Title
The Geometric Structure of Shared Neural Semantics: Evidence from Cross-Subject EEG-Text Alignment

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors
Chen, Xiao
Lan, Yinyu
Zhang, Xin
et al.

Publication Date
2026

The Geometric Structure of Shared Neural Semantics: Evidence from Cross-Subject EEG-Text Alignment

Xiao Chen, Yinyu Lan, Xin Zhang, Xiangsen Liu, Xiao Zhang*,
School of Medicine and Health, Harbin Institute of Technology
{chenxiao25, lanyinyu, 25S028039, xiangsen.liu}@stu.hit.edu.cn, zhangxiao@hit.edu.cn*
* corresponding authors

Abstract

A fundamental challenge in cognitive science is whether the human brain employs a structured, relational representational scheme for semantics that is shared across individuals. While existing electroencephalography (EEG) studies have primarily focused on decoding or correlational analyses, it remains unclear whether sentence-level EEG activity encodes semantic structure at the level of representational geometry, beyond stimulus-specific decoding. This study investigates whether EEG signals recorded during natural reading can be aligned with representations capturing shared semantic structure and systematic correspondence with text-based distributional semantic models. Using the ZuCo 2.0 dataset, we learn an alignment between sentence-level EEG activity and a shared semantic embedding space and evaluate cross-subject generalization under a strict leave-one-subject-out protocol. Representational alignment is quantified using cross-subject semantic retrieval and Representational Similarity Analysis (RSA), linking neural activity to abstract semantic relations. Our results suggest that human EEG activity during language processing is consistent with a stable, relational organization of semantic information, consistent with structural properties of semantic organization captured in computational models, without implying representational equivalence.

Keywords: semantic representation; EEG; natural reading; cross-subject generalization; representational similarity analysis

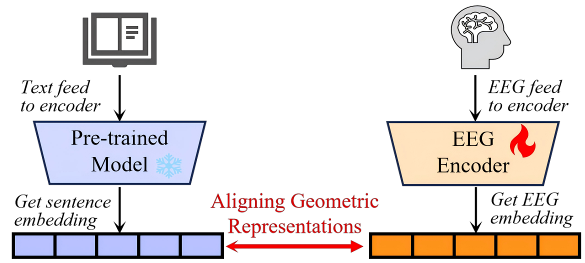

Figure 1: Overview of EEG–text semantic alignment. EEG signals recorded during natural reading are aligned with a fixed text-based semantic embedding space via a shared encoder, enabling comparison of representational geometry across modalities.

Introduction

Understanding how semantic meaning is represented in the human brain remains a central question in cognitive science(Horikawa & Kamitani, 2017; Mishra et al., 2025; Shen et al., 2024). During natural language comprehension, readers construct structured representations that capture not only the meanings of individual words or sentences, but also the relational structure among linguistic units and events(J. Li & Pylkkänen, 2021). While neuroimaging studies have shown that neural activity correlates with linguistic and semantic variables, it remains unclear to what extent neural representations preserve structured semantic geometry that can be meaningfully related to that captured by modern distributional language models(Chen et al., 2021; McGuire & Moshfeghi, 2025).

Recent advances in natural language processing have demonstrated that sentence embeddings learned from large-scale text corpora organize meaning in highly structured vector spaces(Church, 2017; Pennington et al., 2014), where semantic similarity corresponds to geometric proximity. In parallel, cognitive neuroscience has increasingly adopted representational similarity analysis (RSA) to compare the structure of model-derived representations with neural activity(Anderson et al., 2016). This line of work enables a shift from predicting isolated linguistic features to evaluating whether neural representations preserve relational semantic structure at the level of representational geometry(McGuire & Moshfeghi, 2025; Yi et al., 2023). Taken together, this shift reframes neurosemantic modeling as a question of representational geometry rather than decoding performance alone. However, most existing EEG-based studies focus on decoding or task-specific prediction(Ding et al., 2024; Song et al., 2022; Zhang & Li, 2025), leaving open the question of whether such semantic structure reflects shared representational organization across individuals.

