SED-GPT: A Non-Invasive Method for Long-Sequence Fine-Grained Semantics and Emotions Decoding

Wenhao Cui , Zhaoxin Wang , Lei Ma *

Posted Date: 21 August 2025

doi: 10.20944/preprints202508.1518.v1

Keywords: semantic decoding; emotion decoding; functional magnetic resonance imaging (fMRI); large language models (LLMs); Semantic to Brain Response Conversion Module

 doi:10.20944/preprints202508.1518.v1

Article

SED-GPT: A Non-Invasive Method for Long-Sequence Fine-Grained Semantics and Emotions Decoding

Wenhao Cui [1,2], Zhaoxin Wang [1] and Lei Ma [1,*]

[1] School of Information Science and Technology, Nantong University, Nantong 226019, China

[2] Institute of Psychiatry, Psychology & Neuroscience, King's College London, London WC2R 2LS, UK

* Correspondence: malei@ntu.edu.cn

Featured Application

This study introduces the Semantic and Emotion Decoding Generative Pre-trained Transformer (SED-GPT), a non-invasive framework for decoding fine-grained semantics and emotions from long-sequence fMRI. The approach provides potential applications in precise detection of affective disorders, personalized neuromodulation, and human–computer interaction systems that adapt to users' emotional and semantic states.

Abstract

Traditional emotion decoding methods typically rely on short sequences with limited context and coarse-grained emotion categories. To address these limitations, we proposed the Semantic and Emotion Decoding Generative Pre-trained Transformer (SED-GPT), a non-invasive method for long-sequence fine-grained semantics and emotions decoding on extended narrative stimuli. In the encoding stage, we employed a Semantic to Brain Response Conversion Module (SBRCM) to align neural signals with semantic representations, which modeled the relationship between stimulus features and brain responses. To enhance generalization, we constructed a word rate model and estimated noise covariance. In the decoding stage, new semantic sequences were generated according to previous candidate semantic sequences and the prior probabilities of large language models (LLMs). These sequences were converted into brain response sequences through SBRCM, which were then compared with actual brain responses. By combining prior probabilities with likelihood probabilities, the optimal sequence was iteratively derived, generating long-sequence semantic vectors. Finally, the GoEmotions framework was applied to quantify the multiclass emotional distribution of the long-sequence semantic representations. SED-GPT achieves a BERTScore-F1 of 0.650 on semantic decoding and attains a cosine similarity (CS) of 0.504 and a Jensen–Shannon similarity (JSS) of 0.469 for emotion decoding ($p<0.05$). Functional connectivity analyses reveal persistent coupling between the language network and the emotion network, which provides neural evidence for the language-emotion interaction mechanism in Chinese.

Keywords: semantic decoding; emotion decoding; functional magnetic resonance imaging (fMRI); large language models (LLMs); Semantic to Brain Response Conversion Module

1. Introduction

Human emotional experience is complex and involves inner speech activities and self-dialogue during semantic perception [1]. These inner speech processes encode subjective emotional states and elicit neural activation in regions such as the anterior cingulate cortex, insula, and ventromedial prefrontal cortex. These patterns partially overlap with those observed during emotional expression [2–4]. Therefore, emotion constitutes not merely a physiological or behavioral response but also involves cognitive-semantic processing. Emotional states activate their associated semantic networks,

 doi:10.20944/preprints202508.1518.v1

while semantic representations contribute to the refined construction of emotional experience [7–9]. Long-sequence semantic decoding can effectively extract context-dependent information during emotional construction, thereby providing the foundation for fine-grained emotional representation [10,11].

Common neuroimaging modalities for emotion decoding primarily include electroencephalography (EEG), magnetoencephalography (MEG), functional near-infrared spectroscopy (fNIRS), and functional magnetic resonance imaging (fMRI) [12]. EEG and MEG offer millisecond temporal resolution but poor spatial specificity due to volume conduction and field spread, making source localization inherently uncertain [13]. For fNIRS, although it can provide relatively high spatial resolution, its measurement depth is only 2-3 cm [14]. In contrast, fMRI offers high spatial resolution and whole-brain coverage. This technique can not only enable precise localization of emotion related brain regions but also supports simultaneous detection of co-activation patterns between language processing networks and emotion regulation systems via blood oxygenation level dependent (BOLD) signals [15]. These features provide a clear advantage for studying the neural mechanisms of emotion and semantic interaction in the Chinese context.

Traditional emotion classification methods typically rely on short-sequence emotion-induction paradigms and machine learning (ML) algorithms. Kassam et al. elicited nine discrete emotions using word-cued spontaneous emotion induction tasks, achieving a rank accuracy of 0.84 in fMRI-based emotion discrimination [16]. Saarimäki et al. employed multivariate pattern analysis (MVPA) to distinguish six basic emotions, demonstrating good cross-subject generalizability [17]. However, conventional algorithms show significantly reduced decoding accuracy when the number of target emotion categories increases, and these algorithms face constraints from short stimulus sequences, non-naturalistic experimental paradigms, and coarse-grained emotion classification [18].

