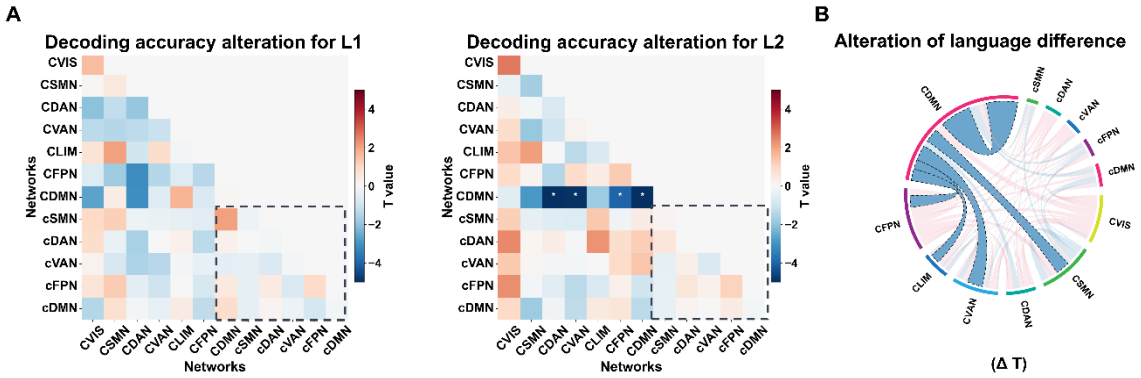
Supplementary Materials

1. Decoding accuracy of lesion of all possible network pairs

In addition to the lesion simulation of cortico-cerebellar network pairs, we performed the analysis after lesion of all possible network pairs. The results revealed that lesion of any network pairs did not lead to significant change of decoding accuracy during L1 narrative comprehension ($p < 0.05$, FDR corrected). However, the lesions between the cortical DMN and other cortical networks including the DAN, VAN, FPN as well as within the DMN resulted in dramatically decrease of decoding accuracy during L2 narrative comprehension ($p < 0.05$, FDR corrected).



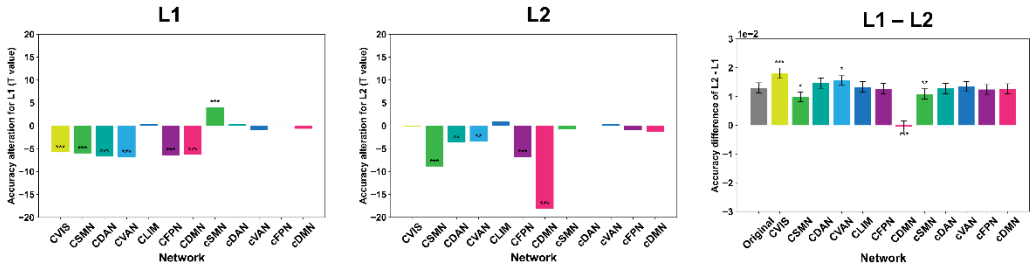
FigS1 Simulation lesions of all the network pairs. A. Decoding accuracy changes after lesion of specific network pair within languages. Black dashed lines indicated within-cerebellar network pair lesions. ***: p < 0.001, **: p < 0.01, *: p < 0.05 after FDR correction. B. Alterations between the language difference after the specific lesion and the original difference. Black dashed lines inferred significant alteration between the language difference before and after the lesions.

2. Lesion of all possible networks

In addition to the lesion simulation of cortico-cerebellar network pairs, we performed the analysis after lesion of all connectivity (not limited to the cortico-cerebellar interactions) related to a given network. After that, we compared the decoding accuracy using the remaining features with the original accuracy from the full features for different languages respectively.