Cross-subject generalization is particularly critical for assessing the cognitive relevance of neural representations, as shared representational geometry would suggest that observed structure reflects properties of human semantic processing rather than subject-specific idiosyncrasies(Z. Li et al., 2022). If semantic structure observed in EEG signals reflects stable properties of human language comprehension rather than idiosyncratic neural patterns, it should be recoverable across participants. Demonstrating cross-subject consistency therefore provides a strong test of whether neural representations capture abstract semantic relationships, independent of individual-specific neural signatures, rather than subject-bound decoding strategies.

In this work, we investigate whether EEG signals recorded

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

during natural reading preserve semantic representations that are (1) relationally aligned with text-based distributional semantic embeddings, and (2) structurally consistent across subjects. Figure 1 illustrates the central hypothesis tested in this work: if EEG activity during natural reading encodes relational semantic information at a population level, then a shared encoder should align EEG signals with a semantic space whose internal geometry mirrors that of text-based representations. Using the ZuCo 2.0 dataset(Hollenstein et al., 2023), we adopt a leave-one-subject-out (LOSO) evaluation protocol to ensure strict cross-subject generalization. Rather than focusing on classification accuracy alone, we evaluate representational correspondence at multiple levels, including sentence retrieval, distributional separation of matched and mismatched pairs, and correspondence of representational geometry via RSA.

Unlike traditional decoding approaches, this work focuses on representational geometry shared across individuals. Our results show that EEG embeddings support above-chance cross-subject retrieval and exhibit reliable separation between matched and mismatched EEG–text pairs. At the representational level, RSA reveals significant correspondence between EEG and text embedding spaces, indicating preservation of relational semantic structure.

Methods

Dataset and Experimental Setting

We evaluate our framework on the ZuCo 2.0 dataset, focusing on the Natural Reading (NR) paradigm. This dataset comprises EEG recordings from 18 participants reading over 300 sentences. Compared to task-driven protocols, the NR paradigm minimizes explicit task demands, allowing neural activity to more directly reflect spontaneous comprehension processes. To assess cross-subject robustness, we adopt a leave-one-subject-out (LOSO) cross-validation protocol, ensuring that model performance reflects generalized representational correspondence rather than subject-specific patterns.

EEG Preprocessing and Representation

Raw EEG signals are preprocessed following the ZuCo pipeline, with the addition of Multivariate Noise Normalization (MVNN) to reduce inter-trial and inter-subject variability. Word-level EEG segments are aggregated into a sentence-level representation to facilitate alignment with the semantic scale of text embeddings.

Let $X_i \in \mathbb{R}^{B \times T}$ denote the sentence-aligned EEG signal for sentence i, where B denotes EEG bands and T indexes time samples. We define a neural encoder f_θ to map these high-dimensional signals into a fixed-dimensional embedding space:

$$e_i = f_\theta(X_i), \quad e_i \in \mathbb{R}^d \tag{1}$$

This formulation allows the derived EEG embeddings e_i to be directly compared with sentence-level text embeddings using similarity-based measures.

EEG Encoder

Proposed EEG Encoder Architecture To address the inherent variability in neural signals and frequency-specific contributions to semantics, the EEG encoder f_θ is implemented as a subject-aware multi-band integrator. Unlike generic architectures, this design incorporates two key mechanisms to facilitate cross-subject alignment.

Subject-Aware Transformation: To account for individual physiological differences, we apply a subject-specific linear projection $P_s \in \mathbb{R}^{D \times D}$ for each subject $s \in S$. For a given input X_i, the transformation is defined as:

$$\hat{X}_{i,s} = P_s X_i, \tag{2}$$

where P_s serves as a learnable alignment layer that maps subject-specific neural manifolds into a common latent space before high-level encoding. During training, subject-specific projections are learned only from training subjects. For the held-out subject in the LOSO setting, no subject-specific parameters are introduced, ensuring that cross-subject generalization is not driven by subject memorization. This design encourages the encoder to disentangle subject-specific variability from shared semantic structure.