In recent years, large language models (LLMs) exemplified by GPT have demonstrated outstanding performance across diverse domains. These models can effectively capture rich contextual semantic representations and exhibit deep-level semantic reasoning capabilities, offering a novel approach to long-sequence emotion decoding [19]. Tang et al. developed an fMRI based generative decoding framework that maps brain activity patterns into the latent semantic space of LLMs, achieving neural reconstruction of continuous English speech (BERTScore = 0.82) and revealing the distributed encoding characteristics of the brain's language network [20]. However, current research has two critical limitations. First, existing methods have not been validated on high-context language systems such as Chinese [21]. Second, prior studies have primarily focused on decoding basic semantic content, failing to further examine the subtle emotional components of language processing [22]. These limitations constrain our understanding of the neural mechanisms that underlie language and emotion interaction in real-world scenarios.

Accordingly, this study aims to develop and validate a fine-grained semantic and emotion decoding framework for extended Chinese narratives, bridging the current gap in understanding the neural mechanisms underlying language-emotion interactions. The innovations of this study include:

- Extended emotion decoding in high-context language: This study investigates the feasibility of long-sequence emotion decoding in Chinese, extending emotion decoding to high-context language systems.
- Fine-grained emotion decoding with SED-GPT: We propose a novel fine-grained decoding framework (SED-GPT) for Chinese narratives, which aligns brain activity with LLM-based semantic vector representations to reconstruct inner speech semantics.
- Dynamic neural interactions in emotion-semantic processing: By systematically examining the dynamic interplay between the language network and emotional systems during Chinese semantic processing, this work provides neural evidence for cognition-emotion coupling.

Overall, this approach is expected to promote the development of affective brain computer interfaces (BCI), demonstrating potential applications in depression treatment and cognitive behavioral therapy.

 doi:10.20944/preprints202508.1518.v1

2. Materials and Methods

2.1. Dataset

This study employed the publicly available dataset SMN4Lang to evaluate our model, which comprises structural MRI and functional MRI data from 12 participants, with each participant contributing 6 hours of fMRI data [23].

At trial onset, a screen displayed the instruction "Waiting for scan", followed by an 8-second blank screen. The instruction then changed to "The audio will begin shortly. Please listen carefully" for 2.65 seconds before the auditory stimulus was presented.

The stimulus set consisted of 60 audio clips (4–7 minutes each) from People's Daily news stories, covering diverse topics including education and culture. All recordings were narrated by the same male speaker. Manual timestamp alignment was performed to ensure precise synchronization between the audio and corresponding textual transcripts.

Structural MRI and fMRI were acquired using a Siemens Prisma 3T scanner equipped with a 64-channel receive coil. T1-weighted images were obtained with a 3D MPRAGE sequence at an isotropic spatial resolution of 0.8 mm^3 with the following parameters: TR = 2.4 s, TI = 1 s, TE = 2.22 ms, flip angle = 8°, FOV = 256 × 256 mm.

fMRI data were collected using a BOLD-sensitive T2*-weighted GE-EPI sequence with the following parameters: TR = 710 ms, TE = 30 ms, flip angle = 54°, in-plane resolution = 2 mm, FOV = 212 × 212 mm.

2.2. Data Preprocessing

In this research, fMRIPrep was used for batch preprocessing of the structural MRI and fMRI data in the Chinese semantic perception dataset. 4 participants were excluded after preprocessing due to extensive occipital lobe damage in their structural scans, leaving eight participants for analysis.

30 long-sequence task fMRI files were randomly sampled to identify brain regions showing significant activation during Chinese semantic perception and in response to twenty-one emotion word categories using second-level general linear model (GLM) [24,25]. Psychophysiological interaction analysis (PPI) then probed the neural pathways underlying semantic perception [26].

In comparing activation between task and resting states, at the individual analysis stage we used a first level GLM based on the hemodynamic response function and an autoregressive noise model. In the group analysis we constructed group level activation statistical maps using a second level GLM. After applying a voxel level correction ($p < 0.001$, $Z > 3.09$) and filtering by a minimum cluster size of 50 voxels, we identified regions showing significant activation [27].

In the study of neural pathways underlying Chinese semantic processing, BOLD signal time series were extracted from predefined brain seed regions. Physiological regressors, psychological regressors, and the original interaction regressor were constructed and then convolved with the hemodynamic response function. A first level GLM was fitted, followed by a second-level GLM group comparison to generate Z-statistic parametric maps. False Discovery Rate (FDR) correction ($p < 0.001$) was applied to identify brain regions exhibiting significant functional coupling with the semantic perception hubs [26].

In this study, we used the Affective Lexicon Ontology to mark timestamps for 21 emotions (joy, calm, respect, praise, trust, love, well wishing, anger, sadness, disappointment, guilt, longing, panic, fear, shame, frustration, disgust, blame, jealousy, doubt, surprise) and constructed an event matrix to compare BOLD signal differences evoked by emotion words and neutral words [28]. After multiple comparison correction and filtering by a minimum cluster size of 10 voxels, we identified the ROIs showing significant activation for each emotion category [27].