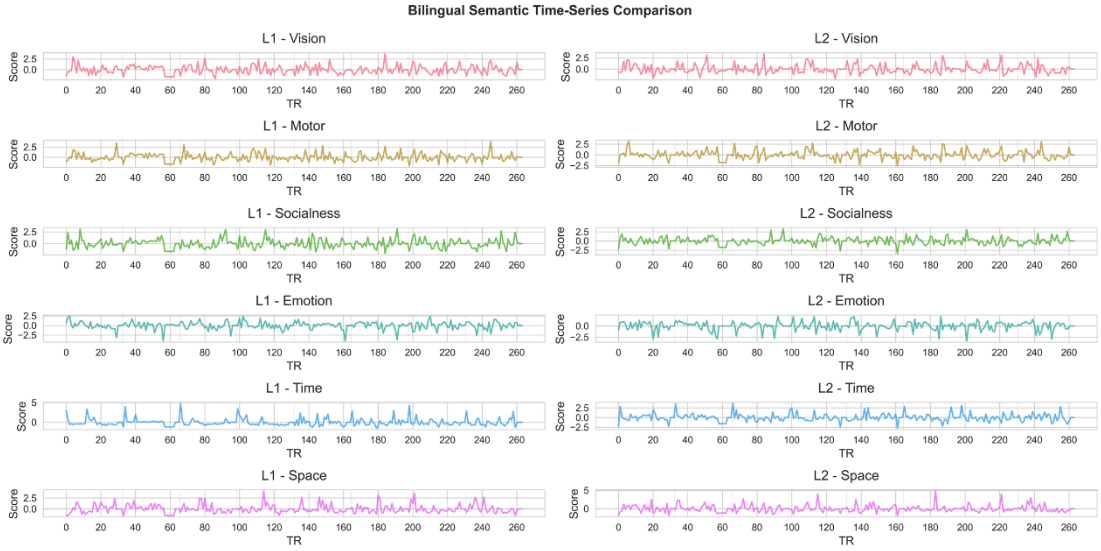
From the results, we can infer that during L1, lesion of any cortical network-related connectivity will lead to poorer performance except the Limbic network. However, lesions of the cerebellar networks did not result in accuracy alteration, except for lesion of the cerebellar SMN will result in a better decoding performance. As for the L2, lesions of the SMN, DAN, VAN, FPN and DMN will lead to poorer performance while lesions of the cortical visual network, limbic network as well as any of the cerebellar networks did affect the overall performance. Comparisons between the languages revealed that lesion of the cortical visual network will lead to larger decoding accuracy difference, while lesions of cortical SMN, VAN and cerebellar SMN will lead to smaller decoding accuracy, while the decoding accuracies were still higher during L2 as compared to L1. Last but not least, lesion of the cortical DMN will result in the largest alteration of the language difference, where the

language difference was not significantly existed.



FigS2 Network lesion. The network lesion was performed by dropping all the connectivity related to the network but not limited to the cortico-cerebellar connectivity. Left: comparison results between the decoding accuracy without lesion and after lesion for L1. Middle: comparison results between the decoding accuracy without lesion and after lesion for L2. Right: alteration of language difference (L2- L1) without lesion and after lesion. ***: $p < 0.001$, **: $p < 0.01$, *: $p < 0.05$ after FDR correction.

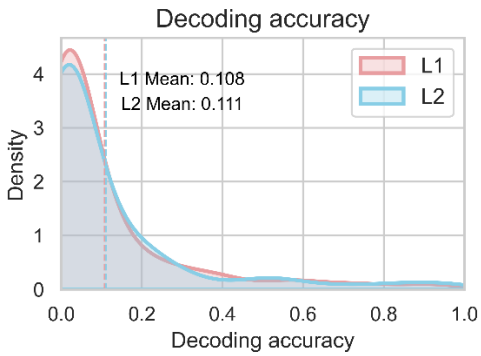
3. Coding of dynamic semantics using large corpus



FigS3 Coding of the semantic time-series. We displayed the original semantic time-series after resampling for the scanning TR without any other temporal manipulations. The coding for the two languages were performed separately.