Frequency-Spectral Weighting: Recognizing that different EEG frequency bands contribute unequally to linguistic processing, we aggregate band-specific representations across $\mathcal{F} = \{\theta, \alpha, \beta, \gamma\}$. Let $\hat{X}_{i,f} \in \mathbb{R}^D$ denote the feature representation corresponding to frequency band f. The integrated representation is computed as a weighted combination:

$$\tilde{X}_i = \sum_{f \in \mathcal{F}} \alpha_f \odot \hat{X}_{i,f}, \tag{3}$$

where $\odot$ denotes element-wise multiplication, and $\alpha_f \in \mathbb{R}^D$ are learnable feature-wise weights. These weights are normalized across frequency bands using a softmax function:

$$\alpha_f = \frac{\exp(\mathbf{w}_f)}{\sum_{f' \in \mathcal{F}} \exp(\mathbf{w}_{f'})}, \tag{4}$$

where $\mathbf{w}_f \in \mathbb{R}^D$ are trainable parameters.

This architectural configuration prioritizes interpretability and controlled inductive biases, enabling the analysis of cross-subject semantic alignment without relying on overly complex architectures. From a cognitive perspective, this design can be interpreted as separating subject-specific neural variability from shared representational structure while emphasizing frequency components relevant to language processing.

Baseline Encoders To isolate whether cross-subject EEG–text correspondence can emerge without explicit structural inductive biases, we consider two minimal baseline models with limited representational assumptions:

- **Linear:** A single linear projection that flattens X_i and maps it to $\mathbb{R}^d$. This baseline tests whether a global linear transformation is sufficient to align EEG representations with semantic embeddings across subjects, without modeling frequency structure or subject variability.

- **MLP:** A multi-layer perceptron with non-linear activations but without subject-aware or frequency-specific mechanisms. This model increases representational capacity while remaining agnostic to EEG-specific structure, allowing us to assess whether generic non-linearity alone can account for alignment performance.

These baselines are intentionally designed to be structurally minimal, enabling us to evaluate whether cross-subject semantic correspondence arises from general function approximation or requires domain-specific inductive biases. All models share the same training objective and evaluation protocol as the proposed encoder, ensuring that performance differences reflect architectural choices rather than optimization or data leakage effects.

Text Embedding Space

For each sentence stimulus, we obtain text embeddings using pretrained transformer-based language models, including Sentence-BERT (SBERT) and RoBERTa(Liu et al., 2019; Reimers & Gurevych, 2019), as operationalized approximations of distributional semantic structure. All text encoders are kept fixed during training and serve as stable reference spaces for evaluating representational correspondence.

Let $s_i \in \mathbb{R}^d$ denote the text embedding of sentence i, with dimensionality matched to that of the EEG embedding for cosine similarity computation. By anchoring EEG representations to multiple pretrained language embedding spaces, we assess whether neural activity during reading exhibits relational structure that generalizes across different computational models of semantics, rather than reflecting alignment to a specific architecture.

EEG–Text Correspondence Objective

To encourage correspondence between EEG embeddings and their corresponding text embeddings, we employ a contrastive objective based on cosine similarity. For a batch of N EEG–text pairs $\{(e_i, s_i)\}_{i=1}^{N}$, the cosine similarity between an EEG embedding and a text embedding is defined as:

$$sim(e_i, s_j) = \frac{e_i^\top s_j}{\|e_i\| \|s_j\|} \tag{5}$$

Positive pairs correspond to matched EEG–sentence pairs $(i = j)$, while negative pairs are formed by mismatched combinations within the same batch $(i \neq j)$. The model is optimized to maximize similarity for positive pairs while minimizing similarity for negative pairs, thereby enforcing a margin between matched and mismatched pairs in the joint embedding space(Radford et al., 2021). This objective encourages EEG embeddings not only to approach their corresponding text embeddings but also to preserve relative relational structure among sentences, enabling evaluation of representational geometry beyond pair-wise matching. Notably, this objective enforces instance-level correspondence but does not explicitly optimize the global similarity structure among sentences, allowing representational geometry to be evaluated as an emergent property.

