mean (S.E.M) across electrodes. Temporal co-I was computed within the corresponding signal (ERP, BB) across time points between -100 to 350 ms after tone presentation. The average of the corresponding electrodes across monkeys is shown for the complete co-I chart (red and blue plots); for positive co-I values (redundancy only; red panel); and for negative co-I values (synergy only; blue plot). The grey-scale plots show the proportion of monkeys showing significant co-I differences in the single electrodes analysis depicted in Figure S2.

of the ERP response is usually considered a confounding factor in the isolation of PE (Parras et al., 2017). For this reason, we also investigated synergy and redundancy in a separate task capable of attenuating the effects of SSA (i.e. the Local/Global Task). In the Local contrast, although we observed temporal synergy in both ERP and BB signals, the off-diagonal synergy was primarily observed between early and late time points of the BB signals in the temporal cortex (Figure 5; grey clusters ~150-350 ms after tone presentation). The ERP signals, on the other hand, showed diagonal synergy in both the temporal (Figure 5; grey clusters ~40-150 ms after tone presentation) and frontal cortex (Figure 5; grey clusters ~150-350 ms after tone presentation).

Another advantage of the Local/Global Task is the possibility of exploring a higher-order PE observed as a violation of the overall sequence of tones (Global contrast; see Figure 1B, and METHODS). This context-dependent deviancy effect has been shown to elicit neural activation in frontal regions (Chao et al., 2018; Jiang et al., 2022). In the case of the Global contrast, we observed temporal synergy across early and late time points but mostly in the BB signals both within the auditory (Figure 5F; grey clusters ~0-350 ms after tone presentation) and frontal electrodes (Figure 5H; grey clusters ~230-330 ms after tone presentation). Taken together, these results suggest that the Local/Global task elicits distributed patterns of PE information across time that are primarily encoded by firing rates.

Spatio-temporal synergy and redundancy

We investigated whether local and higher-order PE are encoded by synergistic information between cortical regions. Thus, we characterized the synergy (and redundancy) between auditory and frontal electrodes for the local and global contrast (Figure 6). Spatio-temporal co-I was computed between the auditory and frontal electrodes over time and averaged across

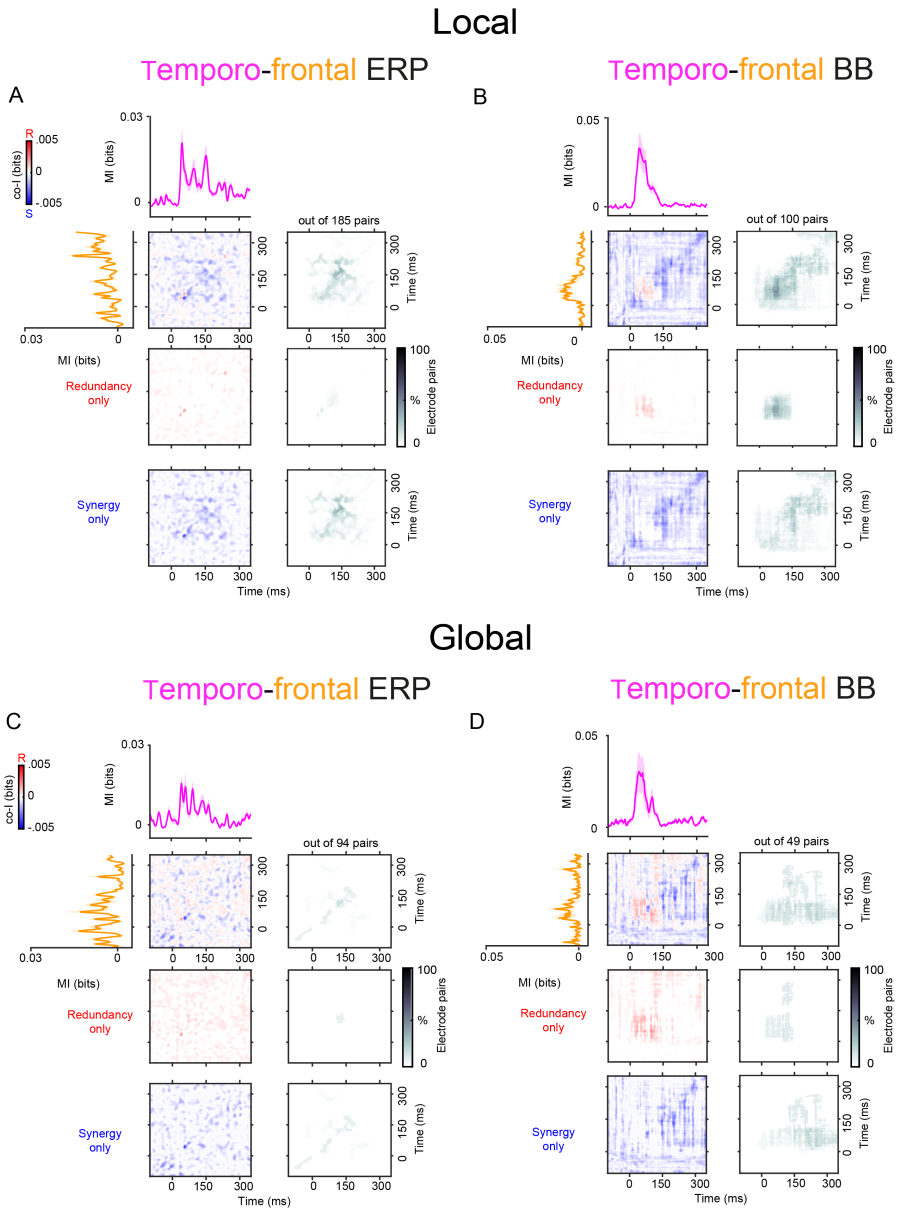

Figure 6: Spatio-temporal synergy and redundancy between auditory and frontal electrodes in the Local/Global Task (Experiment 2). Co-information revealed synergistic and redundant spatio-temporal patterns between auditory and frontal electrodes in the ERP (Panels **A, C**) and BB (Panels **B, D**) signals. MI (solid traces) between standard and deviant trials for temporal (pink color) and frontal (orange color) electrodes. Error bars represent standard error of the mean (S.E.M) across electrodes. Co-I was computed between each pair of electrodes and across time points between -100 to 350 ms after tone presentation. The average of the temporo-frontal pairs across the three monkeys is shown for the complete co-I chart (red and blue panel); for the positive co-I values (redundancy only; red panel); and for the negative co-I values (synergy only; blue panel). The proportion of electrode pairs showing significant co-I differences is shown in the corresponding grey-scale panels. The average co-I charts for the individual monkeys are shown in Figures S4, S5 for ERP signal, and Figures S7, S8 for the BB.

two monkeys separately for each signal (i.e. ERP and BB signals). Consistent with the effects observed in the Roving Oddball Task, we observed multiple patterns of synergistic and redundant information between temporal and frontal regions. We also noticed an interesting difference between the two tasks. While the Roving Oddball Task elicited most of the synergistic interactions between ERP signals (Figure 4A), the Local/Global Task elicited most of the synergy between BB signals (Figure 6B,D).

Although the per-electrode and electrode-pair analyses of synergy and redundancy exploit the optimal spatial resolution of the recording modality across temporal and frontal regions, they could also miss information encoded in the spatial pattern both within and between temporal and frontal areas. They could therefore potentially miss synergy or redundancy that is only apparent when considering multiple electrodes together, either due to a low signal-to-noise ratio within each channel or because of a genuinely distributed informative spatial pattern. This might be particularly relevant for the ERP signals that showed extensive temporal and frontal PE effects (Figure

1A,B). Thus, to account for potential informative spatial patterns of redundancy and synergy in ERP signals, and to reduce any concern about high-order interactions between channels within each region in the pairwise channel analysis, we complemented our analyses by computing co-information based on the response across multiple electrodes (MVCo-I: Multi-Variate Co-information) (Figure S11 and S12). In brief, we have applied a cross-validated multivariate analysis approach that uses machine learning to capture the best linear representation of the prediction error signal across a whole region, and we have repeated our co-information analyses within and between the two brain regions of interest using the classifiers' outputs (frontal and temporal) (see METHODS).

The MVCo-I analyses within-region (Figures S13 and S15) and between-regions (Figures S14 and S16) showed comparable co-I in terms of synergistic and redundant dynamics observed in the per-electrode (Figures 3 and 5) and in the between-electrodes (Figures 4 and 6) analyses, but with increased statistical power (i.e., increased MI).

To sum up, we observed widespread patterns of synergy within and between ERP and BB signals across the auditory cortical hierarchy. The distributed nature of the temporal and spatio-temporal synergistic interactions across the cortical hierarchy raises the question of whether the emergence of synergy is a consequence of the recurrent and feedback connections in the auditory network.

Explaining the presence of synergy using a brain-constrained neurocomputational model

To test the validity of our working hypothesis that synergistic information may be driven mainly by recurrent and feedback connections, we applied an existing neural network model closely reproducing structural and functional properties of relevant areas in the superior-temporal and inferior-frontal lobes of the primate brain (Figure 7A) to simulate auditory PE processing (see METHODS). Our approach was to employ a fully brain-constrained neurocomputational model that accurately replicates critical neurobiological and neuroanatomical features of the mammalian cortex (Pulvermüller et al., 2021) and to stimulate this model using the same Roving Oddball Task adopted in Experiment 1. The network simulated neuronal firing rates, which are the main contributors of the BB signals; we, therefore, honed our modelling efforts on reproducing the BB signals observed in Experiment 1 (Figure 3B, D and Figure 4B). The model we used (see Figure 1E) has been previously applied to successfully simulate and explain automatic auditory change detection and the Mismatch Negativity (MMN) response to familiar and unfamiliar sounds in the human brain (Garagnani and Pulvermüller, 2011) and several other phenomena in the domains of language acquisition and processing, attention, memory, and decision making (Tomasello et al., 2019, 2017; Schomers et al., 2017; Pulvermüller and Garagnani, 2014; Garagnani and Pulvermüller, 2013; Garagnani et al., 2008) — see Pulvermüller et al. (2021) for a recent review.

Here, we recorded the network's responses (measured in each area as the sum of all cells' firing rates) to predefined random patterns simulating standard and deviant tones; following the same procedure used to process the experimental data (see METHODS), we then analysed the resulting PE signals. We observed that, before any adjustment of its parameter values, the network already encoded both redundant and synergistic information, specifically, in the signal from its superior-temporal region (including areas A1, AB, PB). We then further constrained the model's dynamics by fine-tuning three of its parameters (i.e., the strength of the neuronal adaptation, the local inhibition, and the between-area links) so that the temporal and spatial features of synergistic information encoded in the simulated PE responses would closely resemble those we observed experimentally. This process of parameter tuning did not qualitatively change the network's responses but simply improved the fit of the responses with the observed data.

To quantify the similarity of the co-I values between the real and the simulated data, we computed the Structural Similarity Index (SSIM). The SSIM assesses the structural similarity between two images, with values ranging from 0 (dissimilar) to 1 (highly similar). Hence, we converted the co-I plots of the real and simulated data to images and computed SSIM between them. While the SSIM between simulated and experimental co-I was 0.74 (Figure 7B versus 3B, the frontal cortex comparison showed an SSIM of 0.83 between simulated and experimental co-I (Figure 7C versus Figure 3D). Both values were significant above chance level ($p < 0.05$) after comparing them to a distribution of surrogate SSIM values. The surrogate distribution was obtained by computing the SSIM between the experimental co-I image and a shuffled version of the simulated co-I image, and repeating this procedure 1000 times.

Having attained a good match between experimental and simulated data (e.g, compare Figure 7 panels B and C, with Figure 3 panels B and D, respectively), we then moved on to address the main question, i.e., whether the presence of feedback and recurrent links in the underlying neural network has a direct impact on the emergence of synergy. To investigate this, we directly manipulated the model's structural connectivity and analysed the effects of such manipulation on its responses to the same stimuli. Specifically, we used two types of architectures: first, a fully-connected model (FC), having connectivity as shown in Figure 7A, i.e., including both feedforward and feedback between-area, and recurrent within-area, connections. It is important to stress that such connectivity reflects the neuroanatomical links known to exist between (and within) corresponding superior-temporal and inferior-frontal cortices in the macaque brain, as well as between the human's homologue cortical areas (see METHODS for details). Second, we ran a set of feed-forward-only (FF) model simulations, in which we artificially cut all the feedback and recurrent links of the FC architecture, while maintaining all the feedforward ones intact. By *feedback links* here we refer to all the links in the model going from right to left in Figure 7A (i.e., from area PB to A1, from area PF to AB, from PM to PB, from M1 to PF, and from each area to its left-hand side next-neighbour). For each of these two model types, we ran three distinct simulations. In each simulation, the projections linking any two areas (and any area to itself) were *sparse* and established at random, with the proba-

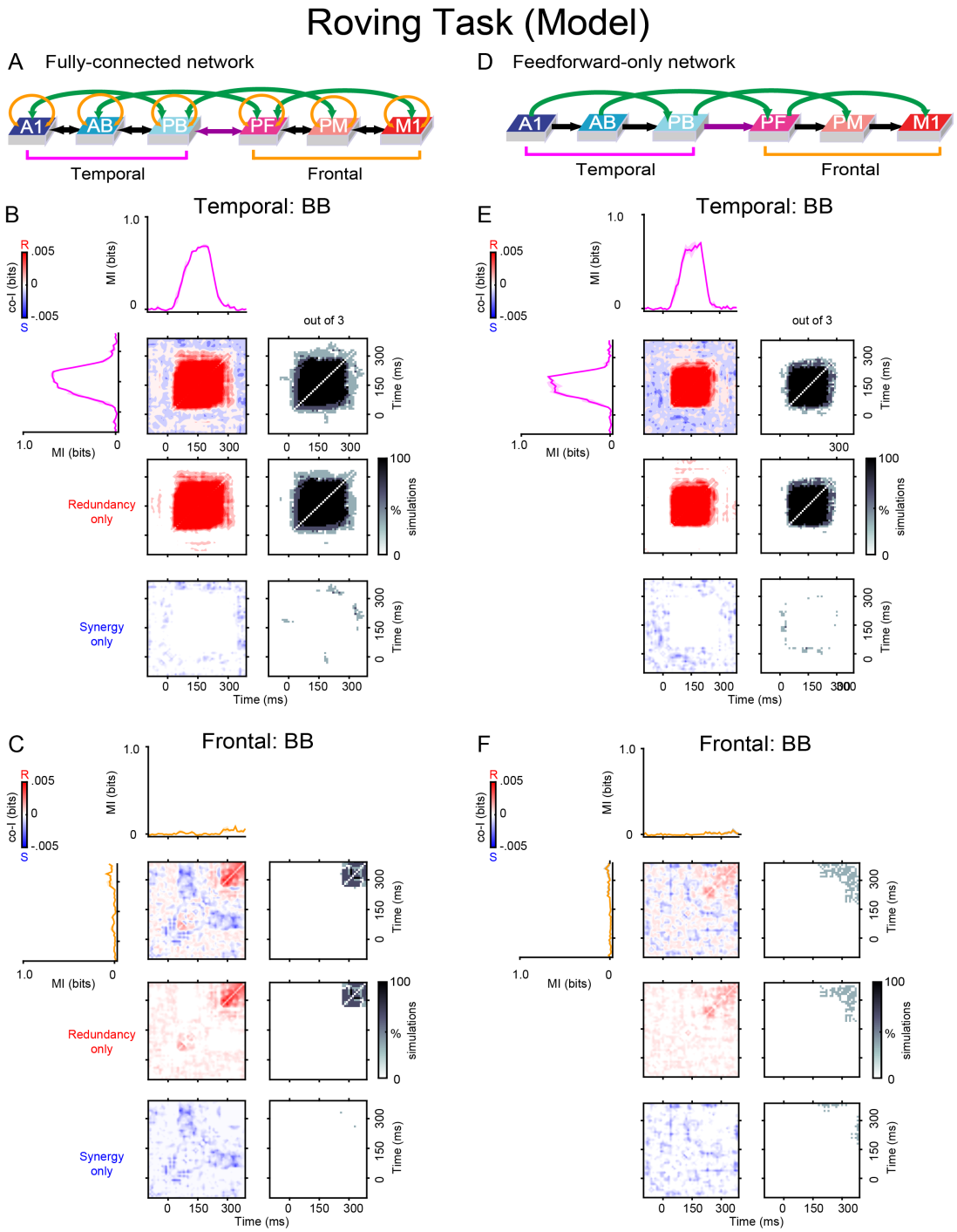

Figure 7: Model architectures and simulation results. A brain-constrained model of temporal and frontal areas of the marmoset brain (see Fig. 1E) was stimulated with simulated tones as in the Roving Oddball Task used in Experiment 1. (**A,D**): Different network architectures used for the simulations (see METHODS). Feedforward and feedback between-area connections are depicted as black and green arrows; recurrent within-area links (panel A only) are shown in gold. Input stimuli were repeatedly presented to area A1 of the network (model correlate of primary auditory cortex) and firing rate responses of each excitatory cell within the six areas were recorded. (**B,C**): Results obtained with networks having a Fully Connected (FC) architecture (shown in panel A), which included both feedforward/feedback links and recurrent connections. (**E,F**): Results obtained using networks having a Feedforward-only (FF) architecture (panel D), in which the feedback and recurrent connections were absent. MI (solid traces) between standard and deviant trials averaged across three simulation runs (each run modelling a single monkey dataset) are plotted for the three temporal (A1, AB, PB: pink curves) and three frontal (PF, PM, M1: orange curves) areas' simulated responses. Error bars represent standard error of the mean (S.E.M). Co-information analyses were performed on the model temporal and frontal areas' signals. Temporal co-I was computed within the simulated firing rates across time points between -100 to 350 ms after stimulus onset. The average of the corresponding electrodes across simulated monkey datasets is shown for the complete co-I chart (red and blue panel), for positive co-I values (redundancy only; red panel) and for negative co-I values (synergy only; blue panel). The grey-scale panels show the proportion of simulated monkey datasets with the highest MI within the temporal (A1, AB, PB) and frontal (PF, PM, M1) regions. Note the similarity (in terms of temporal patterns of synergy and redundancy) between the results obtained from the FC model responses (panels **B** and **C**) and those from the corresponding experimental data (the BB signal shown in Figure 3, panels **B** and **D**, respectively).