Based on the significant activated brain regions and the functionally coupled regions identified above, along with prior knowledge of brain areas involved in Chinese semantic perception [29,30], MNI coordinates were extracted to generate ROI masks. Using these ROI masks, the fMRI response

 doi:10.20944/preprints202508.1518.v1

time series of voxels within each mask were extracted. The 3D coordinates of the voxels were then re-encoded into 1D indices, forming a 2D neural response matrix.

2.3. SED-GPT

In this study, we propose a non-invasive method for long-sequence fine-grained semantics and emotions decoding, termed Semantic and Emotion Decoding Generative Pretrained Transformer (SED-GPT), as shown in Figure 1.

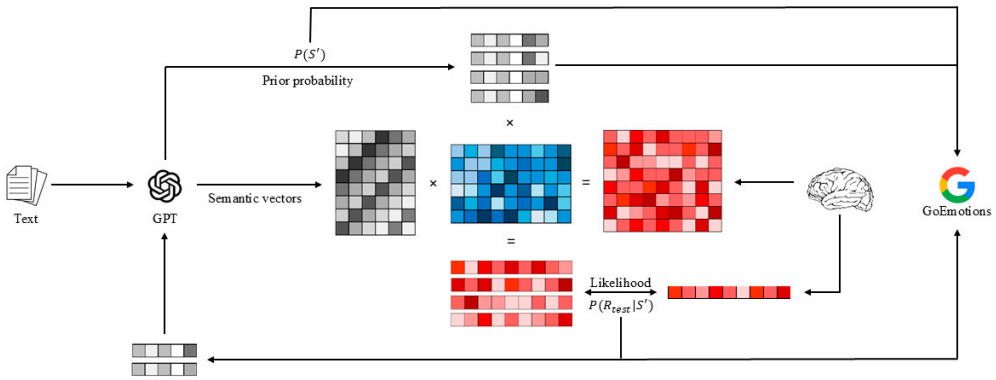

Figure 1. Overall architecture of the SED-GPT model.

In the encoding stage, we employed a Semantic to Brain Response Conversion Module (SBRCM) to minimize the inter-modal distance between linguistic stimuli and brain activity patterns. This module established a mapping between stimulus features and corresponding neural responses, incorporating a word rate model while estimating noise covariance matrices to improve model generalizability.

In the decoding stage, new semantic sequences were generated using linguistic priors from the language model and candidates. These sequences were then projected into neural response space through the SBRCM. The predicted simulated responses were compared with empirically observed brain activity patterns to compute likelihood distributions. Through iterative Bayesian integration of prior probabilities and likelihood estimates, the optimal semantic sequence was reconstructed. This iterative process enabled the generation of extended semantic vectors, whose emotional valence was subsequently classified into multiple categories using the GoEmotions framework [31].

2.4. Fine-Tuning of LLMs

Due to the differences in writing systems and word boundary markers, English text contains clear separators (spaces) between words and exhibits rich morphological variation, making word-level tokenization effective for capturing semantic and syntactic information. In contrast, Chinese lacks explicit word boundaries, and each character carries its own meaning, so character-level tokenization allows flexible combination into complete lexical units [32]. Consequently, English LLMs are typically trained with words as tokens, whereas Chinese LLMs use characters as tokens.

Character-encoded Chinese LLMs successfully avoid the complexity of Chinese word segmentation and require less computation for training [32]. However, this design choice also means that high-performance, open-source Chinese GPT based on word-level tokens remain unavailable [33,34]. Character-level tokenization in Chinese GPT models does not align with word-based semantic encoding mechanisms: neuroimaging studies show that Mandarin users rely on explicit word boundaries, encoding meaning at the word rather than character level [34,35].

Unlike the character-level tokenization commonly used in Chinese LLMs, human language comprehension integrates lexical, syntactic, and contextual information via a distributed semantic network, activating and mapping the corresponding conceptual representations to enable word-level

 doi:10.20944/preprints202508.1518.v1

semantic processing and reasoning. In this process, newly encountered symbols are matched, retrieved and compared against existing network nodes, activating relevant concepts or semantic domains [36]. This semantic information processing mechanism parallels the token-based encoding used in LLMs, where each input token is projected into a high-dimensional semantic space to capture and combine conceptual features [37].

In this study, we indexed semantic feature vectors at the word level. We used GPT-4 to translate the time aligned Chinese transcripts into English in batch and used a GPT-2 model—fine-tuned on the DeepMind Q&A news corpus—to construct the stimulus matrix [38]. The prompt for LLM-based timestamp conversion was: Convert the Chinese text into the most appropriate English. You may split intervals and their corresponding times; ensure that each interval corresponds to exactly one English word and that the xmax value at the end of each sentence remains consistent.

For each word-time pair at every timestamp, the word sequence was fed into the GPT language model, and the semantic feature vector of the target word was extracted from the 9th layer of the model, where the semantic feature vector represents a 768-dimensional semantic embedding [39]. After obtaining the target words and their corresponding semantic feature vectors, these embedding vectors were temporally resampled to match the fMRI data acquisition time points using a three-lobe Lanczos filter [40].