Evaluation Metrics

Retrieval-Based Evaluation We evaluate alignment quality via a cross-modal retrieval task using five key metrics: (1) Top-1 and Top-5 Accuracy, representing the success rate of exact and near-exact retrieval; (2) Mean Rank, indicating the average retrieval position of the ground truth; (3) Mean Cosine Similarity (Mean Cos), measuring the representational closeness. Together, these metrics assess the reliability of EEG–text correspondence across the embedding space under cross-subject evaluation.

Representational Similarity Analysis (RSA) We assess representational correspondence using Representational Similarity Analysis (RSA). For each modality, we compute a pairwise cosine similarity matrix across all sentence embeddings, denoted as R_{EEG} and R_{Text}. RSA correspondence is examined by correlating the vectorized upper triangular entries (excluding the diagonal) of the two matrices. To characterize different aspects of representational alignment, we consider both Spearman's rank correlation (ρ), which reflects monotonic structural correspondence, and Pearson's correlation (r), which reflects linear geometric similarity. These correlations are used to characterize structural correspondence and to support qualitative interpretation of representational geometry, rather than serving as standalone performance metrics.

Statistical Analysis All reported metrics are averaged across LOSO folds. For RSA correspondence, statistical significance is assessed using permutation testing by randomly shuffling sentence labels in one modality and recomputing correlations. Because pairwise similarity values are bounded and non-independent, and may deviate from Gaussian assumptions, we primarily rely on Spearman's rank correlation to assess relational correspondence. Pearson correlation is additionally reported to characterize linear trends in similarity magnitude.

Cluster-Level Geometry Analysis To examine coarse-grained representational structure, we grouped sentences using hierarchical clustering (average linkage) on pairwise cosine distances computed from text embeddings. The number of clusters was fixed to $K = 12$ for visualization and analysis.The number of clusters was chosen to balance interpretability and granularity, and fixed across analyses for consistency. Clustering was performed exclusively in the text embedding space, with EEG representations not used to define cluster assignments.

Based on these assignments, we computed average pairwise similarity between clusters for both EEG and text embeddings, yielding cluster-level representational similarity matrices. Comparing these matrices assesses whether higher-level semantic groupings defined in the text space are pre-

Table 1: Cross-subject EEG–text retrieval and representational correspondence under different pretrained text embedding spaces (mean ± SD across LOSO folds). Results test whether EEG representations generalize across individuals beyond subject-specific neural signatures. Values are reported with two decimal places unless higher precision is required.

EEG Model	Text Model	Top-1↑	Top-5↑	Mean Rank↓	Mean Cos↑	RSA↑
Random	–	0.35%	1.78%	141.59	0.0026	0.0001
Linear	SBERT	1.05% ± 0.34%	6.67% ± 1.88%	25.65 ± 1.28	0.39 ± 0.01	0.41 ± 0.01
	RoBERTa	1.06% ± 0.92%	6.72% ± 1.16%	25.04 ± 1.21	0.40 ± 0.01	0.44 ± 0.01
MLP	SBERT	2.06% ± 1.61%	13.02% ± 2.74%	22.08 ± 1.59	0.47 ± 0.02	0.34 ± 0.01
	RoBERTa	1.64% ± 0.26%	14.45% ± 2.72%	22.36 ± 1.22	0.46 ± 0.01	0.35 ± 0.01
Ours	SBERT	**2.69**% ± 1.57%	**14.39**% ± 2.09%	**18.90** ± 2.23	**0.57** ± 0.01	**0.62** ± 0.01
	RoBERTa	**2.70**% ± 0.67%	**15.62**% ± 0.85%	**19.65** ± 1.18	**0.55** ± 0.02	**0.63** ± 0.02

served in EEG representations and characterizes systematic differences between neural and textual semantic geometry.