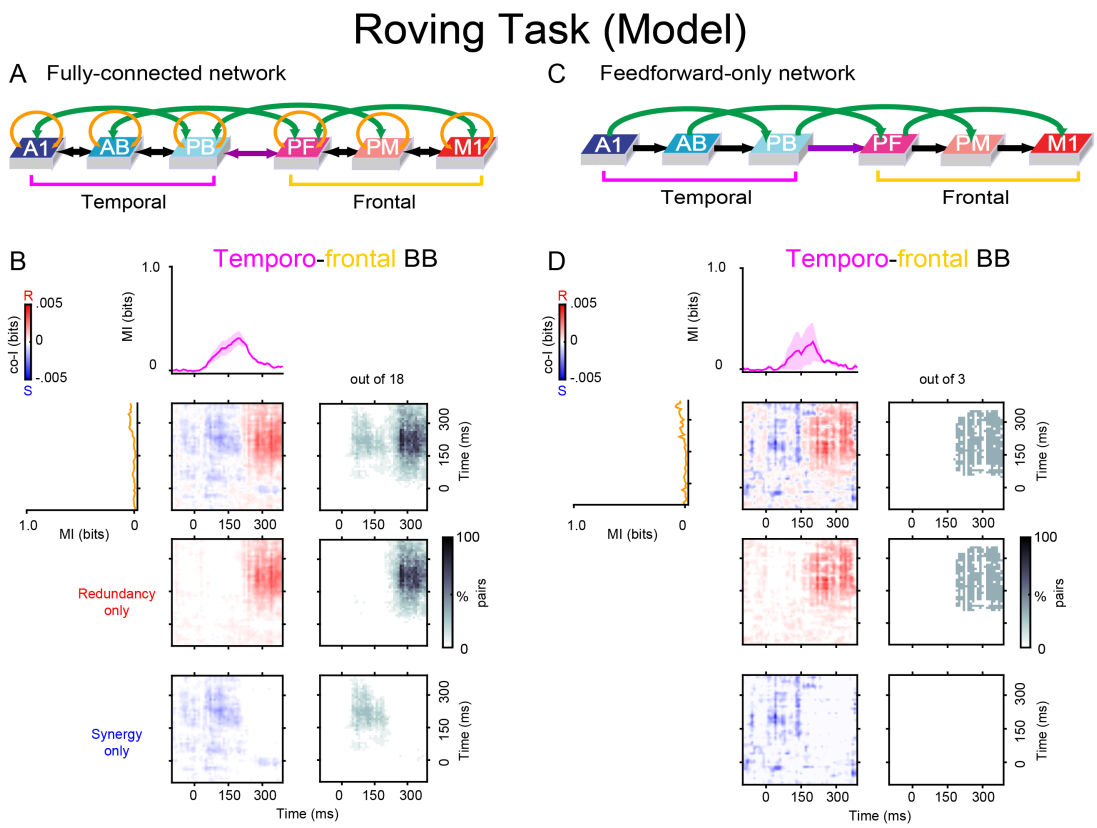

Figure 8: Model architectures and simulation results. The firing rate responses of the networks used to produce the results of Figure 7 were subjected to co-information analyses between the simulated temporal and frontal areas' signals. (**B**) Results obtained using Fully Connected (FC) networks (panel **A**), which included both feedforward and feedback (black and green arrows) links and recurrent (golden arrows) connections. (**D**) Results obtained using Feedforward-only (FF) networks (panel **C**), in which the feedback and recurrent connections were absent (see METHODS). MI (solid traces) between standard and deviant trials averaged across three simulation runs (each run modelling a single monkey dataset) are plotted for the three temporal (A1, AB, PB: pink curves) and three frontal (PF, PM, M1: orange curves) areas' simulated responses. Error bars represent standard error of the mean (S.E.M). Co-information analyses were performed between the model temporal and frontal areas' signals. Temporal co-I was computed from the simulated firing rates across time points between -100 to 350 ms after stimulus onset. The average of the corresponding electrodes across simulated monkey datasets is shown for the complete co-I chart (red and blue panel), for positive co-I values (redundancy only; red panel), and for negative co-I values (synergy only; blue panel). The grey-scale panels show the proportion of significant co-I pairs between superior-temporal (A1, AB, PB) and frontal (PF, PM, M1) areas using areas that showed significant MI between standard and deviant trials. Note the similarity (in terms of spatio-temporal patterns of synergy and redundancy) between the results obtained from the model responses and those from the corresponding BB signals in the experimental data of Figure 4: co-I measures of network responses show significant synergy between temporal and frontal regions (see panel **B**), as observed in real marmoset data (Figure 4B). Also, note that such synergistic effects disappear after the removal of the network's feedback and recurrent links (compare the bottom-right plot of panel **B**, FC architecture, against that of panel **D**, FF architecture).

bility of any two cells being connected by a synapse decreasing with their (modelled) cortical distance. The weights of all the synapses were also set to a small random value comprised between 0 and 0.1 (see METHODS for details).

Given this, we treated each simulation run as the model correlate of a single marmoset in Experiment 1 (Fr, Kr, and Go), as it was produced using a slightly different (random) variation of the same prototype network architecture (FC or FF). During each simulation, we generated and recorded 100 trials (50 deviants and 50 standards) and analysed the co-information within and between the resulting network responses in exactly the same way as in the experimental data. The results showed that the FC model showed highly synergistic interactions between temporal and frontal regions (Figure 8A,B). Crucially for our hypothesis, we observed that the removal of all the between-area feedback projections and recurrent within-area links of the network entirely prevented the emergence of synergistic interactions between frontal and temporal model regions (see Figure 8D). Additional simulations obtained with a version of the architecture containing just nearest-neighbour between-area feedback (and feedforward) links, along with recurrent ones (see Figure S10) again failed to reproduce such synergistic interactions, indicating that it is not simply the presence of feedback projections, but specifically of *higher-order*, or so-called "jumping" (Schomers et al., 2017), cortico-cortical links connecting non-adjacent areas of the processing hierarchy that is needed for synergy to emerge in the model.

DISCUSSION

In this study, we focused on computing temporally-resolved metrics of redundancy and synergy, aiming at investigating the dynamics of the information interactions within and between

cortical signals encoding PE. Due to the interplay between temporal and spatial neural dynamics, our approach revealed a rich repertoire of redundant and synergistic patterns, showing transient and sustained information dynamics distributed across the auditory hierarchy.

There are a wide variety of ways of using information-theoretic, and other measures to study representational interactions in neural coding (Chicharro, 2014). Schneidman et al. (2003) discuss three types of response independence in the context of spiking neuron population coding: activity independence, conditional independence and information independence. Here we focus only on information independence, as we are interested in relating the information representation between areas. Deviations from information independence are best measured with co-information. To date, co-information has been less frequently applied to aggregate signals as we do here (Ince et al., 2017, 2016; Zhan et al., 2019)

Interpreting synergistic interactions

Synergistic information was observed mainly off-diagonally, i.e. between early and late times points after tone presentation for both within (Figures 3-5) and between cortical areas (Figure 6). This indicates that late temporal responses carry information that, in combination with the early one, provides extra information about the identity of the tone (standard or deviant) than when considered in isolation. This raises the question about what is the functional relevance of synergistic information for representing prediction errors. The off-diagonal synergy between early and late time points could be a signature of a neural state shift. It is interesting to note that the synergy remains strong over periods after the PE response is no longer represented (i.e. no MI at those time points). However, the initial representation of the PE may have shifted the local network dynamics into a different state. Then knowing this ongoing state improves the readout of the encoded information at the earlier time point. Thus, the off-diagonal synergy might be an echo of the initial PE representation that is not directly observable in later time points.

Synergy can also arise from a common source of neural noise that is non-stimulus specific. For example, the spatio-temporal synergy between regions could reflect a global change in attention or arousal. In this situation, the readout of one area provides information about the global neural state even when it does not convey information about the PE directly, and this can be used to improve the resolution with which the PE can be decoded from the other area. Although this might be a possibility, the tight timing of the synergy bands observed in both experiments (i.e, diagonal and off-diagonal synergistic patterns) speak more of a transient dynamics rather than global ongoing fluctuations underlying the spatio-temporal synergy.

How does synergistic information emerge from local and distributed neural dynamics? We reason that a plausible neurobiological mechanism for the generation of synergy is through recurrent and feedback neural interactions within and between areas, respectively. This hypothesis is corroborated by the novel neurocomputational results presented here. Specifically, we adapted a 6-layer-deep, brain-constrained neural network model reproducing the neuroanatomy and neurophysiology of language areas in the temporal and frontal cortex of the human brain (Garagnani and Pulvermüller, 2011; Schomers et al., 2017; Garagnani et al., 2008) to simulate and explain the cortical mechanisms underlying the generation of the PE responses that we observed experimentally in the marmoset monkey. In response to oddball stimulation with simulated auditory tones, the model was found to produce responses containing both synergistic and redundant information (here we looked at the network's per-area sum of all cells' firing rates; this measure can be related directly to the experimentally recorded BB signal). By tuning the model parameters, we were able to get the network's spatio-temporal patterns of synergy and redundancy encoding the PE response to closely replicate those found experimentally across the auditory cortical hierarchy; furthermore, manipulation of the model's connectivity revealed that synergistic interactions emerged in it only when strong, higher-order ("jumping") forward and backward links (green arrows in Figure 1E) connecting frontal and temporal regions were present.

Based on this computational result, we conjecture that the cortical homologues of such jumping links (known to exist between corresponding regions of the marmoset brain, see METHODS) may play a similarly crucial role in the emergence of the temporo-frontal synergistic interactions observed in the ECoG data. This prediction awaits further validation through experimental testing.

Interpreting redundant interactions

A different type of dynamics was observed in the case of the redundant information across the cortex. Redundant patterns of information were observed mainly at time points close to the diagonal of the co-I chart, both within signals (Figures 3-5) and between signals (Figure 6). The advantage of computing redundancy is that it reveals to which extent local and inter-areal signals represent the same information about the stimuli category on a trial-by-trial basis. Redundant interactions about tone category (i.e., deviant or standard) were observed in the ERP and BB signals and represented the outcome of the shared information across time points (temporal redundancy) and between areas (spatio-temporal redundancy). These observed redundancy patterns raise the question of what is the functional relevance of redundant information for processing PE across the cortex.

A neurobiological interpretation of redundancy is that the neural populations encoding this type of information share a common mechanism (Ince et al., 2017). From the perspective of cortical dynamics, redundancy then could provide cortical interactions with robustness (Luppi et al., 2022; Olivares et al., 2022), as redundant interdependencies convey information that is not exclusive to any single cortical region. Robustness, understood as the ability to tolerate perturbations that might affect network functionality (Luppi et al., 2022), is a desirable characteristic of cortical networks processing predictions to preserve stimuli separability in the presence of highly variable stimuli features, environmental noise, or endogenous sources of noise such as background neural activity. Thus, our results suggest that redundancy quantifies the robustness of the information processing in the cortex, enabling multiple areas to process

common information about prediction errors. One way this could arise is from multiple regions receiving the same input evidence for processing in different parallel pathways.

Differences in redundancy and synergy between tasks

The employed tasks all showed distinct patterns of synergistic and redundant dynamics. The Roving Oddball Task elicited synergistic information mostly within the ERP signal, while the local deviant in the Local/Global Task displayed temporally distributed synergy within both ERP and BB signals. A possible explanation for this is that the MMN-response for the Roving Oddball Task could primarily reflect stimulus-specific adaptation at the level of the auditory cortex (O'Reilly, 2021), while the Local/Global Task shows smaller effects relating to SSA due to the 20-sequence adaptation period at the start of each testing run (Chao et al., 2018; Jiang et al., 2022).

While the local deviant in the Local/Global Task showed highly distributed synergistic information across brain areas and for both monkeys, the patterns observed for the global deviant were more monkey-dependent. Strong effects were observed within signals (ERP and BB) and between brain areas (temporo-frontal) for monkey Ji, but monkey Nr exhibited minimal effects within and between all cortical regions (See Figure S2). A speculative explanation for the lack of a global effect is that higher-order deviants can be driven by top-down attention (Chennu et al., 2013; Bekinschtein et al., 2009). In this case, the lack of effects in Nr could be simply explained by a lack of interest in the experimental stimuli.

Distributed processing across cortical areas: Implications for predictive coding

Our findings might have ramifications for predictive coding theories. For example, the information encoding PEs was not merely redundant but also highly synergistic across areas. In principle, the lack of redundancy between PEs is in consistent with hierarchical predictive coding (HPC) because HPC entails that prediction errors are independently generated in different levels of the hierarchy (Friston and Kiebel, 2009; Rao and Ballard, 1999). However, synergy corresponds to the extra information obtained when signals are considered together, suggesting that there is a more holistic or complementary representation of PE rather than just "independently" generated PEs, with the correlational structure of the signals conveying additional information. Possibly, this synergy results from the recurrent (including long-distance feedback) interactions across many nodes, as suggested by the results of the simulations obtained with a brain-constrained neurocomputational model.

Conclusion

Our results support the notion that PE information is broadcasted by transient, aperiodic neural activity across the cortex (i.e. ERPs and BB signals) (Vinck et al., 2023). By distinguishing the type of information encoded by these inter-real interactions, we have shown that PEs not only share information across regions but also encode complementary information between distributed signals. Thus, our results demonstrate that distributed representations of prediction error signals across the cortical hierarchy are highly synergistic.

Authorship contributions

Conceptualization: ACJ, RI, and MV. Data analysis: FG, ACJ, LR and JÄ. Visualization: FG, ACJ, and JÄ. Marmoset recordings and surgeries: MK. Software and methods for marmoset- and simulation-data analysis: RI, MJ, LR, and KM. Neurocomputational modelling: MG. Writing and editing: ACJ, FG, JÄ, CU, MG and MV. Supervision: ACJ.

Acknowledgments

ACJ thanks Dr. Daniel Bor, Dr. Valdas Noreika, Dr. Tristan Bekinschtein, and Dr. Dora Hermes for contributing to valuable discussions. MG thanks Dr. Friedemann Pulvermüller for the invaluable discussions and Dr. Thomas Wennekers for his substantial help in developing the original neurocomputational architecture. This project was financially supported by a Bial Foundation Grant (ID: A-29477) to ACJ.

METHODS

Data acquisition

This study used ECoG recordings from five adult male common marmosets (*Callithrix jacchus*). The details of the datasets for three of the monkeys (Kr, Go and Fr) have been described previously in Canales-Johnson et al. (2021); Komatsu et al. (2015), and Jiang et al. (2022) for two of the monkeys (Ji and Nr).

For marmosets Kr, Go and Fr, (i.e., animals that performed the Roving Oddball Task) the ECoG recordings were acquired in a passive listening condition while the monkeys were awake. During the recording sessions, the monkeys Go and Kr sat on a primate chair in a dimly lit room, while monkey Fr was held in a drawstring pouch, which was stabilized in a dark room. Every session lasted for about 15 minutes of which the first 3 minutes of data were used for various standard stimuli and the remaining 12 minutes of data acquisition were dedicated to the Roving oddball sequences. For the data analysis, we acquired a total of three sessions for monkey Fr, which resulted in 720 (240 $\times$ 3) standard and deviant trials, and six sessions for monkeys Go and Kr, resulting in 1440 (240 $\times$ 6) standard and deviant trials. For the recordings, a multi-electrode data acquisition system was used (Cerebus Blackrock Microsystems, Salt Lake City, UT, USA) with a band-pass filter of 0.3–500 Hz and then digitized at 1 kHz. In the signal pre-processing, those signals were re-referenced using an average reference montage, and high-pass filtered above 0.5 Hz, using a 6th-order Butterworth filter.

The recording was done with chronically implanted, customized multielectrode ECoG electrode arrays (Cir-Tech Inc., Japan). Before implantation with the ECoG electrode arrays, the monkeys were anesthetized and further suffering was minimized. All electrodes were implanted in epidural space; 28 in the left hemisphere and an additional 4 in the frontal cortex of the right hemisphere of monkey Fr, 64 in the right hemisphere of monkey Go, and 64 in the right hemisphere of monkey Kr. In the 32-electrode array, each electrode contact was 1 mm in diameter and had an inter-electrode distance of 2.5 - 5.0 mm (Komatsu et al., 2015). In the 64-electrode array, each electrode contact was 0.6 mm in diameter and had an inter-electrode distance of 1.4 mm in a bipolar pair (Komatsu et al., 2019). The electrode arrays covered the temporal, parietal, frontal, and occipital lobes.