2.5. Semantic to Brain Response Conversion Moduler

In the encoding stage, the SBRCM receives input from two modalities: 768-dimensional semantic embeddings extracted from GPT-2 and BOLD signals recorded by fMRI. To account for the temporal-scale difference between these two modalities, the semantic stimulus vectors from 5-10th TRs prior to neural response onset were concatenated to construct a joint feature space. A linear mapping from this feature space to neural signals was then established through L2-regularized regression, with a word rate model being derived by estimating semantic occurrence frequency within individual TRs. To enhance model generalizability, a bootstrapping approach was incorporated in the encoding phase [41].

In the decoding stage, candidate semantic sequences were initially generated based on GPT-2 prior probabilities. These candidate sequences were subsequently projected into neural response space through the SBRCM to produce predicted brain responses. Each candidate's likelihood was computed by evaluating the correspondence between its predicted response and the empirically observed neural signals. Candidate sequences exhibiting significant discrepancies with the actual brain response patterns were iteratively filtered out. The language model priors were then integrated with the neural likelihoods, and beam search was employed to select a new set of candidate tokens, with this process being repeated until multiple complete text passages were generated. Finally, the sequence exhibiting the highest likelihood was selected, and the GoEmotions model was applied to extract an emotion probability distribution, thereby enabling the decoding of emotional content throughout the semantic perception process [31].

2.6. Evaluation Metrics

2.6.1. Semantic Similarity

BERTScore, word error rate (WER) and Euclidean distance (ED) were employed to quantify how closely the decoded text matches the original stimulus.

BERTScore measures semantic overlap by aligning contextual embeddings of the candidate and reference texts to compute precision, recall and F1 score [42]. Precision reflects the matching quality between the two texts. Recall reflects the extent of matching coverage. The F1 score balances both quality and coverage. The metrics are defined as follows:

$$\text{Precision} = \frac{1}{M}\sum_{i=1}^{M} \max_{j\in\{1,\ldots,N\}} \cos(a_i, b_j), \quad (1)$$

 doi:10.20944/preprints202508.1518.v1

$$\text{Recall} = \frac{1}{N}\sum_{j=1}^{N} \max_{i\in\{1,\ldots,M\}} \cos(b_j, a_i), \tag{2}$$

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}, \tag{3}$$

where $A = \{a_1, a_2, \ldots, a_M\}$ denotes the set of candidate text vectors, and $B = \{b_1, b_2, \ldots, b_M\}$ denotes the set of reference text vectors.

ED computes the Euclidean distance between corresponding semantic vectors, where smaller values indicate greater vector proximity and consequently lower semantic reconstruction error [43]. The metric is defined as follows:

$$\text{ED} = \sqrt{\sum_{i=1}^{N}(a_i - b_i)^2}, \tag{4}$$

where $A = \{a_1, a_2, \ldots, a_M\}$ denotes the set of candidate text vectors, and $B = \{b_1, b_2, \ldots, b_M\}$ denotes the set of reference text vectors.

Word error rate (WER) measures, at the word level, the proportion of insertion, deletion, and substitution errors between the predicted text and the reference text relative to the total number of words in the reference [44]. The metric is defined as follows:

$$\text{WER} = \frac{S + D + I}{N}, \tag{5}$$

where S is the number of substitutions in the candidate text, D the number of deletions (words present in the reference but missing in the candidate), I the number of insertions (extra words in the candidate), N the total number of words in the reference text.

2.6.2. Emotional Similarity

For emotion decoding, the GoEmotions framework was applied to extract normalized emotion probability distributions from the decoded text [31].

Cosine Similarity (CS) and Jensen-Shannon (JSS) similarities were then computed between two sets of comparisons: (a) 30 randomly sampled decoded emotion distributions and the true distributions, and (b) 30 randomly generated emotion distributions and true distributions. The emotion-decoding performance of SED-GPT was quantified through direct comparison of these similarity scores.

CS is defined as the cosine of the angle between the predicted and true probability vectors, reflecting the overall alignment of the distributions [45]. To evaluate emotion similarity, we employed the following formula:

$$CS(P, Q) = \frac{P^{\top}Q}{\| P \| \| Q \|}, \tag{6}$$

where the normalized emotion probability distribution of the decoded text is $P = (p_1, \ldots, p_n)$, and the normalized emotion probability distribution of the random text is $Q = (q_1, \ldots, q_n)$.

JSS is defined as one minus the Jensen–Shannon divergence between the two distributions. JSS measures the similarity between the predicted emotion distribution and the true distribution [46]. To evaluate semantic similarity, we employed the following formula:

$$JSS(P, Q) = 1 - \sqrt{\frac{1}{2}\sum_{i=1}^{n}\left(p_i \log_2 \frac{2p_i}{p_i + q_i} + q_i \log_2 \frac{2q_i}{p_i + q_i}\right)}, \tag{7}$$

where the normalized emotion probability distribution of the decoded text is $P = (p_1, \ldots, p_n)$, and the normalized emotion probability distribution of the random text is $Q = (q_1, \ldots, q_n)$.