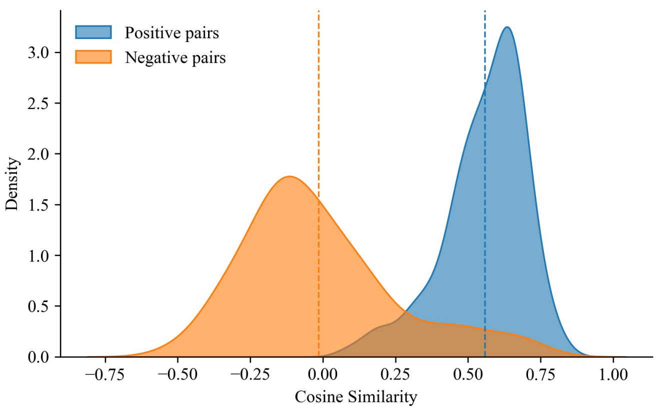

Figure 2: Distribution of EEG–text cosine similarity for matched and mismatched sentence pairs, computed in the SBERT embedding space.The clear separation between distributions indicates reliable cross-subject discrimination between corresponding and non-corresponding sentence representations.

Results

Cross-subject Retrieval Performance

We first evaluate whether EEG embeddings learned under the LOSO setting support cross-subject sentence retrieval above chance based on semantic similarity. Table 1 summarizes retrieval performance across models. Under the same cross-subject protocol, the proposed model consistently outperforms baseline approaches across all metrics. Using SBERT embeddings, Top-1 accuracy improves from 1.05% to 2.69% and mean rank decreases from 25.65 to 18.90 relative to a linear projection, indicating more reliable cross-subject matching. Comparable improvements are observed with RoBERTa embeddings (Table 1), suggesting that the gains are not specific to a particular text model. Performance gains over the MLP baseline further indicate that improvements cannot be attributed solely to generic non-linear transformations. In addition, the proposed model achieves higher mean cosine similarity, suggesting better separation between matched and mismatched sentences in the embedding space.

Although the absolute Top-1 accuracy appears modest, it significantly exceeds the theoretical chance level in a zero-shot cross-subject retrieval task. This gap demonstrates that even with high neural noise, a robust shared semantic structure is preserved across brains. Together, these results indicate that EEG signals recorded during natural reading support cross-subject semantic correspondence, and that the proposed encoder yields more reliable alignment than linear or shallow non-linear mappings. Retrieval performance is interpreted here as evidence of cross-subject correspondence rather than high-fidelity semantic decoding from EEG signals.

Similarity Distributions

To characterize EEG–text correspondence at the distributional level, we analyze cosine similarity distributions for matched and mismatched sentence pairs (Figure 2). Matched pairs exhibit a sharply peaked distribution at higher similarity values, whereas mismatched pairs are centered near zero with a broader spread. This yields a clear separation with minimal overlap, indicating that semantic correspondence can be reliably discriminated from EEG signals. The separation is statistically robust (Cohen's d = 2.63, AUC = 0.944, $p < 0.001$), confirming a substantial margin between matched and mismatched EEG–text representations.

This result establishes that EEG–text correspondence exceeds chance-level similarity, motivating subsequent analyses of representational geometry to test whether relational structure is preserved beyond pairwise discrimination.

Low-level Statistical Control

To assess whether the observed EEG–text correspondence could be explained by coarse spectral statistics, we conducted a low-level control analysis in which EEG signals were averaged within each frequency band across time, discarding all temporal structure, and linearly projected into the text embedding space under the same LOSO protocol. This control yields near-chance retrieval performance (Top-1 = 0.65%) and a dramatic reduction in representational similar-