For marmosets Ji and Nr, (i.e., the animals that performed the Local/Global Task) the EcoG recordings were also acquired in a passive listening condition while the monkeys were fully awake. The monkeys were seated in sphinx position with their head fixed in a sound-attenuated and electrically shielded room. The recording was done with chronically implanted, multielectrode (96) ECoG electrode arrays (Cir-Tech Inc., Japan). For data analysis, electrodes in the temporal and frontal cortices of the marmosets were used. This was done due to the public availability of the data from these electrodes (Jiang et al., 2022). Monkey Ji had a total of 27 electrodes (16 temporal, 11 frontal), and monkey Nr had a total of 39 electrodes (25 temporal, 14 frontal). The data was recorded with a Grapevine NIP system (Ripple Neuro, Salt Lake City, UT) with a sampling rate of 1khz.

All surgical and experimental procedures were performed following the National Institutes of Health Guidelines for the Care and Use of Laboratory Animals and approved by the RIKEN Ethical Committee (No. H26-2-202, for monkeys Kr, Go, and Fr and No. W2020-2-008(2) for monkeys Ji and Nr). The locations of the implanted electrodes of each monkey are found in Figure 2.

Experimental tasks

For the Roving Oddball Task, monkeys Kr, Go and Fr were subjected to a Roving oddball paradigm (Canales-Johnson et al., 2021). Trains of 3, 5, or 11 repetitive single-tones of twenty different frequencies (250-6727 Hz with intervals of 1/4 octave) were presented in a pseudo-random order. Within each tone train the presented tones had the same frequency, but between tone trains the frequency was different. As the tone trains followed each other continuously, the first tone of a train was considered an unexpected deviant tone, because the preceding tones were of a different frequency, while the expected standard tone was defined as the final tone in a train because the preceding tones were of the same frequency (Figure 1A). The presented tones were pure sinusoidal tones that lasted for 64 ms (7 ms rise/fall) and the time between stimulus onsets was 503 ms. Stimulus presentation was controlled by MATLAB (MathWorks Inc., Natick, MA, USA) using the Psychophysics Toolbox extensions (Brainard and Vision, 1997). Two audio speakers (Fostex, Japan) were used to present the tones with an average intensity of 60 dB SPL around the animal's ear.

For the local/global task, monkeys Ji and Nr were subjected to a standard local/global auditory oddball paradigm (Jiang et al., 2022). The monkeys heard tone trains with either a local regularity (five identical tones played in a sequence; xxxxx) or global regularity (five tones, the first four of which were identical, and where the fifth was of a different frequency; xxxxY). To create a local deviant, the last tone of the local tone train (xxxxx) was sometimes played at a different frequency as the earlier tones in the train (local deviant; xxxxY). To create a global deviant, the last tone of the global tone train (xxxxY) was sometimes played with the same frequency as the earlier tones in the train (global deviant; xxxxx). The frequencies for the tones x or Y were either 707 or 4000 Hz. The presented tones were pure sinusoidal tones that lasted for 50 ms with an intertone interval of 150 ms, and they were presented to the monkeys bilaterally with two speakers (Fostex, Japan) from the distance of approximately 20 cm from the head with the average intensity of 70 DB.

Each testing period started with a 14-second resting phase, which was followed by a habitation period during which the specified standard (local or global) was presented 20 times to ensure that the monkey learns the regularity of the tone trains. For a testing run, three blocks of 25-tone trains were played, with a 14s resting phase in between. Out of the 25 trials, 20 (80 percent) were of the specified standard (local or global) and five (20 percent) were deviants. For the global deviants, more than one local standard was always played after to ensure global

consistency. Each run lasted for 6 minutes and 46 seconds, and each session consisted of 3-4 local standard and 3-4 global standard runs, depending on the marmoset's performance during the day. The order of the tasks was randomised, and the frequencies for tones x and Y were balanced. For the analysis, the number of trials for standard and deviant trials had to be equal. This resulted in 330 (local deviant) and 243 (global deviant) trials for Monkey Ji, 251 (local deviant), and 212 (global deviant) for Monkey Nr.

ERP and BB analyses

For further analysis, the raw ECoG voltage responses have been transformed into ERP and BB as described in Canales-Johnson et al. (2021). In brief, common average referencing was used to re-reference the ECoG recordings across all electrodes, and the data was downsampled to 500 Hz. For obtaining ERPs, a low-pass filter of 1-40 Hz was applied for the ERP analysis. Standard and deviant tones were categorized as described before. Epochs of -100 ms to 350 ms around the onset of the tones were taken, and a baseline correction was applied by subtracting the mean voltage during the 100 ms period before the stimulus onset from the total epoch.

To obtain the BB, spectral decoupling of the raw ECoG was carried out (Canales-Johnson et al., 2021; Miller, 2019). To extract the course of broadband spectral activity, the spectral decoupling of the raw ECoG signal was carried out. As for the ERP analysis, common average referencing was used to re-reference the ECoG potentials of all the electrodes. Epochs of -100 ms to 350 ms around the onset of the tones were used to calculate discrete samples of power spectral density (PSD). Trials from both conditions were grouped and individual PSDs were normalized with an element-wise division by the average power at each frequency, and the obtained values were log-transformed. To identify components of stimulus-related changes in the PSD, a principal component method is applied. This consists of calculating the covariance matrix between the frequencies. The eigenvectors of this decomposition are called Principal Spectral Components (PSCs), and reveal distinct components of neural processing, hence enabling us to identify stimulus-related changes in the PSD. Afterward, the time series were z-scored per trial to get intuitive units, then exponentiated and subtracted by 1. Finally, a baseline correction was performed by subtracting the mean value of the pre-stimulus period of -100 to 0 ms.

Both for the ERP and BB signals some electrodes were excluded from further analysis. This was done because the signal was absent or clearly erroneous. Electrode 18 in Fr was excluded from the ERP analysis, while electrodes 18 in Fr, 30, 44, 45 in Go, and 30 in Kr were excluded from the BB analysis.

Mutual Information analyses

To quantify the MI between the stimulus class and the ECoG signal (both ERP and BB), the GCMI toolbox (Gaussian Copula Mutual Information) (Ince et al., 2017) was used. This toolbox calculates the MI based on the Gaussian copula the raw ERP or BB data transforms to. The approach combined a permutation test with 1000 permutations with a method of maximum statistics to correct for multiple comparisons. Using all available trials, the signal at every time point was permuted 1000 times for each electrode, randomly assigning the stimulus class labels each time. The maximum value at each time point was taken, and the 95th percentile of this value was used as the threshold for significance. This method corrects for multiple comparisons and provides a Family-Wise Error Rate (FWER) of 0.05. Electrodes with significant mutual information between standard and deviant trials were selected as electrodes of interest, and the co-information between them was estimated for the ERP and broadband signals separately.

Co-information analyses

We quantified co-information (co-I) within signals (single electrodes) and between signals (between pairs of electrodes) using the GCMI toolbox (Ince et al., 2017). The co-I was calculated by comparing signals on a trial-by-trial basis. This resulted in a quantification of the information content, redundant or synergistic, between the two signals. The co-information (co-I) was calculated in the following way:

$$coI(X;Y;S) = I(X;S) + I(Y;S) - I(X,Y;S)$$

For each time point, $I(X;S)$ corresponds to the mutual information (MI) between the signal at recording site X and stimuli class S. $I(Y;S)$ corresponds to the MI between the signal at recording site Y and stimuli class S. Finally, $I(X,Y;S)$ corresponds to the MI between stimuli class S combining signals from recording sites X and Y.

For each neural marker of auditory PE (i.e., ERP and BB), co-information was computed for each pair of tones (standard and deviants) within recording sites in A1 and frontal regions (Figures 3, 5 and Figures S1-S2), and between A1 and frontal regions (Figures 4, 6 and Figures S3-S8,). Positive co-information shows that signals between recording sites contain redundant, or overlapping, information about the stimuli. Negative co-information corresponds to the synergy between the two variables: the information when considering the two variables jointly is larger than when considering the variables separately.

Figure 1C shows a schematic representation of co-I between two signals. It shows the independent information that response 1 and response 2 (both in white) contain. If there is an overlap in the information that is being represented by the two signals, there is a redundancy (red color) in the information that the two responses contain. If the two signals considered together contain more information than could be expected based on the information present in the individual signals, there is synergy (blue color). Statistical analyses of co-I charts were performed by using a permutation test with 1000 permutations and using the same maximum statistics method described for the MI analyses, resulting in an FWER of 0.05.

Note that MI and co-I values are reported in units of bits. A value of 1 bit corresponds to a halving of uncertainty of the trial state when observing the neural response. It is important to keep in mind though that these information values are the average per sample. Here we use a sampling rate of 500 Hz,

so a value of 0.01 bits/sample corresponds to an approximate information rate of 5 bits/second.

Neurocomputational experiments

Model architecture and function

To investigate the neural mechanisms underlying the generation of the PE responses observed experimentally in the marmoset brain during oddball presentation of auditory tones we took an existing six-layer-deep, neural network architecture (Garagnani and Pulvermüller, 2013; Garagnani et al., 2008; Schomers et al., 2017) closely mimicking neuroanatomy and neurophysiology of six perisylvian areas in the left hemisphere of the human brain involved in spoken language and auditory processing, and adapted it for the present study's needs.

This choice was motivated by the observation that, like other non-human primates, marmosets are known to be highly vocal and exhibit active vocal communication among conspecifics (Fukushima et al., 2019; Miller et al., 2009; Takahashi et al., 2013); furthermore, the existing architecture has been previously used to simulate and explain well-documented neurophysiological patterns of event-related potentials observed during language processing and oddball stimulation with familiar and unfamiliar sounds (Garagnani et al., 2008; Garagnani and Pulvermüller, 2011). The model closely reflects functional and structural features of the mammalian cortex, and incorporates the following neurobiological and neurophysiological constraints:

1) Six cortical areas are modelled, three in the superior temporal and three in the inferior frontal lobes, constituting the marmoset homologues of Brodmann Areas (BAs) 41 (labelled A1 in Fig. 1E), 42 (labelled AB), and 22 (PB) in the superior temporal gyrus, and of BAs 44 and 45 (labelled PF), 6V (PM), and 4 (M1) in the inferior frontal gyrus in humans;

2) Between-area links in the model (green, purple, and black arrows in Figure 1E) reflect known neuroanatomical links between corresponding brain areas in the marmoset (see next section below); recurrent (within-area) connections (golden arrows) are also modelled, in line with known properties of the mammalian cortex (Douglas and Martin, 2004; Braitenberg and Schüz, 1998);

3) Between- and within-area links do not implement *all-to-all* connectivity between cells, but sparse, patchy, and topographic projections, with synaptic links established probabilistically (the probability of two cells being connected decreasing with the distance; see (Kaas, 1997; Amir et al., 1993; Braitenberg and Schüz, 1998) and initialised to weak and random efficacy values;

4) Local lateral inhibition (Eysel et al., 1987; Yuille and Geiger, 1998) and area-specific global regulation mechanisms (referred to as local and global inhibition, respectively) (Yuille and Geiger, 1998; Palm et al., 2014; Braitenberg and Schüz, 1998);

5) Single cells' neurophysiological dynamics, including sigmoid transformation of membrane potentials into neuronal outputs, as well as adaptation and temporal summation of inputs (Matthews, 2000);

6) Constant presence of uniform uncorrelated white noise (simulating spontaneous baseline neuronal firing) in all model neurons (Rolls and Deco, 2010).

A first difference from the human language cortex is that the location of the marmoset homologue of BA 44 – one of the major components of Broca's area (Petrides, 2013) – still has not been definitively agreed upon (Fukushima et al., 2019). However, area 6Vb in the marmoset – which, like in man and macaque, is just caudal to 45 – exhibits cytoarchitectonic features (a scattered, agranular layer 4) that make it a potential candidate for the BA 44 homologue (Fukushima et al., 2019). In addition, area 6Vb shows a pattern of neuroanatomical connectivity different from that of its dorsal (and more caudal) counterpart 6Va, a premotor area (Burman et al., 2015). In previous "human" versions of the architecture, area PF (modelling prefrontal cortex) was defined as including mainly BA45 (and 46v), whereas BA 44 was subsumed by model area PM. Given the above, and the fact that BA 44 is generally considered a prefrontal cortex area, here we decided to treat Marmoset's area 6Vb as the homologue of the insofar missing BA 44, and to label *both* 45 and 6Vb as model area PF (hence limiting PM to include just area 6Va, homologue of BA 6V).

Structurally, each model area consists of two neuronal layers, one of excitatory and one of inhibitory cells, each containing 625 (25x25) cells (see Figure 1E, schematic on the right). Functionally, model cells are graded-response units, each representing a cluster of excitatory pyramidal cells or inhibitory interneurons. The specifics of the computational implementation (including the within-area structure and single-cell functional features) are analogous to those implemented in previously published versions of the architecture (for details, see e.g., Garagnani and Pulvermüller (2011, 2013) and can be found in the Supplementary Text for completeness.

A second crucial aspect that distinguishes human, macaque, and marmoset brains is the structural connectivity between the relevant homologue areas. Our present approach, which builds upon and is in line with several previous studies carried out with this neurocomputational architecture (Garagnani and Pulvermüller, 2011, 2013; Pulvermüller and Garagnani, 2014; Schomers et al., 2017; Garagnani et al., 2008; Henningsen-Schomers and Pulvermüller, 2022; Tomasello et al., 2019; Pulvermüller et al., 2021), is to implement a fully brain constrained model. More precisely, we impose that, for any two model areas, synaptic projections between them are realised only if experimental evidence indicates the presence of neuroanatomical links between the two corresponding cortical areas in the marmoset brain. In the following section, we provide such evidence and the rationale based on which the present network architecture (shown in Figure 1E) was adopted.

Connectivity of the simulated brain areas

The implemented model areas can be thought of as grouped into two sub-systems (frontal and temporal), each simulating a hierarchy of three cortical areas consisting of a primary cortex (motor and auditory, respectively), the adjacent higher secondary, and associative multimodal regions. Neuroanatomical

studies in the mammalian brain indicate that adjacent cortical areas tend to be reciprocally connected (Pandya and Yeterian, 1985; Young et al., 1994). We implemented such next-neighbor connections (black arrows in the network architecture shown in Fig. 1E) in each of the two subsystems based on known evidence from nonhuman primates (including marmosets): within the frontal / motor (PF–PM–M1) (Pandya and Yeterian, 1985; Burman et al., 2015) and within the temporal / auditory (A1–AB–PB) (Kaas and Hackett, 2000; Pandya, 1995; Rauschecker and Tian, 2000; de la Mothe et al., 2006, 2012; Reser et al., 2009). The links connecting the parabelt (area PB) with prefrontal cortex (PF), shown in purple in Figure 1E, are also realised in line with evidence on known long-distance cortico-cortical white matter fibres in the monkey (arcuate fascicle and extreme capsule) connecting posterior-lateral parts of the temporal cortex (area PB) and inferior prefrontal cortex (area PF) (Petrides and Pandya, 2002; Petrides et al., 2012a; Rilling, 2014; Rilling et al., 2008; Romanski et al., 1999b; Petrides et al., 2012b; Burman et al., 2015). Finally, although less strong and richly developed than in humans (Rilling et al., 2012; Thiebaut de Schotten et al., 2012), the presence of higher-order "jumping" connections between non-adjacent areas in the model (green arrows in Figure 1E) has been documented also in monkeys (including in marmosets). Specifically, neuroanatomical studies indicate that A1 is directly connected to PB (Pandya and Yeterian, 1985; Scott et al., 2017; de la Mothe et al., 2006, 2012; Reser et al., 2009), that AB is connected to PF (Romanski et al., 1999a; Kaas and Hackett, 2000; Petrides and Pandya, 2009; Rauschecker and Scott, 2009; Smiley et al., 2007), that PB and PM are linked (Rilling et al., 2008; Suzuki et al., 2015) and that PF – here including areas 45 and 6Vb – is also directly connected to M1 (Deacon, 1992; Schmahmann et al., 2009).

A previous modelling study using a neurocomputational architecture analogous to the present one looked at the effects of qualitative and quantitative differences between monkey's and human's perisylvian areas connectivity on verbal working memory (Schomers et al., 2017). In that study, the architecture used to simulate the monkey brain did not implement any of the jumping links that we included in the present model of the marmoset's cortex. Schomers et al. (2017), however, did acknowledge that the extant evidence does not imply a "*complete absence of jumping links in nonhuman primates*". In addition, none of the neuroanatomical studies used to constrain that modelling work included data from marmoset monkeys (the evidence about non-human primate connectivity used relied on macaques or chimpanzees). Finally, as Schomers and colleagues clarified, their work focused on modelling the major structural connectivity differences between monkeys and humans rather than on modelling the full complexity of the neuroanatomical connections of either species (Schomers et al., 2017). Hence, given that – albeit weaker and less developed than in humans – jumping links in non-human primates (including in the marmoset) do appear to exist, the Schomers et al. (2017) study and the present one should not be considered as in conflict, but simply as using the same network architecture to address different computational questions.