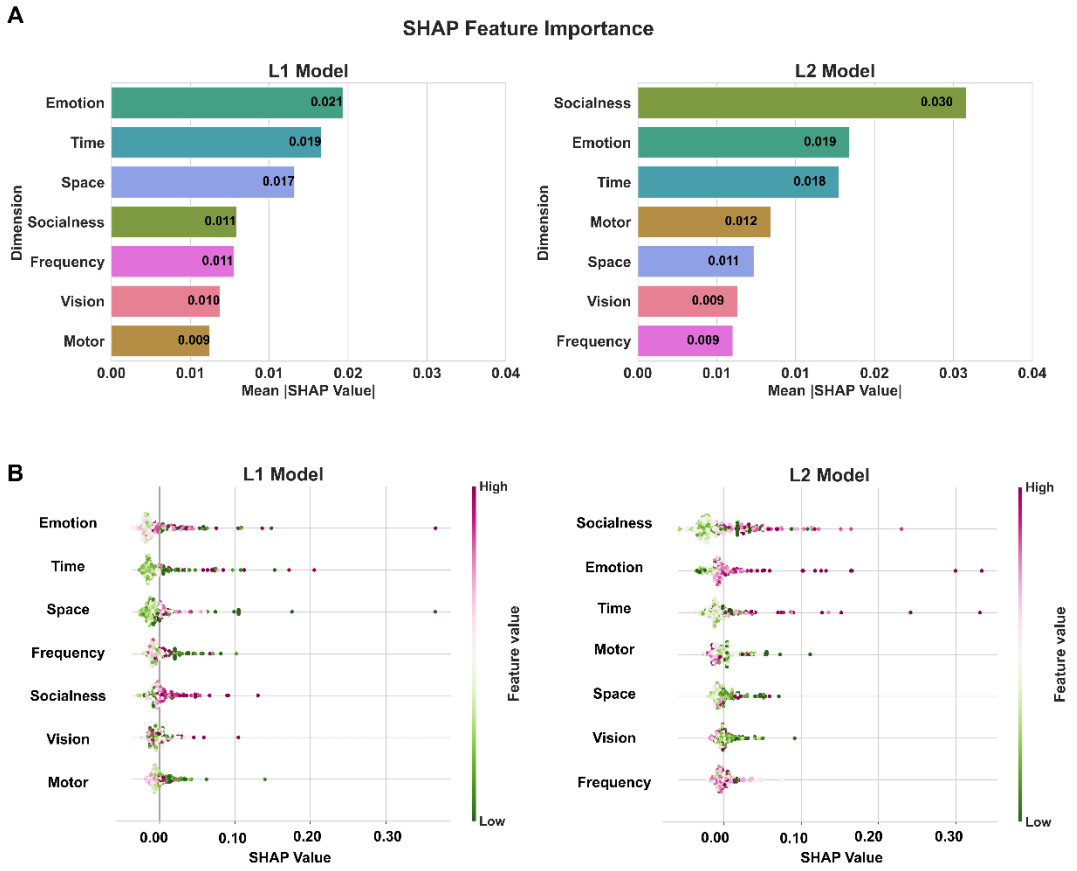
4. Distribution of the decoding accuracy across timepoints

We computed the decoding accuracy of each timepoint by computing the proportion of correct time label during all the iterations for every timepoint. The distribution of decoding accuracy was a long-tail distribution with most of the timepoints were with relatively low decoding accuracy or even unidentifiability.



FigS4 Distribution of the decoding accuracy from the cortico-cerebellar interaction features. The decoding accuracy among the whole timescale demonstrated an abnormal distribution with two peaks, one of the peaks was at the position of zero. This kind of distribution leaded the nonlinear hypothesis for the correlation analysis.