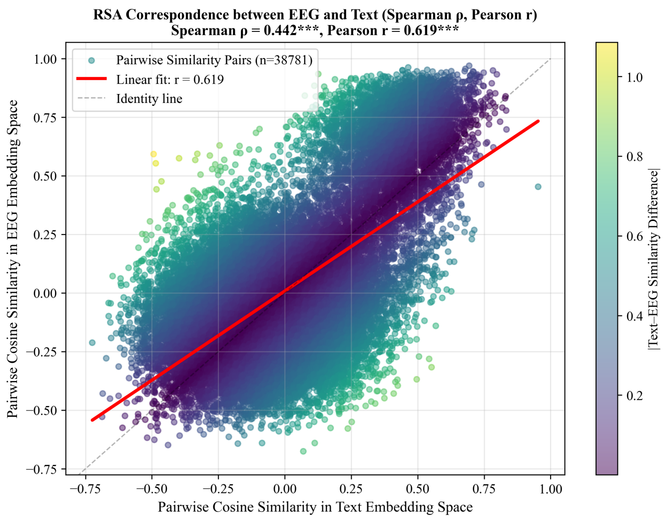

Figure 3: Representational similarity correspondence between EEG and text embeddings. Each point represents a pairwise cosine similarity between two sentences in EEG and SBERT text embedding spaces. Spearman and Pearson correlations quantify structural and linear correspondence between modalities.

ity correspondence (RSA = 0.05), compared to the full linear model. Although average cosine similarity is high (Mean Cos = 0.8020), the minimal margin between matched and mismatched pairs indicates a collapsed representational structure lacking meaningful relational organization. These results rule out an explanation based on band-wise power distributions alone.

Representational Similarity Correspondence

While retrieval metrics assess whether EEG embeddings can identify corresponding sentences, they do not reveal whether EEG and text embeddings share similar relational structure. We therefore perform Representational Similarity Analysis (RSA) to compare pairwise relationships across modalities. Consistent with our methodological framing, RSA correlations are reported to characterize the degree of structural correspondence between modalities, rather than to function as standalone performance metrics.

Pairwise Similarity Correspondence Figure 3 shows pairwise cosine similarities in the EEG embedding space as a function of text embedding similarities across all sentence pairs. A positive relationship is observed, with significant Spearman's rank correlation ($\rho = 0.442$) and Pearson's correlation ($r = 0.619$). The Spearman correlation indicates that relative similarity ordering among sentences is partially preserved, suggesting that EEG embeddings capture coarse-grained relational structure at the population level. The Pearson correlation further indicates partial alignment of absolute similarity values, while deviations from the identity line remain, consistent with modality-specific distortions in representational geometry rather than simple linear mapping. Overall, these results provide converging evidence that EEG embeddings preserve systematic relational organization aligned with text-based semantic structure, while undergoing principled geometric transformations rather than metric equivalence.

Cluster-level Representational Geometry To examine representational structure at a more coarse-grained level, we compare cluster-level similarity matrices derived from EEG and text embeddings (Figure 4). When organized according to text-defined clusters, both modalities exhibit a clear block-diagonal structure, indicating internally coherent groupings. Notably, clusters with high narrative and emotional salience, such as Cluster 3 (centered on interpersonal relationships and personality traits), show relatively strong internal consistency in the EEG space.

While clusters that are proximal in the text embedding space tend to exhibit elevated similarity in the EEG space, the EEG matrix displays attenuated contrasts and localized distortions. The corresponding difference matrix highlights systematic divergences; for example, EEG representations show increased similarity—reflecting relative compression—for Cluster 6 (historical events, business, and political associations). This pattern suggests that while SBERT preserves finer-grained distinctions among socio-political categories, EEG representations may reflect a more integrated encoding of such content.

In contrast, the elevated within-cluster similarity observed for Cluster 3 is consistent with the idea that neural representations during natural reading may emphasize event-level integration and experiential relevance. Taken together, these results indicate that EEG representations do not simply mirror text-based semantic organization, but instead reflect structured transformations shaped by the cognitive demands of language comprehension, with greater sensitivity to socially and narratively salient content than to abstract institutional distinctions.

Summary of Results

Overall, the results show that EEG signals recorded during natural reading support structured embeddings that enable reliable cross-subject correspondence. Retrieval analyses demonstrate functional alignment, while distributional and representational similarity analyses indicate that EEG and text embeddings share partially aligned representational geometry. Together, these findings suggest that EEG–text correspondence extends beyond sentence-level matching and is consistent with structured relationships among stimuli in the learned embedding space.