Procedures

To simulate the Roving Task of Experiment 1, the network was repeatedly presented with stimulus patterns to its auditory cortex (area A1). A stimulus pattern (simulating an auditory tone) consisted of a pre-determined set of 31 cells chosen at random amongst the 25-by-25 cells of area A1 (5% of cells). We used 12 different randomly generated stimulus patterns; presenting a stimulus involved activating the 31 cells of the chosen pattern in A1. A single trial consisted of a baseline (ten simulation time-steps long) with no input, followed by 20 timesteps of stimulus presentation, and 20 timesteps of inter-trial interval (no input); stimulus onset asynchronicity was therefore 50 simulation time-steps. A roving paradigm was used, in which 89% of standard (STD) trials were intermixed with 11% of deviant (DEV) trials. A new DEV trial was always preceded by 6-to-10 identical STD stimuli; the new DEV stimulus was chosen at random. The network's output (firing rates of all cells of the six areas) was recorded from the start of the last STD trial to the end of the critical (DEV) trial following it. For each simulation run (a model correlate of a monkey recording) we collected a total of 50 STD and 50 DEV trials. MI and co-I analyses on the simulated data, as well as the statistical contrasts between STD and DEV tones, were performed exactly as described for the in-vivo data.

Multivariate Co-Information Method (MVCo-I)

Mutual information quantifies statistical dependence on the meaningful effect size of bits. Crucially, these values are additive when combining independent representations. This allows us to quantify representational interactions between electrodes as synergistic or redundant using co-I as described in the manuscript. However, estimating mutual information on high-dimensional responses is challenging. Multi-Variate Pattern Analysis (MVPA) is an approach that has been widely adopted in neuroimaging and neuroscience to deal with high-dimensional signals (King and Dehaene, 2014). MVPA uses techniques from the field of machine learning: namely, supervised learning algorithms and cross-validation, to learn informative patterns in high-dimensional data, and evaluate their generalisation performance (i.e. how well the model could predict in new data). Here we use linear-discriminant analysis to learn the most informative linear combination of channel activity in each region to predict the binary class of the stimulus (i.e. deviant vs standard). There are various metrics for evaluating the cross-validated predictive performance of classification algorithms, for example, overall accuracy, or measures like Area under the ROC curve (Poldrack et al., 2020). These metrics can be used to rank models based on different features (i.e. compare the amount of information in temporal vs frontal regions).

Here, we combine MVPA with information-theoretic co-information to quantify the representational interactions in predictions made from cross-validated multivariate models. To do this, we first apply MVPA in the typical way (here using MVPALight toolbox; (Treder, 2020)) using 10-fold cross-validation (CV). In a 10-fold CV, the overall dataset is randomly separated into 10 disjoint subsets. Then, a model is fit on 9

of those subsets and tested on the 10th, and this is repeated for each of the 10 subsets. Here, we take the decision value of the learned classifier (the value of the linear combination of the weights and the data, which would then be thresholded to make the classification) for each test set trial. This quantifies how strongly the informative pattern the classifier had learned was present in the data on that trial. We combine the test-set decision values from all 10 different CV repetitions and calculate the mutual information between these out-of-sample decision values and the true stimulus value on each trial (Yan et al., 2023b,a). We have used MVPA to reduce the activity of the multi-channel region into a single scalar value: the decision value (d-val).

We can repeat the MVPA analysis for each time point of the stimulus-locked epochs. Often, temporal cross-decoding (King and Dehaene, 2014) is employed with MVPA to compare the consistency of the informative patterns over time. For this method, a classifier is trained at time t, and then tested (in the hold-out test folds) at other times. If it can decode, it shows the same pattern that was learned at time t, which is information at other times. However, this can only compare between data sets or conditions that are in the same physical space: i.e. we can cross-decode across time within one brain region, but we cannot compare between two different brain regions, because there is no way to apply the linear weights learned in the frontal region to the completely different temporal electrodes. Combining MVPA with co-Information (MVCo-I) overcomes this limitation. We compute co-information between the cross-validated decision values of different classifiers. This admits the same interpretation as for the channel-wise analysis. Redundancy shows that there is common information accessed by the two decoding models. Synergy means that there is a super-additive boost in the information available when considering the pattern activation from both models together. When estimating the joint information for the co-information calculation we take the maximum of the individual region MI (because the data processing inequality tells us this is a lower bound on the information that can be extracted from the joint response), the MI from the combined d-vals (2D signal; this has the advantage of being a low dimensional response for MI calculation, and being the optimal informative signal from each region) and the MI from a joint MVPA model fit to the combination of channels from both regions (1D d-vals, but which has the possibility to include synergistic information between the regions which we want to capture with this measure) (Figure S11).

We apply this methodology here in two ways. First, we look at within-area MVCo-I. For this, we train CV classifier models separately at each time point. We then calculate the co-information between two-time points using the cross-validated decision values of the two models. Note, that a crucial difference between this and the temporal cross-decoding method is that we always use the model that is learned to optimally decode information at that time point. Cross-decoding can tell if the same pattern is informative, but we can see redundant information even when the informative pattern changes. We can then compute MVCo-I between regions in the same way.

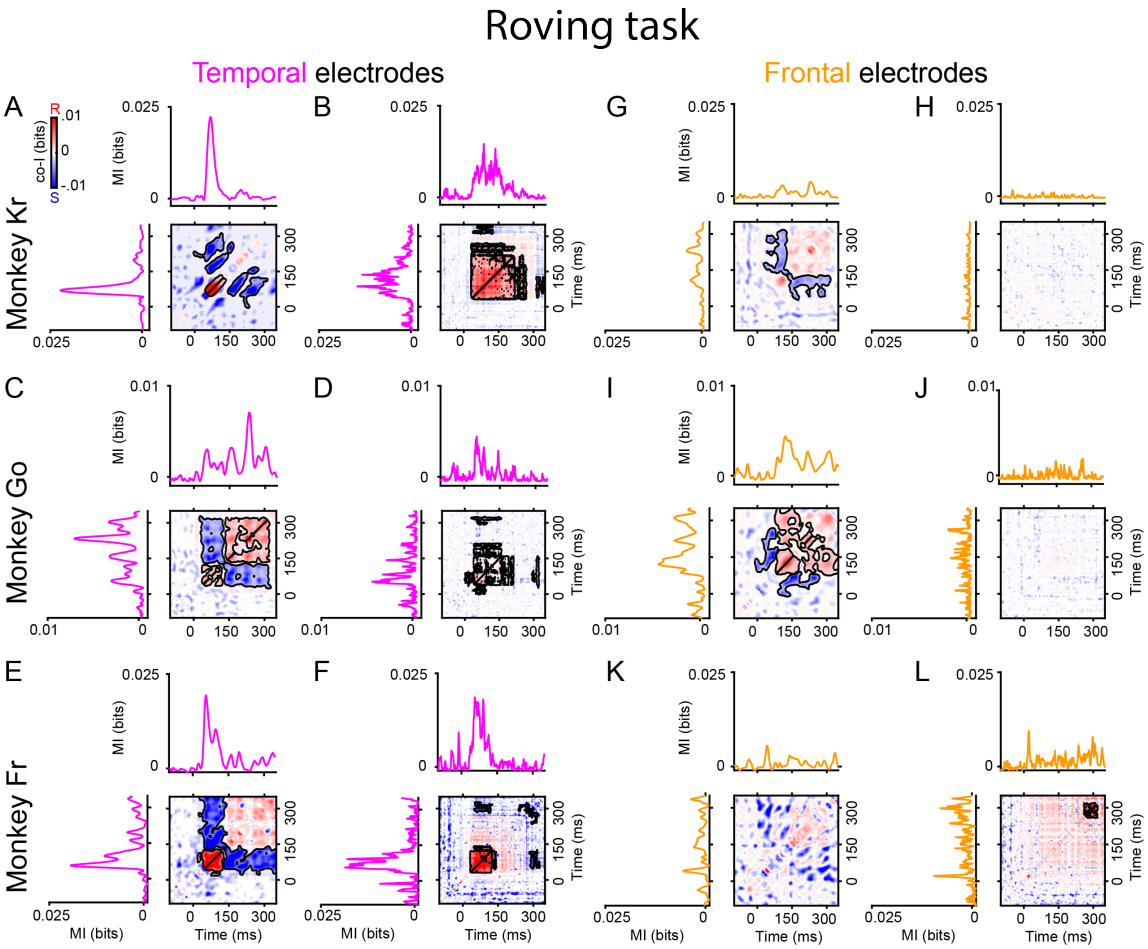

Figure S1: Synergy and redundancy within ERP and within BB signals in temporal and frontal electrodes with the highest MI for the roving task for monkeys Kr, Go and Fr. Co-information within auditory **(A, C, E)**, and frontal **(G, I, K)** electrodes in the ERP signal. Co-information within auditory **(B, D, F)**, and frontal **(H, J, L)** electrodes in the BB signal. MI (solid traces) between standard and deviant trials for temporal (pink color) and frontal (orange color) electrodes. Co-I was computed between each pair of electrodes and across time points between -100 to 350 ms after tone presentation. Significant temporal clusters after a permutation test (see Methods) are depicted in black contours.

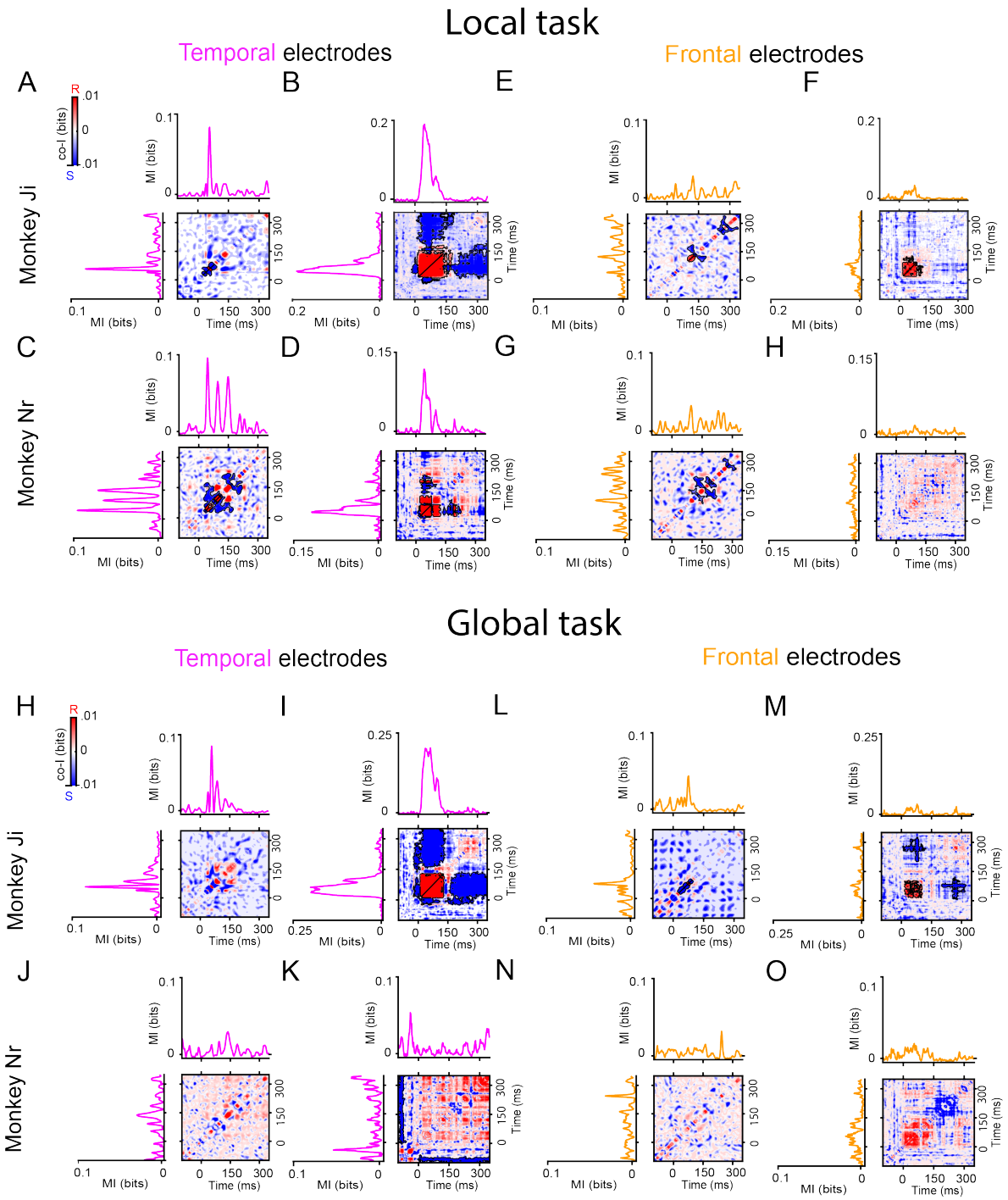

Figure S2: Synergy and redundancy within ERP and within BB signals in temporal and frontal electrodes with the highest MI for the local and global deviants of the local-global task Co-information within auditory **(A, C)**, and frontal **(E,G)** electrodes in the ERP signal for the local task. Co-information within auditory **(B,D)**, and frontal **(F,H)** electrodes in the BB signal for the local task. Co-information within auditory **(H, J)**, and frontal **(L,N)** electrodes in the ERP signal for the global task. Co-information within auditory **(I,K)**, and frontal **(M,O)** electrodes in the BB signal for the global task. MI (solid traces) between standard and deviant trials for temporal (pink color) and frontal (orange color) electrodes. Co-I was computed between each pair of electrodes and across time points between -100 to 350 ms after tone presentation. Significant temporal clusters after a permutation test (see Methods) are depicted in black contours.

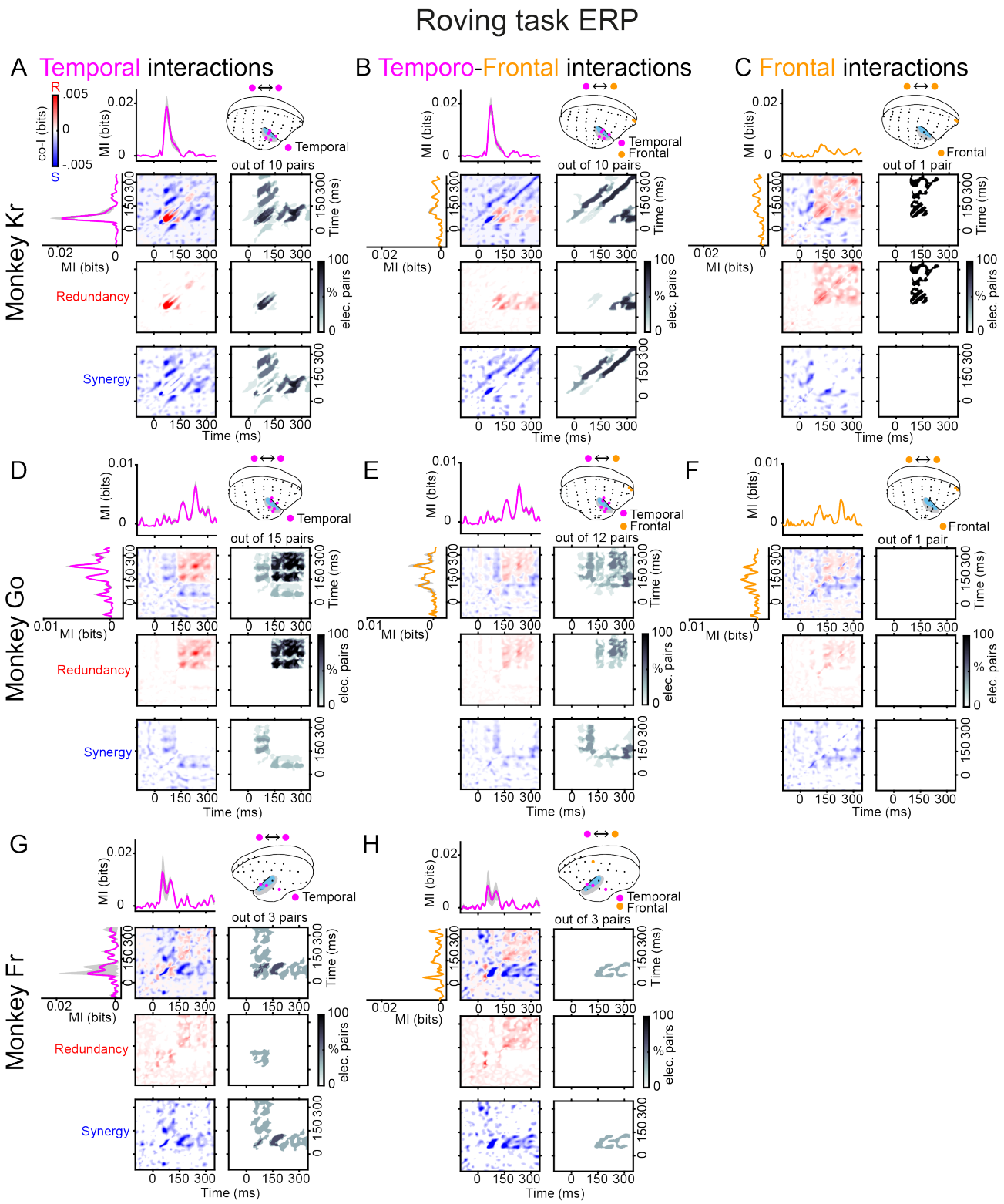

Figure S3: Synergy and redundancy between ERP signals and across cortical areas in marmosets Kr, Go and Fr. Co-information revealed synergistic and redundant PE patterns across temporal **(A, D, G)**, temporo-frontal **(B, E, H)**, and frontal **(C, F)** electrodes. MI (solid traces) between standard and deviant trials for temporal (pink color) and frontal (orange color) electrodes. Error bars represent standard error of the mean (S.E.M) across electrodes. Co-I was computed between each pair of electrodes and across time points between -100 to 350 ms after tone presentation. The average of the corresponding electrode pairs per (i.e. temporal, temporo-frontal, and frontal) is shown for the complete co-I values (red and blue panel), for positive co-I values (redundancy only; red panel), and negative co-I values (synergy only; blue panel).