The similarity between the generated text and the reference text was quantified for each emotion category using GoEmotions with multi-emotion classifications. A similarity metric was defined through computation of the normalized ratio for each emotion category. The closer this ratio is to 1,

 doi:10.20944/preprints202508.1518.v1

the more consistent the predicted probabilities for that emotion between the two texts [47]. To evaluate semantic similarity, we employed the following formula:

$$\mathrm{Sim}(p_i, q_i) = \frac{\min(p_i, q_i)}{\max(p_i, q_i)}, \tag{8}$$

where the normalized emotion probability distribution of the decoded text is $P = (p_1, \ldots, p_n)$, and the normalized emotion probability distribution of the random text is $Q = (q_1, \ldots, q_n)$.

3. Results

3.1. Brain Activation and Functional Connectivity of Chinese Semantic Perception

3.1.1. Brain Activation of Chinese Semantic Perception

Comparative results between task-state and resting-state activation during the Chinese semantic perception task are shown in Figure 2 and Table 1.

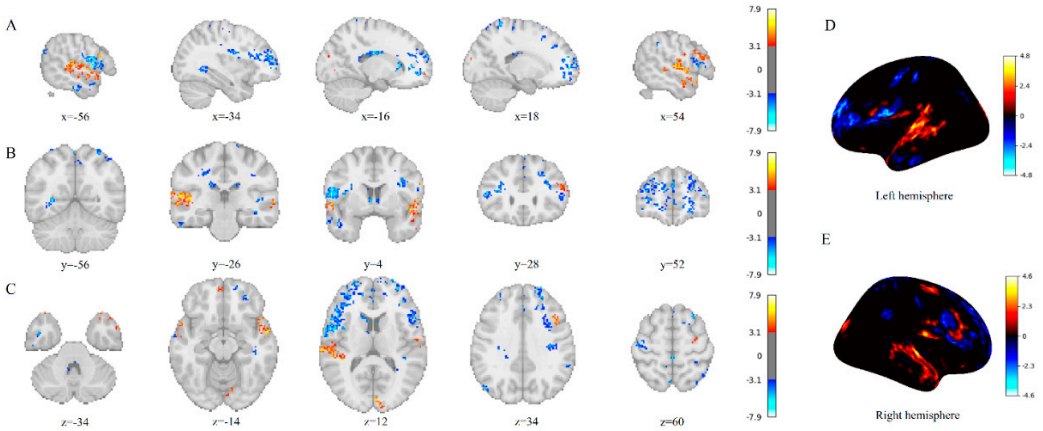

Figure 2. Activation comparison between Chinese semantic perception task state and resting state. (A) Sagittal view; (B) Coronal view; (C) Axial view; (D) Left hemisphere; (E) Right hemisphere.

Table 1. Activation and deactivation clusters in response to task versus resting state.

Regions	x	y	z	Z-peak	size
Activated brain regions					
FP	-26	62	-8	6.15	2225
STGpd	-60	-10	8	7.26	990
STGpd	56	4	-2	7.89	880
FP	26	68	-2	5.26	520
MFG	52	18	24	6.20	180
FP	-56	4	-2	6.41	112
Deactivated brain regions					
FP	36	10	34	-3.09	730
SPL	30	-40	78	-3.09	346
PCunC	0	-38	52	-3.09	280
PoCG	-34	-48	76	-3.09	269
ParaCG	0	44	14	-3.10	249
PreCG	-52	-14	50	-3.09	236

 doi:10.20944/preprints202508.1518.v1

During Chinese semantic processing task, significant activation was observed in the classical language network and semantic processing brain regions, accompanied by deactivation patterns in the default mode network (DMN) and primary sensorimotor cortices. Specifically, enhanced neural activity was identified in the bilateral frontal poles and posterior superior temporal gyri.

3.1.2. Functional Connectivity of Chinese Semantic Perception

Based on the above activation results, we selected bilateral frontal poles and posterior superior temporal gyri as seed ROIs and performed PPI analysis to investigate the functional connectivity mechanisms underlying semantic processing, as shown in Figure 3 and Table 2.