Discussion and Conclusion

Discussion

This study examined whether neural activity recorded during natural reading encodes structured semantic representations that generalize across individuals. By aligning sentence-level

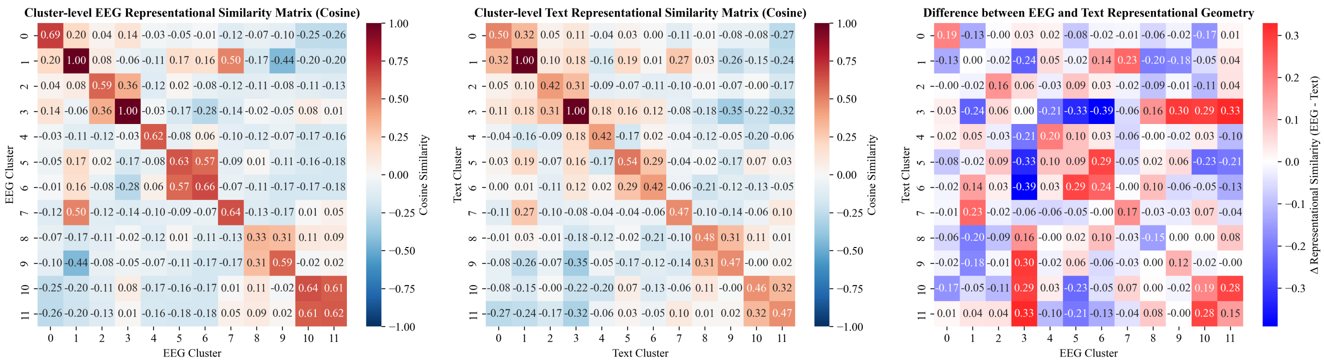

Figure 4: Cluster-level representational similarity matrices for SBERT text embeddings, EEG embeddings, and their difference. Clusters are defined via hierarchical clustering in the SBERT text embedding space. Left: corresponding similarity structure in the EEG embedding space. Middle: cosine similarity between semantic clusters. Right: difference between EEG and text similarity matrices. Overall, the preserved block structure indicates that coarse-grained semantic relationships among sentence clusters are maintained in EEG representations across individuals.

EEG representations with pretrained text embeddings under a strict leave-one-subject-out protocol, we assessed semantic correspondence at both functional and representational levels. EEG embeddings exhibited reliable alignment with text-based semantic structure beyond chance, supporting the presence of shared relational organization in neural representations of meaning.

Importantly, this alignment emerges despite fundamental architectural differences between biological neural systems and transformer-based language models. While large language models are optimized for next-token prediction over large-scale text corpora, the human brain processes language through a distributed biological system shaped by evolutionary constraints and real-time interaction. The presence of shared representational geometry across these systems suggests a form of functional convergence, in which distinct learning mechanisms give rise to partially similar relational structures for organizing high-level semantic information.

Cross-subject retrieval demonstrates that EEG representations contain information systematically related to sentence-level semantics across participants. However, retrieval performance alone cannot distinguish between stimulus-specific matching and structured representation. Representational similarity analysis addresses this limitation by revealing correspondence between the relative geometric organization of EEG and text embedding spaces. Critically, this correspondence is unlikely to be explained by low-level spectral properties: a control analysis based on band-averaged power yields near-chance performance, ruling out an account based on coarse statistical features and supporting the interpretation that the observed geometry reflects higher-order organization.

At the same time, representational analyses reveal systematic divergences between neural and text-based semantic geometry. Cluster-level results clarify these differences: while SBERT preserves fine-grained symbolic distinctions across domains, EEG representations appear more sensitive to experiential and social salience. For example, higher internal consistency in Cluster 3 (interpersonal relationships and emotional traits) contrasts with relative compression in Cluster 6 (institutional events and business-related content). This pattern is consistent with the idea that neural semantic representations may preferentially encode dimensions related to personal relevance and event structure—features grounded in human experience that are less explicitly emphasized in text-only models.