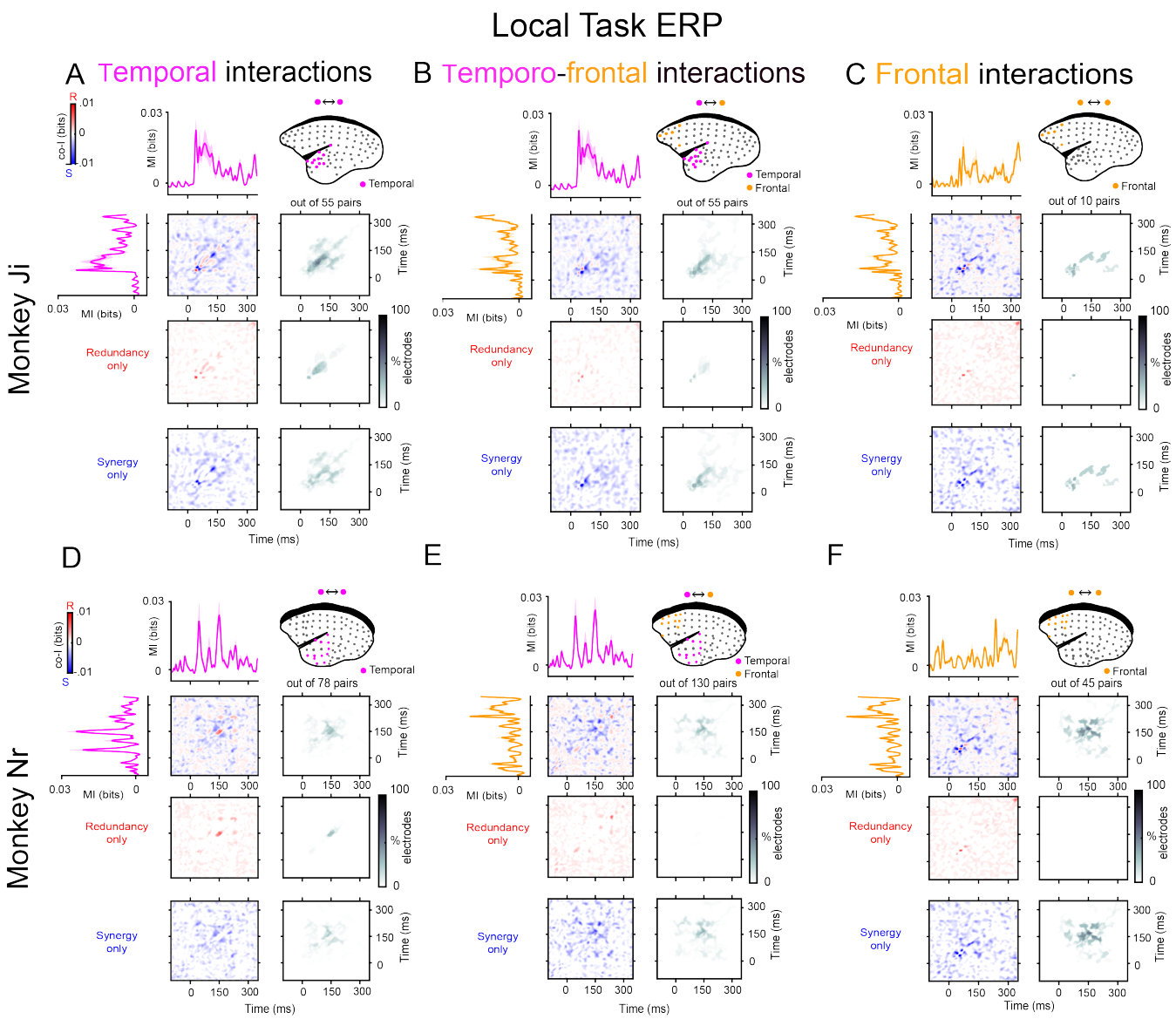

Figure S4: Synergy and redundancy between ERP signals and across cortical areas for marmosets Ji and Nr for the local deviant of the Local-Global task. Co-information revealed synergistic and redundant PE patterns across temporal **(A, D)**, temporo-frontal **(B, E)**, and frontal **(C, F)** electrodes. MI (solid traces) between standard and deviant trials for temporal (pink color) and frontal (orange color) electrodes. Error bars represent standard error of the mean (S.E.M) across electrodes. Co-I was computed between each pair of electrodes and across time points between -100 to 350 ms after tone presentation. The average of the corresponding electrode pairs per (i.e. temporal, temporo-frontal, and frontal) is shown for the complete co-I values (red and blue panel), for positive co-I values (redundancy only; red panel), and negative co-I values (synergy only; blue panel).

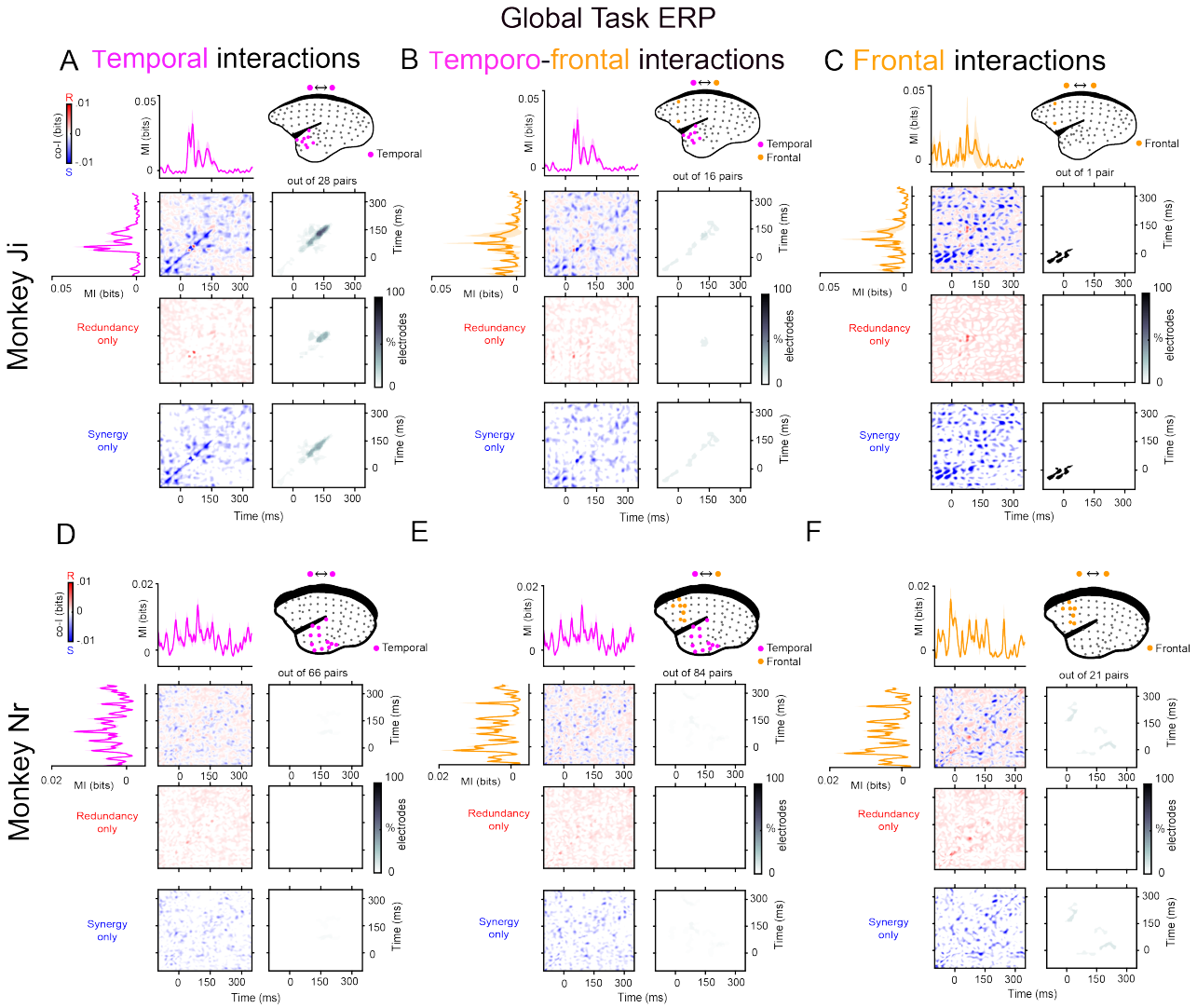

Figure S5: Synergy and redundancy between ERP signals and across cortical areas for marmosets Ji and Nr for the Global deviant of the Local-Global task. Co-information revealed synergistic and redundant PE patterns across temporal **(A, D)**, temporo-frontal **(B, E)**, and frontal **(C, F)** electrodes mostly for marmoset Ji, while Nr only showed relatively weak synergetic patterns between frontal electrodes. MI (solid traces) between standard and deviant trials for temporal (pink color) and frontal (orange color) electrodes. Error bars represent standard error of the mean (S.E.M) across electrodes. Co-I was computed between each pair of electrodes and across time points between -100 to 350 ms after tone presentation. The average of the corresponding electrode pairs per (i.e. temporal, temporo-frontal, and frontal) is shown for the complete co-I values (red and blue panel), for positive co-I values (redundancy only; red panel), and negative co-I values (synergy only; blue panel).

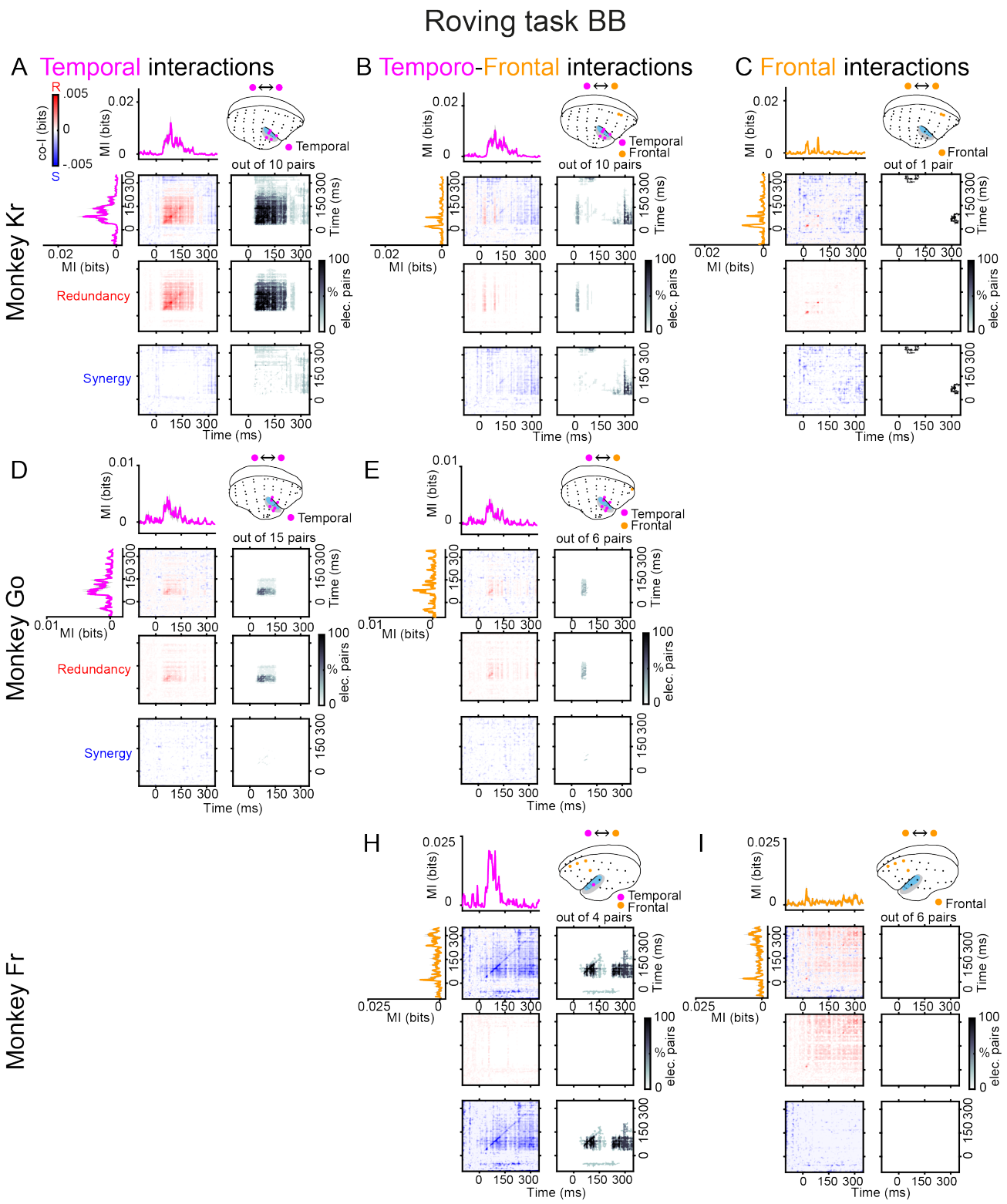

Figure S6: Synergy and redundancy between BB signals and across cortical areas in marmosets Kr, Go and Fr in the Roving task. Co-information revealed synergistic and redundant PE patterns across temporal **(A, D)**, temporo-frontal **(B, E, H)**, and frontal **(C, I)** electrodes. MI (solid traces) between standard and deviant trials for temporal (pink color) and frontal (orange color) electrodes. Error bars represent standard error of the mean (S.E.M) across electrodes. Co-I was computed between each pair of electrodes and across time points between -100 to 350 ms after tone presentation. The average of the corresponding electrode pairs per (i.e. temporal, temporo-frontal, and frontal) is shown for the complete co-I values (red and blue panel), for positive co-I values (redundancy only; red panel), and negative co-I values (synergy only; blue panel).

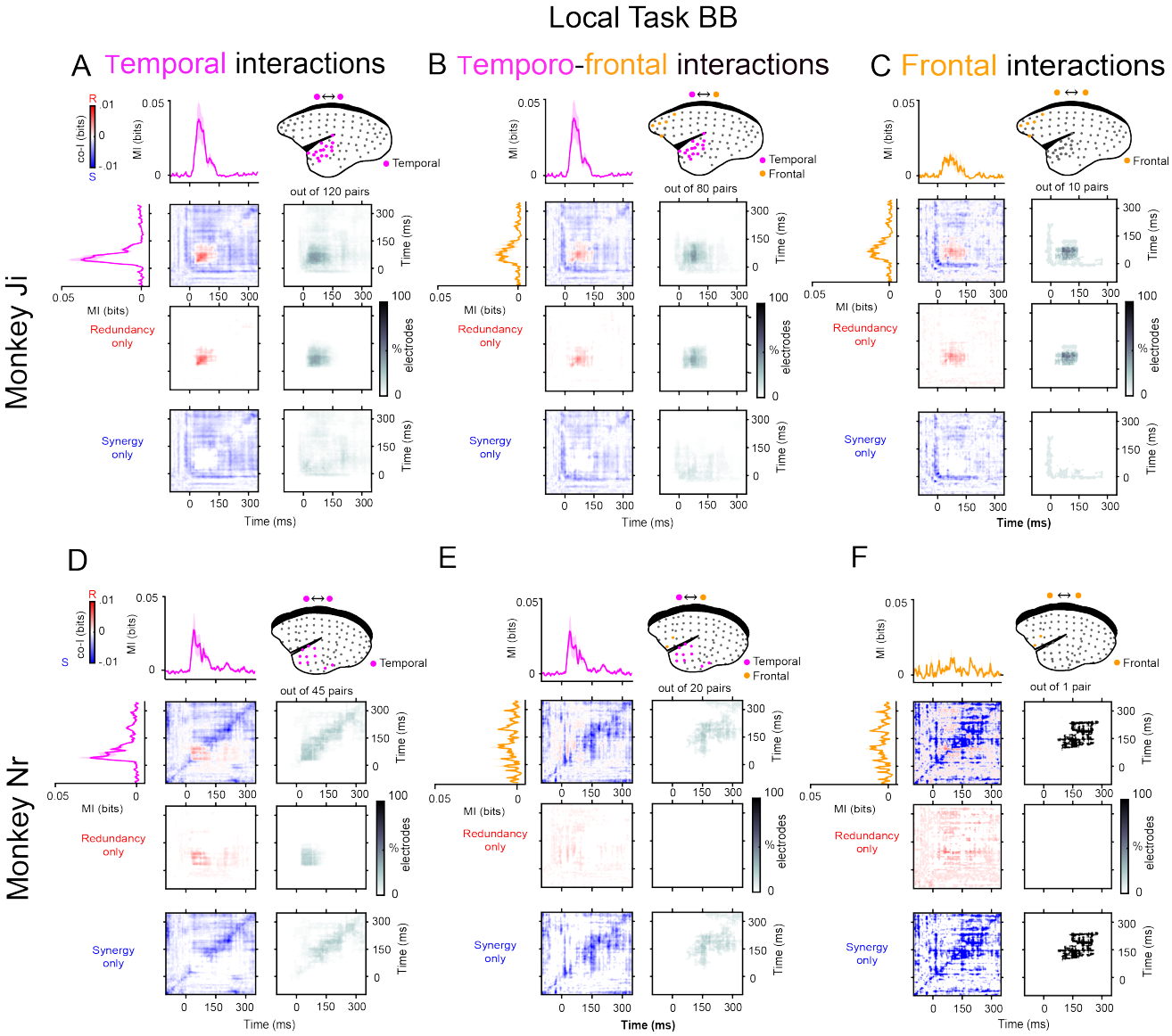

Figure S7: Synergy and redundancy between BB signals and across cortical areas for marmosets Ji and Nr for the local deviant of the Local-Global task. Co-information revealed synergistic and redundant PE patterns across temporal **(A, D)**, temporo-frontal **(B, E)**, and frontal **(C, F)** electrodes. MI (solid traces) between standard and deviant trials for temporal (pink color) and frontal (orange color) electrodes. Error bars represent standard error of the mean (S.E.M) across electrodes. Co-I was computed between each pair of electrodes and across time points between -100 to 350 ms after tone presentation. The average of the corresponding electrode pairs per (i.e. temporal, temporo-frontal, and frontal) is shown for the complete co-I values (red and blue panel), for positive co-I values (redundancy only; red panel), and negative co-I values (synergy only; blue panel).