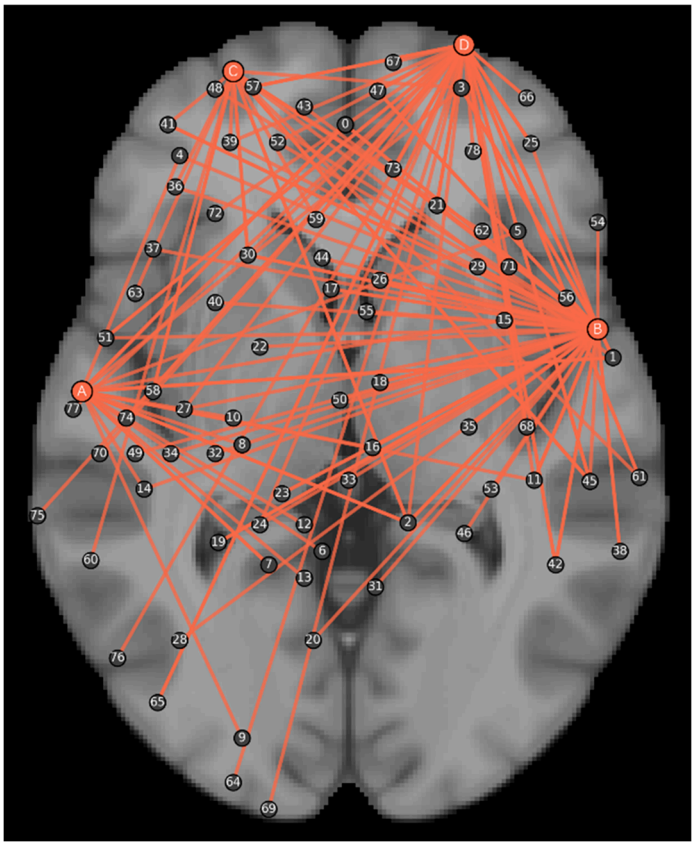

Figure 3. PPI network analysis results. A represents the posterior part of the left superior temporal gyrus, B represents the posterior part of the right superior temporal gyrus, C represents the left frontal pole, and D represents the right frontal pole, and points with similar projection distances in the transverse plane are classified into the same cluster.

Table 2. PPI functional connectivity index table.

Index	Regions	Index	Regions	Index	Regions
0	FMC, ParaCG	27	PreCG	54	IFGpt
1	MTG, PoCG, STG	28	LOCsup	55	ACC
2	PCG	29	MFG	56	IFGoper
3	FP	30	FOC, SFG	57	FP
4	FP	31	PCunC	58	COC, PreCG
5	TP	32	PreCG	59	FOC, SFG
6	PoCG	33	PCC	60	MTGto
7	SPL	34	PoCG	61	ITGpd, MTGpd

 doi:10.20944/preprints202508.1518.v1

8	PreCG	35	PreCG	62	MFG
9	LOCsup	36	FOC, MFG	63	IFGoper
10	PreCG	37	IFGpt	64	OP
11	SCC, PoCG	38	ITGto	65	LOCsup
12	PoCG	39	FP	66	FP
13	PCunC	40	MFG	67	FP
14	PoCG	41	FP	68	IC, PoCG
15	COC, FP	42	AG	69	OP
16	PCG, PreCG	43	ParaCG	70	SCC
17	ACC, ParaCG	44	SFG	71	IC
18	SMA	45	pITG, pSMG	72	FOC
19	SPL	46	SPL	73	ACC, ParaCG
20	PCunC	47	FP	74	COC
21	FP, SFG	48	FP	75	MTGpd
22	SFG	49	PoCG	76	LOCsup
23	PoCG	50	ACC	77	PoCG
24	PoCG	51	MTGad, PreCG	78	FP
25	FP	52	FP, ParaCG		
26	ParaCG, SFG	53	PoCG		

The PPI analysis revealed that core semantic processing regions exhibited significant task-state functional coupling with distributed brain areas. These semantic processing regions showed significant functional coupling with the 79 distinct connectivity clusters (FDR, $p<0.001$) including paracingulate gyrus, anterior cingulate cortex and insular cortex.

These connections support processes such as semantic retrieval, contextual and narrative maintenance, motor simulation, attentional control, visual imagery and spatial scene construction, and affective and self-referential processing [48–54].

3.2. Brain Regions Activated by Emotional Words

The comparative activation results between each category of emotion words and neutral words during Chinese semantic perception tasks are shown in Table 3.

Table 3. Activation and deactivation in response to emotional versus neutral words.

Regions	x	y	z	Z-peak	Size	Regions	x	y	z	Z-peak	Size
Joy						FP	30	36	50	-3.52	15
PreCG	-62	-4	22	3.31	16	**Fear**					
LOCid	-52	-68	6	3.31	13	LOCs	-14	-64	66	3	14
ParaCG	0	52	10	4.18	17	LOCs	18	-64	58	3.77	13
SPL	42	-50	62	3.93	18	FP	-30	62	-6	3.33	12
FP	46	50	4	3.57	24	**Disgust**					
FP	48	44	6	4.1	14	FP	34	62	8	3.4	11
PreCG	52	-10	40	3.29	15	PostCG	44	-32	40	3.22	13
Affection						**Boredom**					
SFG	-14	-6	74	3.39	11	SFG	-18	26	62	3.44	12