Methodologically, these findings highlight the importance of combining cross-subject evaluation with representational analyses. Evaluating representational geometry under subject-independent constraints provides a more stringent test of shared semantic structure than within-subject decoding alone. More broadly, this framework opens the possibility of examining whether other physiological signals, such as eye movements or reading times, exhibit similar geometric organization, offering a pathway toward understanding the multimodal nature of semantic representation.

Limitations and Future Work

Despite these limitations, EEG's low signal-to-noise ratio and spatial resolution constrain retrieval precision. Future work should adopt more biologically plausible or perceptually grounded models than SBERT, isolate the roles of specific frequency bands and temporal dynamics, and integrate eye-tracking and reading-time measures to assess representational alignment across multiple language-related modalities.

Acknowledgement

This work was supported by the Noncommunicable Chronic Diseases-National Science and Technology Major Project (Grant No. 2024ZD0524402).

References

Anderson, A. J., Zinszer, B. D., & Raizada, R. D. S. (2016). Representational similarity encoding for fmri. *NeuroImage*, *128*, 44–53.

Chen, X., Martin, R., & Fischer-Baum, S. (2021). Challenges for using representational similarity analysis to infer cognitive processes. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *43*.

Church, K. W. (2017). Word2vec. *Natural Language Engineering*, *23*, 155–162.

Ding, Y., Zhang, S., Tang, C., & Guan, C. (2024). Masatcn: Multi-anchor space-aware temporal convolutional neural networks for eeg emotion recognition. *IEEE Journal of Biomedical and Health Informatics*, *28*, 3953–3964.

Hollenstein, N., Tröndle, M., Plomecka, M., Kiegeland, S., Özyurt, Y., Jäger, L. A., & Langer, N. (2023). The zuco benchmark on cross-subject reading task classification with eeg and eye-tracking data. *Frontiers in Psychology*, *13*, 1028824.

Horikawa, T., & Kamitani, Y. (2017). Generic decoding of seen and imagined objects using hierarchical visual features. *Nature Communications*, *8*, 15037.

Li, J., & Pylkkänen, L. (2021). Disentangling semantic composition and semantic association in the left temporal lobe. *Journal of Neuroscience*, *41*, 6526–6538.

Li, Z., Zhu, E., Jin, M., Fan, C., He, H., Cai, T., & Li, J. (2022). Dynamic domain adaptation for class-aware cross-subject and cross-session eeg emotion recognition. *IEEE Journal of Biomedical and Health Informatics*, *26*(12), 5964–5973.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. *arXiv*.

McGuire, N., & Moshfeghi, Y. (2025). Towards brain passage retrieval: An investigation of eeg query representations. *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 613–623.

Mishra, A., Shukla, S., Torres, J., Gwizdka, J., & Roychowdhury, S. (2025). Thought2text: Text generation from eeg signal using large language models. *Findings of the Association for Computational Linguistics: NAACL*, 3747–3759.

Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 1532–1543.

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., & Clark, J. (2021). Learning transferable visual models from natural language supervision. *Proceedings of the International Conference on Machine Learning*, 8748–8763.

Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv*.

Shen, G., Zhao, D., He, X., Feng, L., Dong, Y., Wang, J., Zhang, Q., & Zeng, Y. (2024). Neuro-vision to language: Enhancing brain recording-based visual reconstruction and language interaction. *Advances in Neural Information Processing Systems*, *37*, 98083–98110.

Song, Y., Zheng, Q., Liu, B., & Gao, X. (2022). Eeg conformer: Convolutional transformer for eeg decoding and visualization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *31*, 710–719.

Yi, K., Wang, Y., Ren, K., & Li, D. (2023). Learning topology-agnostic eeg representations with geometry-aware modeling. *Advances in Neural Information Processing Systems*, *36*, 53875–53891.

Zhang, H., & Li, H. (2025). Transformer-based eeg decoding: A survey. *arXiv preprint arXiv:2507.02320*.