5. Semantic correlation of different neural features

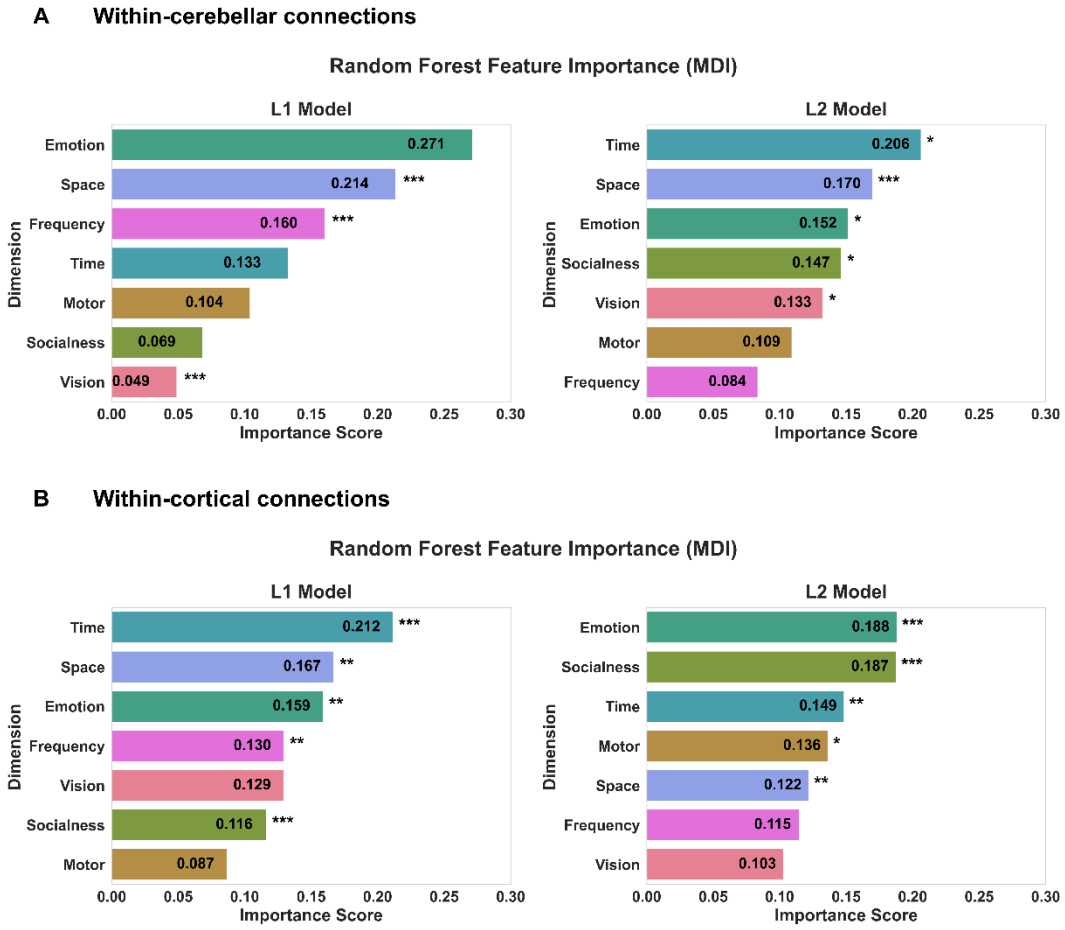


FigS5 Evaluation of feature importance using SHAP value. A. The absolute mean SHAP value was computed for each dimension in both the L1 and L2 models, reflecting the average contribution of each feature to model output regardless of direction. Higher values indicate greater importance. B. The SHAP beeswarm plots illustrate the distribution of SHAP values for individual timepoints (samples) across all features. Each dot represents a single timepoint, colored by the original feature value: green indicates lower values, while pink denotes higher values. The position along the x-axis represents the SHAP value, where negative values indicate a decreasing

effect on the model output, and positive values indicate an increasing effect. This visualization captures both the magnitude and direction of each feature's influence on the prediction.

6. Semantic correlation of different neural features

We utilized the random forest regression to fit the decoding accuracy from the within-cerebellar connectivity and within-cortical connectivity with the semantic time-series as well. The results revealed that the model fitted well with the accuracy of within-cerebellar connectivity, with the R^2 of 0.848 for L1 and 0.804 for L2. The regression model then considered that the space, frequency as well as the vision dimensions contributed to the L1 model fitting significantly while all the features except the motor and frequency dimensions contributed to L2 model fitting. The model for predicting the within-cortical connectivity underperformed as compared to the within-cerebellar features, with the R^2 of 0.828 for L1 and 0.796 for L2. The within-cortical model considered all the features except the vision and motor dimensions to be the contributors of model fitting for L1. However, all the features except the frequency and vision were regarded as significant predictors for L2.



FigS6 The correlation between the decoding accuracy from other connectivity features and the semantic time-series. A. The feature importance for predicting the within-cerebellar decoding accuracy. B. The feature importance for predicting the within-cortical decoding accuracy. ***: $p < 0.001$, **: $p < 0.01$, *: $p < 0.05$ after FDR correction. ·: $p < 0.05$ without multiple correction.