 doi:10.20944/preprints202508.1518.v1

CunC	16	-76	28	3.24	11	FP	-46	36	14	-3.2	15
Longing						LOCs	-26	-66	38	-3.55	11
IFGpt	-54	34	0	-3.29	11	**Guilt**					
PCunC	-8	-56	42	-4.34	13	SFG	-26	62	-8	4.04	30
PCunC	2	-68	30	-3.32	14	SMGp	36	10	34	-3.47	18
Trust						**Disappointment**					
AG	-58	-56	28	3.81	33	LOCid	44	-78	8	3.21	12
Praise						MFG	-50	34	34	-3.4	18
SMGa	-54	-30	50	3.51	15	FP	48	48	2	-3.73	11
LOCid	-54	-68	6	4.41	17	**Doubt**					
PostCG	-30	-40	70	4.39	35	MFG	-26	32	50	-3.38	15
SFG	-10	-6	78	3.6	19	PCunC	-4	-64	54	-3.65	12
LOCs	-8	-60	68	3.52	20	FP	16	48	44	-4.32	16
SFG	6	-4	78	4.27	21	FP	14	64	18	-3.55	15
LOCs	12	-58	66	3.6	35	**Criticism**					
SPL	22	-46	76	3.55	21	Cereb.L	-22	-90	-28	3.72	11
LOCid	52	-66	-10	3.35	15	FP	-16	58	24	3.45	12
STGp	-66	-42	14	-3.94	30	Cereb.R	24	-92	-28	3.31	12
MTGto	-60	-46	6	-4.51	104	ITGto	-60	-50	-12	-3.4	16
MTGto	-64	-48	6	-4.18	18	SMGp	-56	-48	52	-4.06	32
MFG	-48	10	44	-5.27	320	ITGto	-58	-40	-12	-3.49	16
AG	-54	-62	22	-4.07	16	COC	-58	-22	12	-3.93	12
MTGp	-54	-26	-6	-4.48	37	PreCG	-48	6	36	-4.12	26
PreCG	-46	0	52	-4.46	21	FP	-48	38	12	-3.9	57
LOCs	-40	-74	48	-4.14	42	SMGp	-50	-46	52	-3.58	12
OP	-6	-94	0	-4.23	48	PreCG	-46	0	52	-3.21	15
LOCs	32	-76	40	-3.65	26	SMGp	-44	-50	46	-4.59	16
LOCs	40	-64	54	-3.87	21	LOCs	-26	-72	42	-4.87	130
LOCs	38	-60	46	-3.06	16	SMGa	-40	-24	38	-3.31	11
MFG	44	14	34	-4.7	18	LOCs	-34	-60	42	-4.59	43
MTGto	60	-44	-10	-2.97	15	SFG	4	18	58	-3.01	12
Goodwill						OP	28	-100	4	-4.1	37
MFG	-44	12	56	3.66	12	OP	26	-94	10	-3.08	13
MFG	-30	28	50	3.23	12	LOCs	30	-70	44	-3.31	11
MFG	-30	30	44	3.25	13	LOCs	40	-60	50	-3.6	35
Sadness						AG	44	-46	42	-4.19	42
PCunC	-12	-74	38	-3.06	12	MFG	40	32	36	-3.59	22
PCunC	0	-50	64	-3.18	18	FP	46	42	28	-3.78	25
PCunC	0	-44	46	-2.99	11	MFG	46	22	28	-3.08	14
PCunC	6	-62	54	-3.33	11	IFGop	52	14	26	-4.26	17
LOCs	30	-62	60	-3.23	29	SMGp	58	-44	42	-3.7	22

 doi:10.20944/preprints202508.1518.v1

Under conditions of continuous natural language stimulation, all emotion categories elicit widely distributed neural activation patterns. Specifically, the processing of emotional words engaged not only classic limbic regions associated with affective processing, but also significantly activated: primary visual cortex, sensorimotor cortices, facial expression-modulation regions and high-level cognitive cortices [55–57].

3.3. Semantic Decoding Performance

To assess semantic decoding performance, the text outputs generated by the decoder were quantitatively compared with both the original stimulus texts (ground truth) and randomly generated control texts, as shown in Table 4.

Table 4. Comparison of semantic decoding performance.

Metrics	EXP	RM	U	P
BERT	0.650 ± 0.151	0.326 ± 0.074	5872	p<0.001
ED	12.432 ± 2.896	14.528 ± 0.673	1767	p<0.001
WER	0.924 ± 0.028	0.989 ± 0.051	862	p<0.001

EXP denotes the experimental group. RM denotes the random group. BERT denotes BERTScore F1.

Across all three metrics, the differences between the experimental and random groups were highly significant ($p < 0.001$). This indicates that the text produced by our semantic decoder significantly outperformed the random baseline at capturing and reconstructing semantic information in long Chinese narratives.

3.4. Emotion Decoding Performance

The emotion recognition performance was quantitatively assessed by computing both CS and JSS between the decoded emotion distributions and the corresponding true distributions. These metric values were compared against a random baseline condition, as shown in Table 5.

Table 5. Comparison of Emotion Recognition Similarity Metrics.

Metrics	EXP	RM	U	P
CS	0.504 ± 0.348	0.233 ± 0.248	645	p<0.05
JSS	0.469 ± 0.227	0.323 ± 0.148	620	p<0.05

EXP denotes the experimental group. RM denotes the random group.

For emotion recognition evaluation, we conducted multidimensional affective analysis of the decoded results and compared them with random baseline. Overall, the experimental group demonstrated significantly higher scores in both CS and JSS compared to the random group ($p < 0.05$).