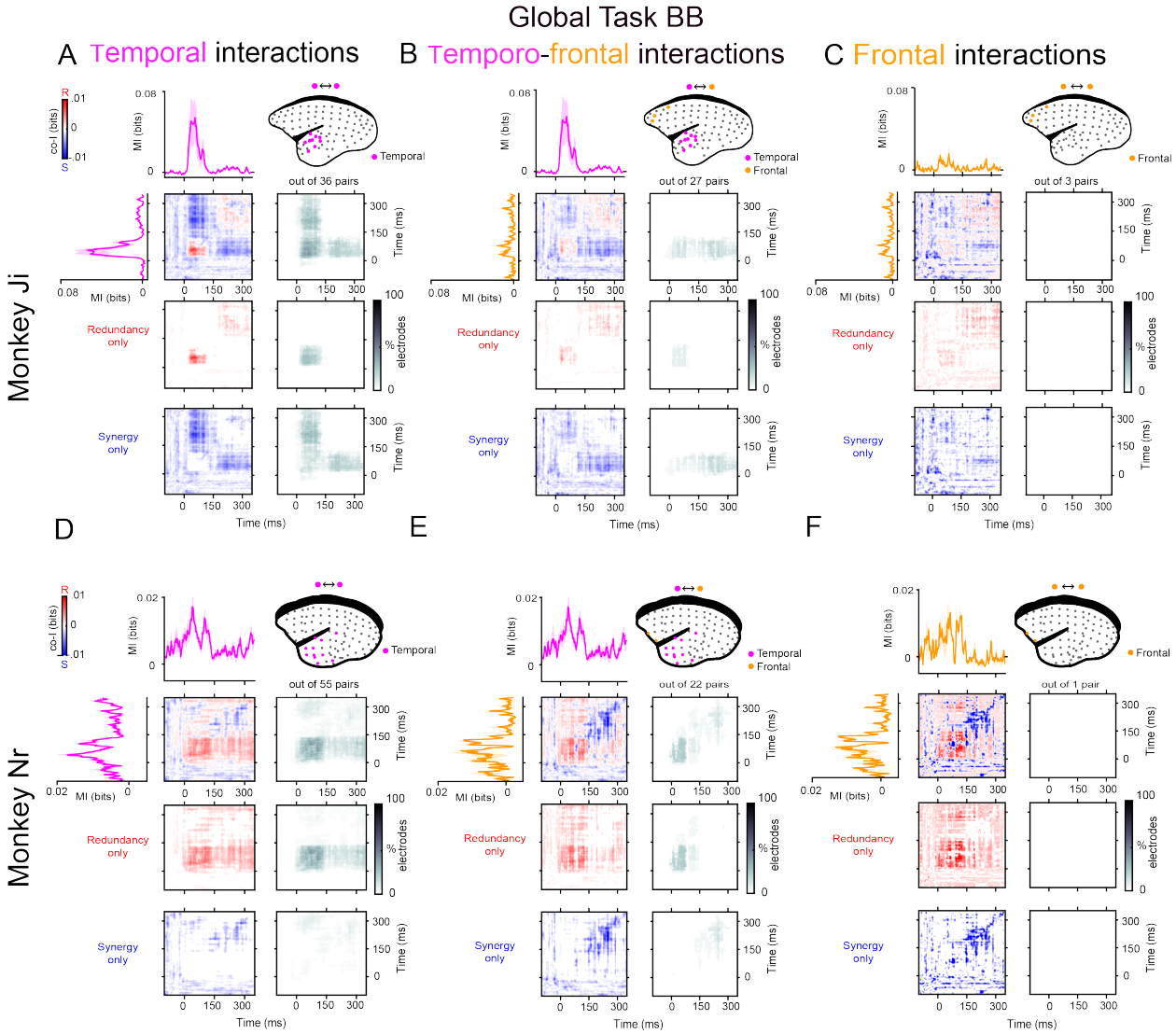

Figure S8: Synergy and redundancy between ERP signals and across cortical areas in marmosets Ji and Nr for the global deviant of the Local-Global task. Co-information revealed synergistic and redundant PE patterns across temporal **(A, D)**, temporo-frontal **(B, E)**, and frontal **(C, F)** electrodes. MI (solid traces) between standard and deviant trials for temporal (pink color) and frontal (orange color) electrodes. Error bars represent standard error of the mean (S.E.M) across electrodes. Co-I was computed between each pair of electrodes and across time points between -100 to 350 ms after tone presentation. The average of the corresponding electrode pairs per (i.e. temporal, temporo-frontal, and frontal) is shown for the complete co-I values (red and blue panel), for positive co-I values (redundancy only; red panel), and negative co-I values (synergy only; blue panel).

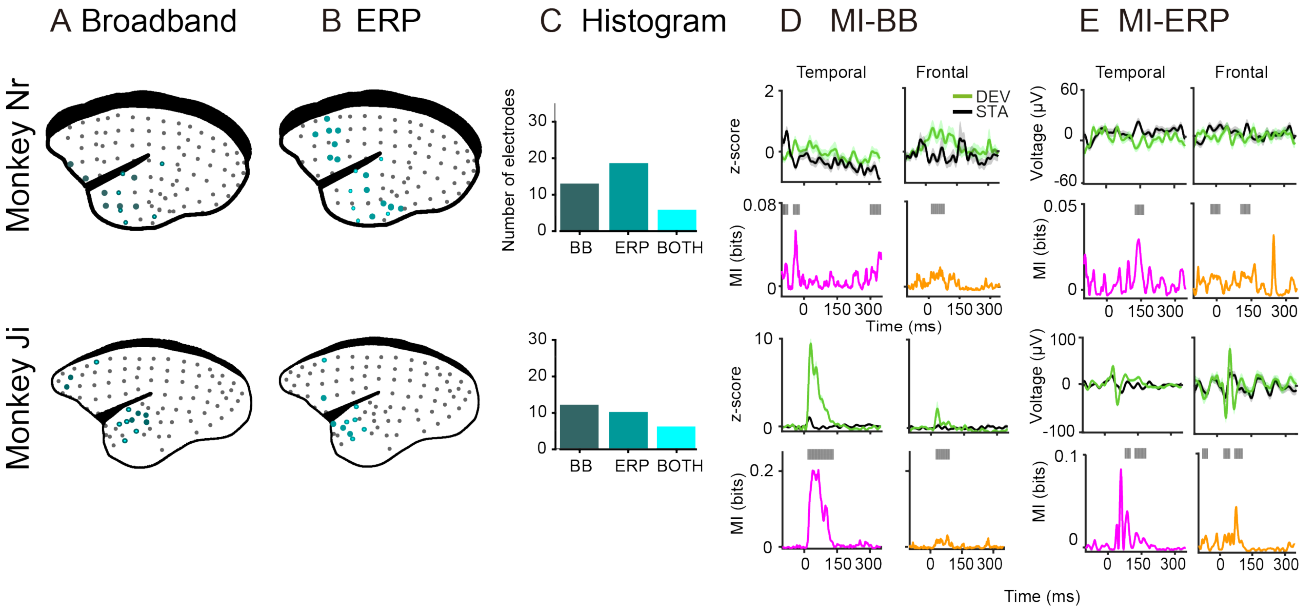

Figure S9: Broadband and ERP markers of PE across the monkey brain for monkeys Ji and Nr for the global task. Electrode locations for marmoset Nr (96 electrodes; upper panel) and Ji (96 electrodes; lower panel). Electrodes showing significant PE effect after computing MI between standard and deviant trials for the **(A)** Broadband (dark green circles) and **(B)** ERP (light green circles) markers of auditory prediction error in both monkeys. Electrodes showing significant MI for both markers are depicted in cyan. **(C)** Histogram of electrodes showing significant MI between tones for BB (left), ERP (middle), and both markers (right) for each animal. **(D)** Electrodes with the highest MI in the temporal and frontal cortex showing the BB signal for deviant and standard tones. Error bars represent standard error of the mean (S.E.M) across trials. Deviant tone (green) and standard tone (black), and the corresponding MI values in bits (effect size of the difference) for the temporal (pink trace) and frontal (orange trace) electrodes. Significant time points after a permutation test are shown as grey bars over the MI plots. **(E)** Electrodes with the highest MI in the temporal and frontal cortex showing the ERP signal for deviant and standard tones. Color codes are the same as in **C**.

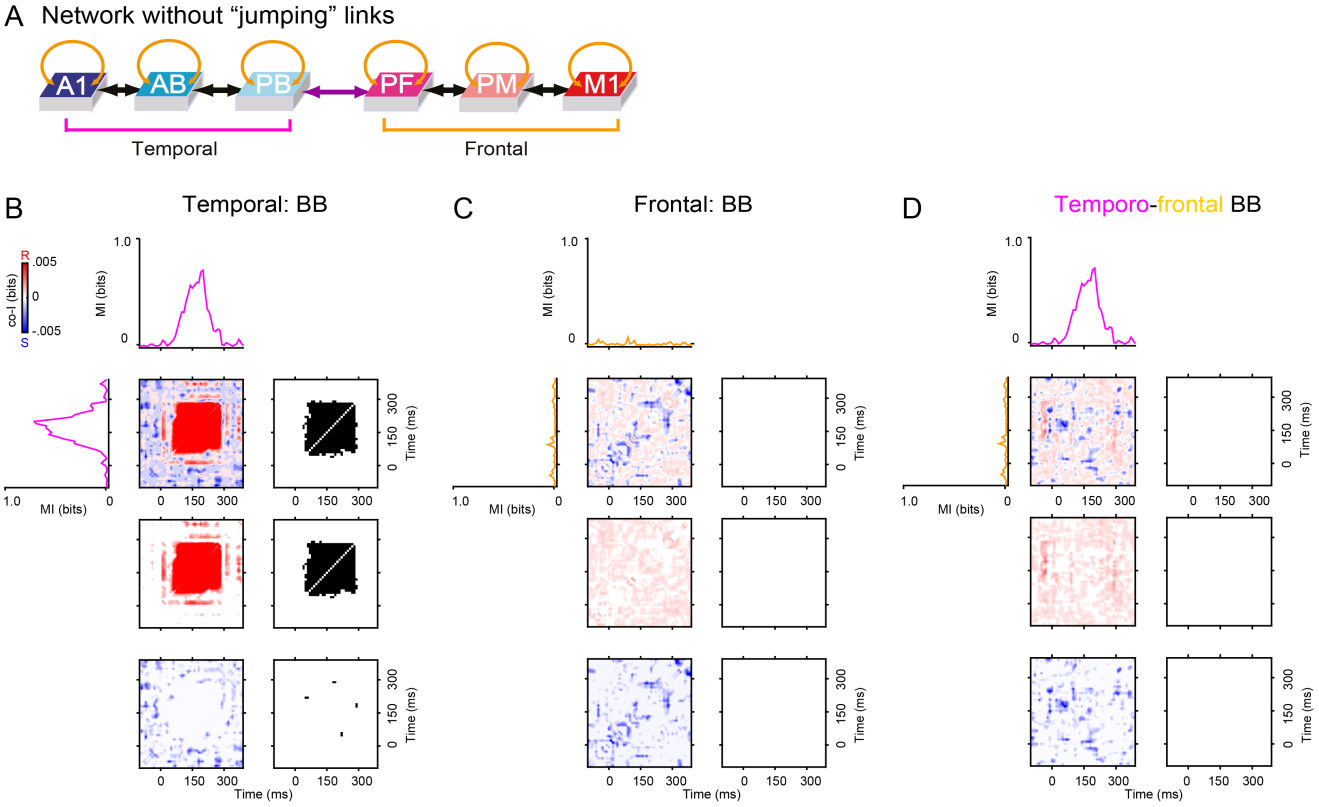

Figure S10: Results of network simulations. (**A**) Model architecture: fully connected (FC) with “serial” connectivity structure (i.e., only next-neighbour between-area links). Connections included both feedforward and feedback (black and green arrows) and recurrent (golden arrows) links, but, unlike in previous simulations, no higher-order “jumping” links (simulating cortico-cortical projections between non-adjacent areas A1-PB, AB-PF, PB-PM, and PB-M1). (**B, C, D**) Results obtained by stimulating the network of Panel **A** using a simulated Roving Task paradigm (as in Experiment 1). MI (solid traces) between standard and deviant trials are plotted for the three temporal (A1, AB, PB: pink curves) and three frontal (PF, PM, M1: orange curves) areas’ simulated responses. Co-information analyses were performed between the simulated temporal and frontal areas’ signals. Temporal co-I was computed from the simulated firing rates across time points between -100 to 350 ms after stimulus onset. The average of the corresponding electrodes for the simulated responses is shown for the complete co-I chart (red and blue panel), for positive co-I values (redundancy only; red panel), and for negative co-I values (synergy only; blue panel). The grey-scale panels show significant clusters of co-I for areas with the highest MI. Note the almost entire absence of synergistic (or redundant) information between temporal and frontal regions (see panel **D**).

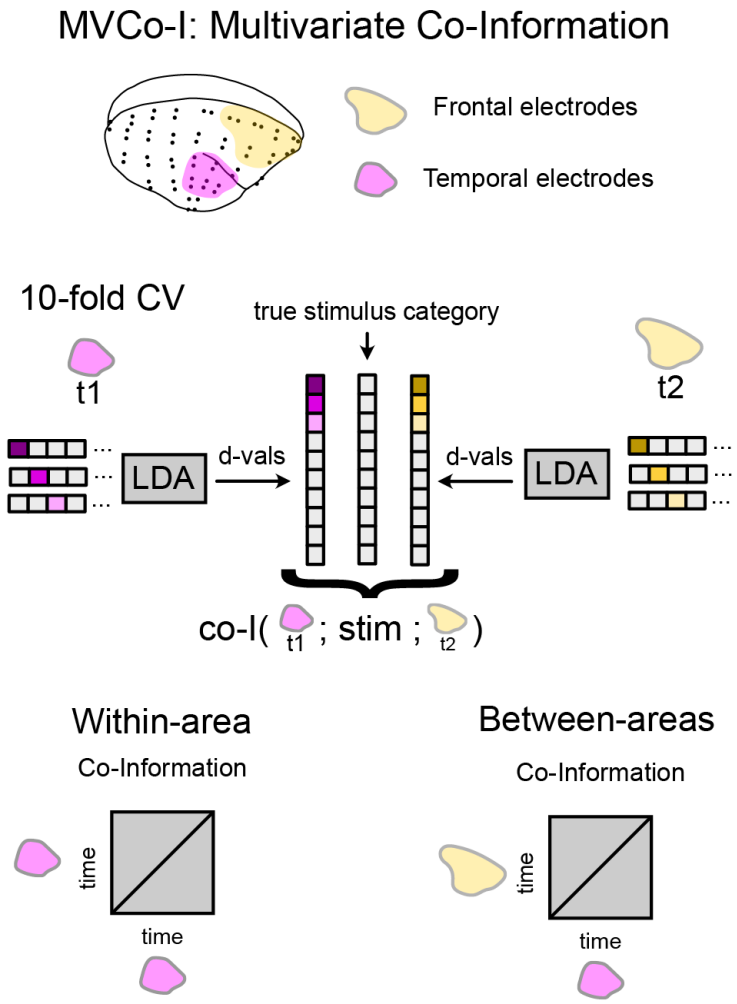

Figure S11: Schematic of the MVCo-I method. To calculate co-information between multivariate responses we use 10-fold cross-validation (CV) with Linear Discriminant Analysis (LDA). For each of the 10 hold-out folds, we train an LDA classifier to discriminate the stimulus category of a trial (i.e. deviant vs standard tone). We then compute the classifier decision values (d-vals; i.e. the linear combination of the learned pattern weights and the raw data) on each trial for the hold-out fold. We then concatenate the CV d-vals from all folds. We have used LDA as a cross-validated supervised dimensionality reduction method to obtain a one-dimensional representation of the region's activity at that time point, which is maximally discriminative between conditions. We can then compute the co-information between the ground-truth stimulus class of each trial, and the CV d-vals from two different classifiers (i.e. either from the same region at different time points, within-area coI, or from different regions, between-areas co-I). This calculation is the same as for channel-wise analysis (see Methods).

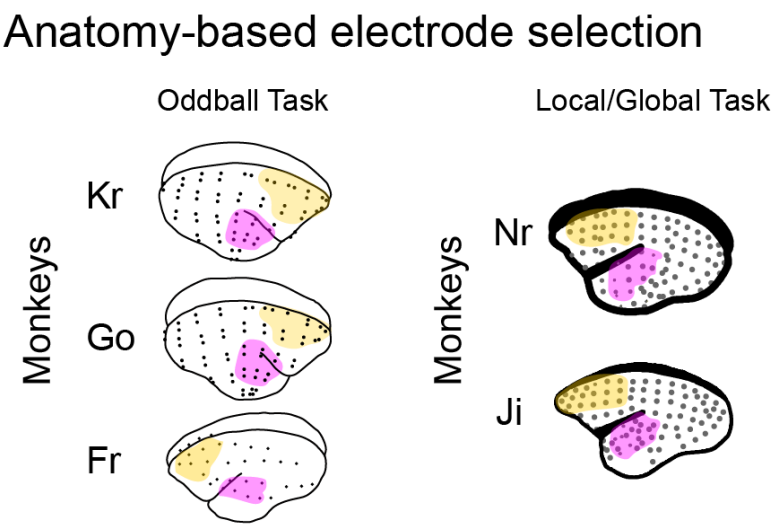

Figure S12: Anatomy-based electrode selection for the MVCo-I Method per marmoset. Roving Oddball Task (3 marmosets: Kr, Go and Fr) and Local/Global Task (2 marmosets: Ji and Nr).

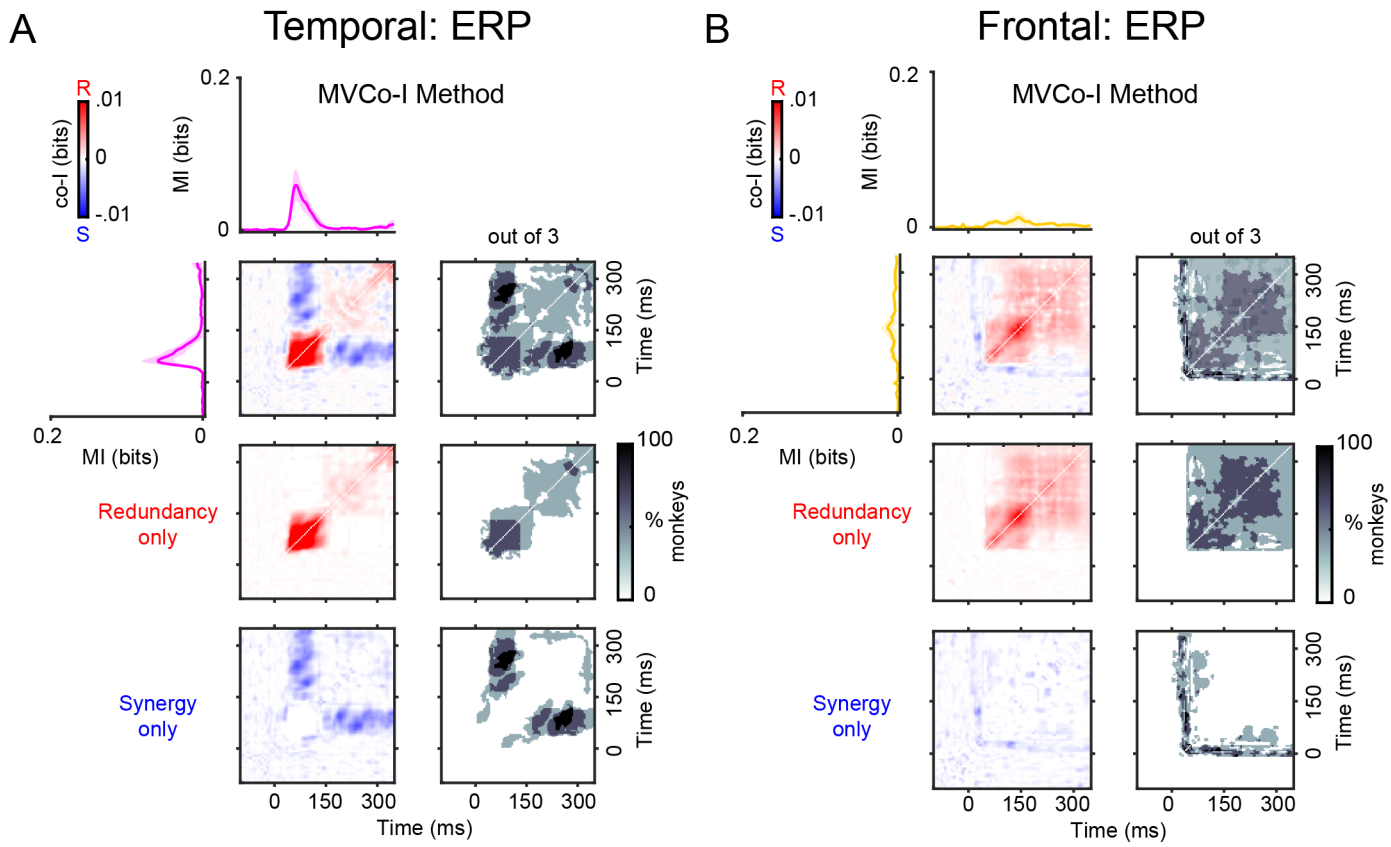

Figure S13: Temporal synergy and redundancy within ERP signals in the auditory and frontal electrodes using the MVCo-I Method (Experiment 1: Roving Oddball Task). MVCo-I revealed synergistic and redundant temporal patterns within Temporal ERP (Panel **A**) and Frontal ERP (Panel **B**) signals in the auditory cortex. MI (solid traces) between standard and deviant trials for auditory (pink color) and frontal (orange color) responses averaged across the three monkeys. The corresponding electrodes used for the MVCo-I method are depicted in 1B. Error bars represent standard error of the mean (S.E.M). Temporal co-I was computed within the corresponding signal (ERP) across time points between -100 to 350 ms after tone presentation. The average of the corresponding electrodes across monkeys is shown for the complete co-I chart (red and blue plots); for positive co-I values (redundancy only; red panel); and for negative co-I values (synergy only; blue plot).

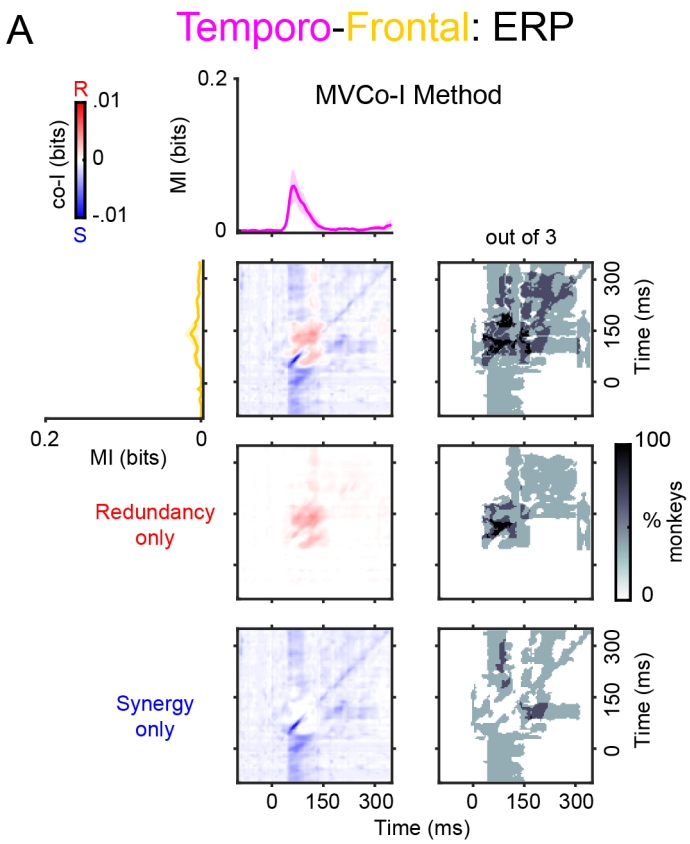

Figure S14: Spatio-temporal synergy and redundancy between auditory and frontal ERP signals using the MVCo-I Method (Experiment 1: Roving Oddball Task). (**A**) MVCo-I revealed synergistic and redundant temporal patterns between temporal and frontal signals. MI (solid traces) between standard and deviant trials for auditory (pink color) and frontal (orange color) responses averaged across the three monkeys. The corresponding electrodes used for the MVCo-I method are depicted in 1B. Error bars represent standard error of the mean (S.E.M). Temporal co-I was computed within the corresponding signal (ERP) across time points between -100 to 350 ms after tone presentation. The average of the corresponding electrodes across monkeys is shown for the complete co-I chart (red and blue plots); for positive co-I values (redundancy only; red panel); and for negative co-I values (synergy only; blue plot).

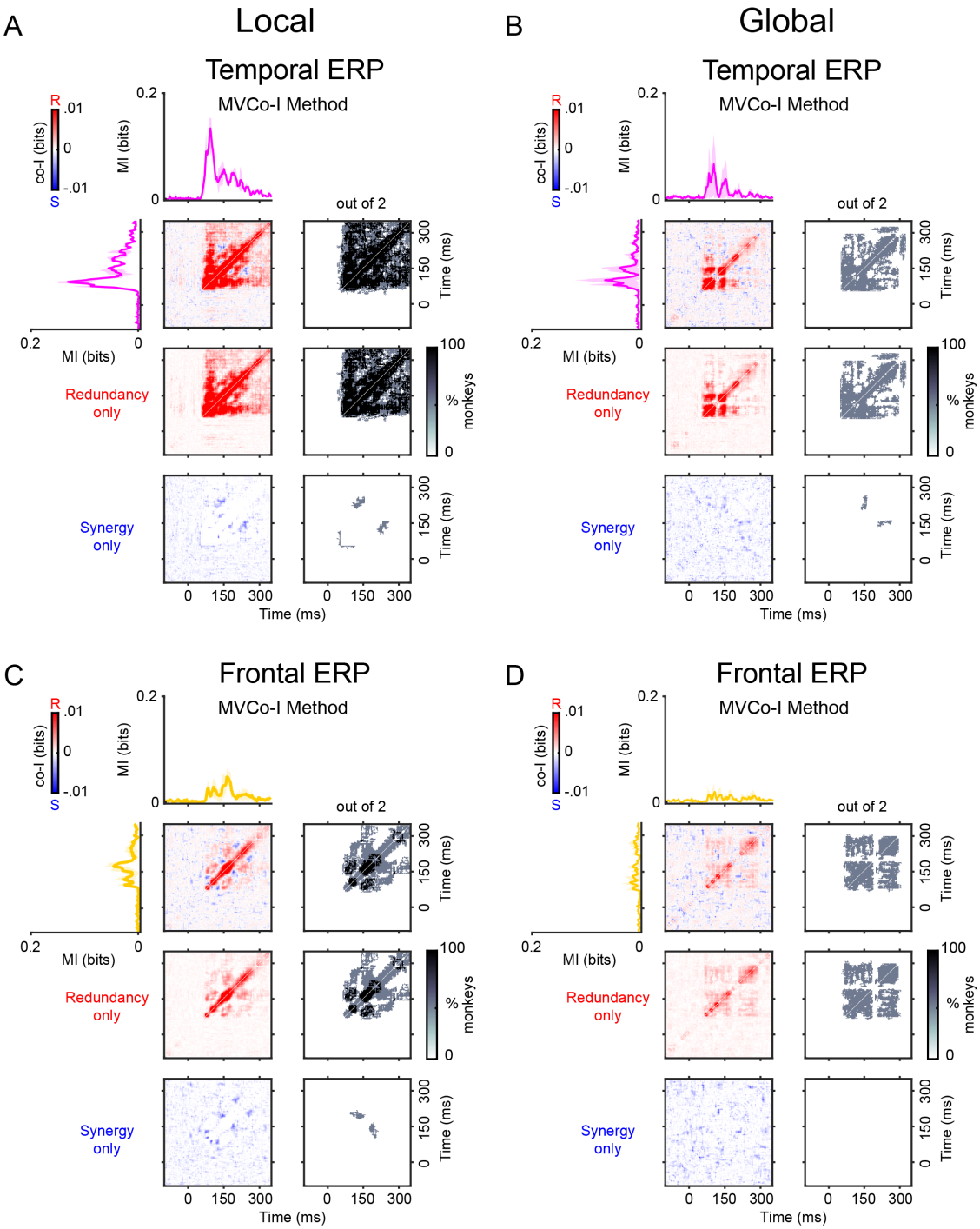

Figure S15: Temporal synergy and redundancy within ERP signals in the auditory and frontal electrodes using the MVCo-I Method (Experiment 2: Local/Global Task). MVCo-I revealed synergistic and redundant temporal patterns within auditory ERP signals in the Local (Panel **A**) and Global (Panel **B**) contrasts; and within frontal ERP signals in the Local (Panel **C**) and Global (Panel **D**) contrasts. MI (solid traces) between standard and deviant trials for auditory (pink color) and frontal (orange color) responses averaged across the two monkeys. The corresponding electrodes used for the MVCo-I method are depicted in 1B. Error bars represent standard error of the mean (S.E.M). Temporal co-I was computed within the corresponding signal (ERP) across time points between -100 to 350 ms after tone presentation. The average of the corresponding electrodes across monkeys is shown for the complete co-I chart (red and blue plots); for positive co-I values (redundancy only; red panel); and for negative co-I values (synergy only; blue plot).

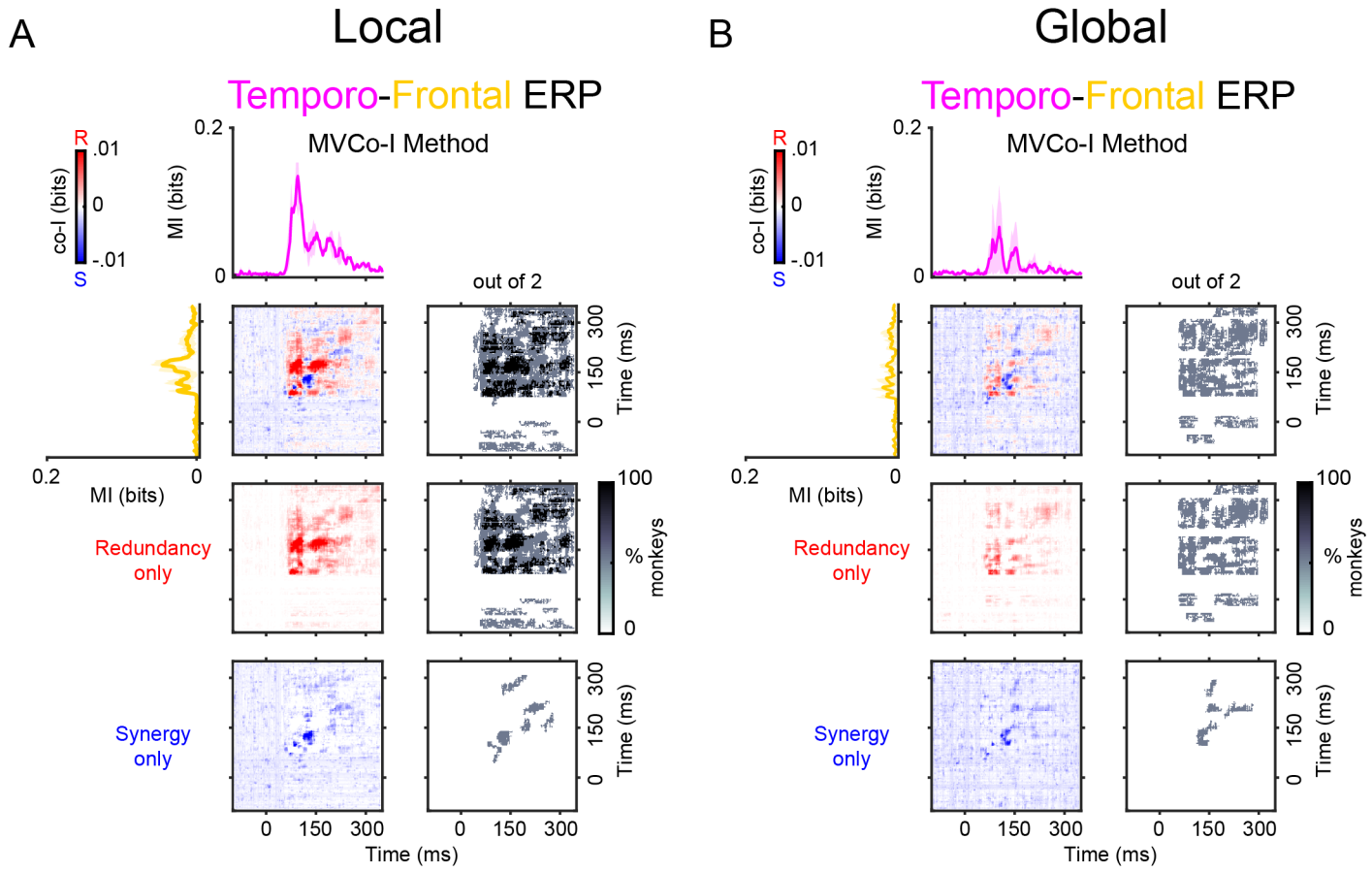

Figure S16: Spatio-temporal synergy and redundancy between temporal and frontal ERP signals using the MVCo-I Method (Experiment 2: Local/Global Task). MVCo-I revealed synergistic and redundant temporal patterns between auditory and frontal ERP signals in the Local (Panel **A**) and Global (Panel **B**) contrasts. MI (solid traces) between standard and deviant trials for auditory (pink color) and frontal (orange color) responses averaged across the two monkeys. The corresponding electrodes used for the MVCo-I method are depicted in 1B. Error bars represent standard error of the mean (S.E.M). Temporal co-I was computed within the corresponding signal (ERP) across time points between -100 to 350 ms after tone presentation. The average of the corresponding electrodes across monkeys is shown for the complete co-I chart (red and blue plots); for positive co-I values (redundancy only; red panel); and for negative co-I values (synergy only; blue plot).

References

Amir, Y., Harel, M., Malach, R., 1993. Cortical hierarchy reflected in the organization of intrinsic connections in macaque monkey visual cortex. Journal of Comparative Neurology 334, 19–46.

Bekinschtein, T.A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., Naccache, L., 2009. Neural signature of the conscious processing of auditory regularities. Proceedings of the National Academy of Sciences 106, 1672–1677.

Bimbard, C., Sit, T.P.H., Lebedeva, A., Reddy, C.B., Harris, K.D., Carandini, M., 2023. Behavioral origin of sound-evoked activity in mouse visual cortex. Nat. Neurosci. .

Blenkmann, A.O., Collavini, S., Lubell, J., Llorens, A., Funderud, I., Ivanovic, J., Larsson, P.G., Meling, T.R., Bekinschtein, T., Kochen, S., Endestad, T., Knight, R.T., Solbakk, A.K., 2019. Auditory deviance detection in the human insula: An intracranial eeg study. Cortex 121, 189–200.

Braitenberg, V., Schüz, A., 1998. Cortex: Statistics and Geometry of Neuronal Connectivity. Springer.

Breakspear, M., 2017. Dynamic models of large-scale brain activity. Nat. Neurosci. 20, 340–352.

Burman, K.J., Bakola, S., Richardson, K.E., Yu, H.H., Reser, D.H., Rosa, M.G., 2015. Cortical and thalamic projections to cytoarchitectural areas 6va and 8c of the marmoset monkey: connectionally distinct subdivisions of the lateral premotor cortex. Journal of Comparative Neurology 523, 1222–1247.

Canales-Johnson, A., Teixeira Borges, A.F., Komatsu, M., Fujii, N., Fahrenfort, J.J., Miller, K.J., Noreika, V., 2021. Broadband dynamics rather than frequency-specific rhythms underlie prediction error in the primate auditory cortex 41, 9374–9391.

Chao, Z.C., Takaura, K., Wang, L., Fujii, N., Dehaene, S., 2018. Large-scale cortical networks for hierarchical prediction and prediction error in the primate brain. Neuron 100, 1252–1266.e3.

Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibánez, A., Owen, A.M., Bekinschtein, T.A., 2013. Expectation and attention in hierarchical auditory prediction. Journal of Neuroscience 33, 11194–11205.

Chicharro, D., 2014. A causal perspective on the analysis of signal and noise correlations and their role in population coding. Neural Comput. 26, 999–1054.

Deacon, T.W., 1992. Cortical connections of the inferior arcuate sulcus cortex in the macaque brain. Brain research 573, 8–26.

Douglas, R.J., Martin, K.A., 2004. Neuronal circuits of the neocortex. Annual Review of Neuroscience 27, 419–451.

Eysel, U.T., Wörgötter, F., Pape, H.C., 1987. Local cortical lesions abolish lateral inhibition at direction selective cells in cat visual cortex. Experimental Brain Research 68, 606–612.

Friston, K., Kiebel, S., 2009. Predictive coding under the free-energy principle. Philos. Trans. R. Soc. Lond. B Biol. Sci. 364, 1211–1221.

Fukushima, M., Ichinohe, N., Okano, H., 2019. Chapter 3 - neuroanatomy of the marmoset, in: Marini, R., Wachtman, L., Tardif, S., Mansfield, K., Fox, J. (Eds.), The Common Marmoset in Captivity and Biomedical Research. Academic Press. American College of Laboratory Animal Medicine, pp. 43–62.

Garagnani, M., Pulvermüller, F., 2011. From sounds to words: a neurocomputational model of adaptation, inhibition and memory processes in auditory change detection. Neuroimage 54, 170–181.

Garagnani, M., Pulvermüller, F., 2013. Neuronal correlates of decisions to speak and act: Spontaneous emergence and dynamic topographies in a computational model of frontal and temporal areas. Brain Lang. 127, 75–85.

Garagnani, M., Wennekers, T., Pulvermüller, F., 2008. A neuroanatomically grounded hebbian-learning model of attention-language interactions in the human brain. Eur. J. Neurosci. 27, 492–513.

Giordano, B.L., Ince, R.A.A., Gross, J., Schyns, P.G., Panzeri, S., Kayser, C., 2017. Contributions of local speech encoding and functional connectivity to audio-visual speech perception. eLife 6, e24763.

Henningsen-Schomers, M.R., Pulvermüller, F., 2022. Modelling concrete and abstract concepts using brain-constrained deep neural networks. Psychol. Res. 86, 2533–2559.

Ince, R.A., Giordano, B.L., Kayser, C., Rousselet, G.A., Gross, J., Schyns, P.G., 2017. A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula. Human Brain Mapping 38, 1541–1573.

Ince, R.A.A., Jaworska, K., Gross, J., Panzeri, S., van Rijsbergen, N.J., Rousselet, G.A., Schyns, P.G., 2016. The deceptively simple N170 reflects network information processing mechanisms involving visual feature coding and transfer across hemispheres. Cereb. Cortex 26, 4123–4135.

Jiang, Y., Komatsu, M., Chen, Y., Xie, R., Zhang, K., Xia, Y., Gui, P., Liang, Z., Wang, L., 2022. Constructing the hierarchy of predictive auditory sequences in the marmoset brain. eLife 11, e74653.

Kaas, J.H., 1997. Topogrphic maps are fundamental to sensory processing. Brain Research Bulletin 44, 107–112.

Kaas, J.H., Hackett, T.A., 2000. Subdivisions of auditory cortex and processing streams in primates. Proceedings of the National Academy of Sciences 97, 11793–11799.

King, J.R., Dehaene, S., 2014. Characterizing the dynamics of mental representations: the temporal generalization method. Trends Cogn. Sci. 18, 203–210.

Komatsu, M., Kaneko, T., Okano, H., Ichinohe, N., 2019. Chronic implantation of whole-cortical electrocorticographic array in the common marmoset. J. Vis. Exp. .

Komatsu, M., Takaura, K., Fujii, N., 2015. Mismatch negativity in common marmosets: Whole-cortical recordings with multi-channel electrocorticograms. Sci. Rep. 5, 15006.

Lachaux, J.P., Axmacher, N., Mormann, F., Halgren, E., Crone, N.E., 2012. High-frequency neural activity and human cognition: Past, present and possible future of intracranial eeg research. Progress in Neurobiology 98, 279–301. High Frequency Oscillations in Cognition and Epilepsy.

Lohuis, M.N.O., Marchesi, P., Olcese, U., Pennartz, C., 2022. Triple dissociation of visual, auditory and motor processing in primary visual cortex. bioRxiv .

Luppi, A.I., Mediano, P.A.M., Rosas, F.E., Holland, N., Fryer, T.D., O'Brien, J.T., Rowe, J.B., Menon, D.K., Bor, D., Stamatakis, E.A., 2022. A synergistic core for human brain evolution and cognition. Nat. Neurosci. 25, 771–782.

Matthews, G.G., 2000. Neurobiology: molecules, cells and systems. Wiley-Blackwell.

McGill, W.J., 1954. Multivariate information transmission. Psychometrika 19, 97–116.

Miller, C.T., Eliades, S.J., Wang, X., 2009. Motor planning for vocal production in common marmosets. Animal Behaviour 78, 1195–1203.

Miller, K.J., 2019. A library of human electrocorticographic data and analyses. Nat. Hum. Behav. 3, 1225–1235.

de la Mothe, L.A., Blumell, S., Kajikawa, Y., Hackett, T.A., 2006. Cortical connections of the auditory cortex in marmoset monkeys: core and medial belt regions. Journal of Comparative Neurology 496, 27–71.

de la Mothe, L.A., Blumell, S., Kajikawa, Y., Hackett, T.A., 2012. Cortical connections of auditory cortex in marmoset monkeys: lateral belt and parabelt regions. The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology 295, 800–821.

Nigam, S., Pojoga, S., Dragoi, V., 2019. Synergistic coding of visual information in columnar networks. Neuron 104, 402–411.e4.

Olivares, J., Orio, P., Sadílek, V., Schmachtenberg, O., Canales-Johnson, A., 2022. Neural oscillations across olfactory regions encode odorant information in the teleost olfactory system. bioRxiv doi:10.1101/2022.04.15.488302.

O'Reilly, J.A., 2021. Roving oddball paradigm elicits sensory gating, frequency sensitivity, and long-latency response in common marmosets. IBRO Neuroscience Reports 11, 128–136.

Palm, G., Knoblauch, A., Hauser, F., Schüz, A., 2014. Cell assemblies in the cerebral cortex. Biol. Cybern. 108, 559–572.

Pandya, D.N., 1995. Anatomy of the auditory cortex. Revue neurologique 151, 486–494.

Pandya, D.N., Yeterian, E.H., 1985. Architecture and connections of cortical association areas. Association and auditory cortices , 3–61.

Panzeri, S., Moroni, M., Safaai, H., Harvey, C.D., 2022. The structures and functions of correlations in neural population codes. Nat. Rev. Neurosci. 23, 551–567.

Park, H., Ince, R.A.A., Schyns, P.G., Thut, G., Gross, J., 2018. Representational interactions during audiovisual speech entrainment: Redundancy in left posterior superior temporal gyrus and synergy in left motor cortex. PLOS Biology 16, 1–26.

Parras, G.G., Nieto-Diego, J., Carbajal, G.V., Valdés-Baizabal, C., Escera, C., Malmierca, M.S., 2017. Neurons along the auditory pathway exhibit a hierarchical organization of prediction error. Nat. Commun. 8, 2148.

Petrides, M., 2013. Neuroanatomy of language regions of the human brain.

Academic Press.
Petrides, M., Pandya, D., 2002. Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. European Journal of Neuroscience 16, 291–310.
Petrides, M., Pandya, D.N., 2009. Distinct parietal and temporal pathways to the homologues of broca's area in the monkey. PLoS biology 7, e1000170.
Petrides, M., Tomaiuolo, F., Yeterian, E.H., Pandya, D.N., 2012a. The prefrontal cortex: comparative architectonic organization in the human and the macaque monkey brains. cortex 48, 46–57.
Petrides, M., Tomaiuolo, F., Yeterian, E.H., Pandya, D.N., 2012b. The prefrontal cortex: Comparative architectonic organization in the human and the macaque monkey brains. Cortex 48, 46–57.
Poldrack, R.A., Huckins, G., Varoquaux, G., 2020. Establishment of best practices for evidence for prediction: A review. JAMA Psychiatry 77, 534–540.
Pulvermüller, F., Garagnani, M., 2014. From sensorimotor learning to memory cells in prefrontal and temporal association cortex: a neurocomputational study of disembodiment. Cortex 57, 1–21.
Pulvermüller, F., Tomasello, R., Henningsen-Schomers, M.R., Wennekers, T., 2021. Biological constraints on neural network models of cognitive function. Nat. Rev. Neurosci. 22, 488–502.
Rao, R.P., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. 2, 79–87.
Rauschecker, J.P., Scott, S.K., 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nature neuroscience 12, 718–724.
Rauschecker, J.P., Tian, B., 2000. Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proceedings of the National Academy of Sciences 97, 11800–11806.
Reser, D.H., Burman, K.J., Richardson, K.E., Spitzer, M.W., Rosa, M.G., 2009. Connections of the marmoset rostrotemporal auditory area: Express pathways for analysis of affective content in hearing. European Journal of Neuroscience 30, 578–592.
Rilling, J.K., 2014. Comparative primate neuroimaging: insights into human brain evolution. Trends in cognitive sciences 18, 46–55.
Rilling, J.K., Glasser, M.F., Jbabdi, S., Andersson, J., Preuss, T.M., 2012. Continuity, divergence, and the evolution of brain language pathways. Frontiers in Evolutionary Neuroscience 3, 11.
Rilling, J.K., Glasser, M.F., Preuss, T.M., Ma, X., Zhao, T., Hu, X., Behrens, T.E., 2008. The evolution of the arcuate fasciculus revealed with comparative dti. Nature neuroscience 11, 426–428.
Rolls, E.T., Deco, G., 2010. The Noisy Brain: Stochastic Dynamics as a Principle of Brain Function. Oxford University Press.
Romanski, L.M., Bates, J.F., Goldman-Rakic, P.S., 1999a. Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. Journal of Comparative Neurology 403, 141–157.
Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S., Rauschecker, J.P., 1999b. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nature neuroscience 2, 1131–1136.
Saleem, A.B., Diamanti, E.M., Fournier, J., Harris, K.D., Carandini, M., 2018. Coherent encoding of subjective spatial position in visual cortex and hippocampus. Nature 562, 124–127.
Schmahmann, J.D., Schmahmann, J., Pandya, D., 2009. Fiber pathways of the brain. OUP USA.
Schneidman, E., Bialek, W., Berry, 2nd, M.J., 2003. Synergy, redundancy, and independence in population codes. J. Neurosci. 23, 11539–11553.
Schomers, M.R., Garagnani, M., Pulvermüller, F., 2017. Neurocomputational consequences of evolutionary connectivity changes in perisylvian language cortex. J. Neurosci. 37, 3045–3055.
de Schotten, M.T., Forkel, S.J., 2022. The emergent properties of the connected brain. Science 378, 505–510.
Scott, B.H., Leccese, P.A., Saleem, K.S., Kikuchi, Y., Mullarkey, M.P., Fukushima, M., Mishkin, M., Saunders, R.C., 2017. Intrinsic connections of the core auditory cortical regions and rostral supratemporal plane in the macaque monkey. Cerebral cortex 27, 809–840.
Shenoy, K.V., Kao, J.C., 2021. Measurement, manipulation and modeling of brain-wide neural population dynamics. Nat. Commun. 12, 633.
Smiley, J.F., Hackett, T.A., Ulbert, I., Karmas, G., Lakatos, P., Javitt, D.C., Schroeder, C.E., 2007. Multisensory convergence in auditory cortex, i. cortical connections of the caudal superior temporal plane in macaque monkeys. Journal of Comparative Neurology 502, 894–923.
Steinmetz, N.A., Zatka-Haas, P., Carandini, M., Harris, K.D., 2019. Distributed coding of choice, action and engagement across the mouse brain. Nature 576, 266–273.
Suzuki, W., Banno, T., Miyakawa, N., Abe, H., Goda, N., Ichinohe, N., 2015. Mirror neurons in a new world monkey, common marmoset. Frontiers in neuroscience 9, 459.
Takahashi, D., Narayanan, D., Ghazanfar, A., 2013. Coupled oscillator dynamics of vocal turn-taking in monkeys. Current Biology 23, 2162–2168.
Thiebaut de Schotten, M., Dell'Acqua, F., Valabregue, R., Catani, M., 2012. Monkey to human comparative anatomy of the frontal lobe association tracts. Cortex 48, 82–96. Frontal lobes.
Tomasello, R., Garagnani, M., Wennekers, T., Pulvermüller, F., 2017. Brain connections of words, perceptions and actions: A neurobiological model of spatio-temporal semantic activation in the human cortex. Neuropsychologia 98, 111–129.
Tomasello, R., Wennekers, T., Garagnani, M., Pulvermüller, F., 2019. Visual cortex recruitment during language processing in blind individuals is explained by hebbian learning. Sci. Rep. 9, 3579.
Treder, M.S., 2020. MVPA-Light: A classification and regression toolbox for multi-dimensional data. Front. Neurosci. 14, 289.
Urai, A.E., Doiron, B., Leifer, A.M., Churchland, A.K., 2022. Large-scale neural recordings call for new insights to link brain and behavior. Nat. Neurosci. 25, 11–19.
Varley, T.F., Sporns, O., Schaffelhofer, S., Scherberger, H., Dann, B., 2023. Information-processing dynamics in neural networks of macaque cerebral cortex reflect cognitive state and behavior. Proc. Natl. Acad. Sci. U. S. A. 120, e2207677120.
Vinck, M., Uran, C., Spyropoulos, G., Onorato, I., Broggini, A.C., Schneider, M., Canales-Johnson, A., 2023. Principles of large-scale neural interactions. Neuron 111, 987–1002.
Voitov, I., Mrsic-Flogel, T.D., 2022. Cortical feedback loops bind distributed representations of working memory. Nature 608, 381–389.
Yan, Y., Zhan, J., Garrod, O., Cui, X., Ince, R.A.A., Schyns, P.G., 2023a. Neural representation strength of predicted category features biases decision behavior. Current Biology (In Press) .
Yan, Y., Zhan, J., Ince, R.A.A., Schyns, P.G., 2023b. Network communications flexibly predict visual contents that enhance representations for faster visual categorization. J. Neurosci. 43, 5391–5405.
Young, M.P., Scanneil, J.W., Burns, G.A., Blakemore, C., 1994. Analysis of connectivity: neural systems in the cerebral cortex. Reviews in the Neurosciences 5, 227–250.
Yuille, A.L., Geiger, D., 1998. Winner-Take-All Mechanisms. MIT Press, Cambridge, MA, USA. p. 1056–1060.
Zhan, J., Ince, R.A.A., van Rijsbergen, N., Schyns, P.G., 2019. Dynamic construction of reduced representations in the brain for perceptual decision behavior. Current Biology. 29, 319–326.e4.