The normalized ratios between the emotion distributions of decoded texts and the true emotion distributions were calculated. These ratios were then compared against random baseline distributions, as detailed in Table 6.

Table 6. Comparison of fine-grained emotion decoding.

Emotions	EXP	RM	U	p
admiration	0.128 ± 0.208	0.313 ± 0.299	238	p<0.05
amusement	0.311 ± 0.244	0.211 ± 0.154	539	p=0.191
anger	0.246 ± 0.277	0.060 ± 0.117	654	p<0.05
annoyance	0.429 ± 0.312	0.381 ± 0.325	488	p=0.579

 doi:10.20944/preprints202508.1518.v1

approval	0.323 ± 0.317	0.399 ± 0.255	352	p=0.149
caring	0.260 ± 0.273	0.088 ± 0.098	589.5	p<0.05
confusion	0.527 ± 0.304	0.335 ± 0.220	609	p<0.05
curiosity	0.201 ± 0.258	0.175 ± 0.165	388	p=0.363
desire	0.350 ± 0.255	0.213 ± 0.215	623	p<0.05
disappointment	0.200 ± 0.212	0.146 ± 0.245	539	p=0.191
disapproval	0.293 ± 0.232	0.322 ± 0.289	470.5	p=0.767
disgust	0.506 ± 0.277	0.233 ± 0.242	699	p<0.001
embarrassment	0.442 ± 0.302	0.175 ± 0.152	672.5	p<0.001
excitement	0.354 ± 0.217	0.291 ± 0.171	531	p=0.234
fear	0.420 ± 0.292	0.119 ± 0.150	732	p<0.001
gratitude	0.271 ± 0.302	0.232 ± 0.241	456	p=0.935
grief	0.302 ± 0.287	0.052 ± 0.091	772.5	p<0.001
joy	0.344 ± 0.324	0.145 ± 0.187	646	p<0.01
love	0.212 ± 0.256	0.061 ± 0.098	622	p<0.05
nervousness	0.107 ± 0.107	0.028 ± 0.033	760	p<0.001
neutral	0.358 ± 0.308	0.096 ± 0.174	733	p<0.001
optimism	0.236 ± 0.242	0.231 ± 0.228	429	p=0.762
pride	0.192 ± 0.176	0.199 ± 0.244	471	p=0.762
realization	0.035 ± 0.059	0.039 ± 0.060	450.5	p=0.997
relief	0.331 ± 0.284	0.255 ± 0.219	497	p=0.492
remorse	0.276 ± 0.226	0.120 ± 0.153	659	p<0.01
sadness	0.345 ± 0.362	0.078 ± 0.186	699.5	p<0.001
surprise	0.370 ± 0.303	0.208 ± 0.210	579	p=0.057

EXP denotes the experimental group. RM denotes the random group.

The decoding accuracy rates for anger, disgust, embarrassment, fear, grief, joy, nervousness, neutral, remorse, sadness, caring, confusion, desire and love in the experimental group was significantly above the random baseline ($p < 0.05$). These results underscore the decoder's robust sensitivity to a wide spectrum of emotional states under naturalistic language conditions.

4. Discussion

Conventional emotion decoding methods are constrained by factors such as short-sequence stimuli, coarse-grained categories, and low-context language systems, making it difficult to capture the interaction mechanisms between language and emotion in real-world scenarios [16]. Our study demonstrates fine-grained emotion decoding in long-sequence narratives and provides a methodological basis for investigating the neural mechanisms underlying the interaction between emotion and semantic processing. It may address the bottleneck in dynamic emotion monitoring for depression treatment and cognitive behavioral therapy (CBT) [58].

In this research, we introduce SED-GPT, a fine-grained emotion decoding framework designed for long-sequence Chinese language processing, which establishes a neural alignment between brain activity and LLM-based semantic vector representations to enable inner speech semantic reconstruction. Moreover, our findings reveal the dynamic interplay between language-related cortical networks and affective neural systems during Chinese emotional semantic processing, providing novel neurocognitive evidence for thought-emotion integration mechanisms.

 doi:10.20944/preprints202508.1518.v1

For semantic decoding, SED-GPT achieved a BERTScore F1 of 0.65, an ED of 12.432 and a WER of 0.924. For emotion decoding, it attained a CS of 0.504 and a JSS of 0.469. All decoding performances significantly surpass random baseline.

The GLM results from both task and resting states demonstrated that during Chinese semantic processing, the coordinated activation of the left frontopolar cortex and right medial frontopolar cortex likely involves top-down attentional modulation and cross-modal semantic integration [59]. Bilateral posterior superior temporal gyri were engaged in acoustic feature analysis and semantic primitive extraction, and bilateral temporal poles participated in abstract semantic representation and integration of social context [60,61]. Concurrently, the suppression of the DMN and other non-task regions (e.g., the default mode network) indicates directed allocation of cognitive resources to core language networks [62]. Activity along the left central sulcus during Chinese lexical ambiguity resolution suggests enhanced phonological working memory [63]. The co-activation of the working memory network (left superior frontal gyrus to right superior parietal lobule) with the attentional network (right precuneus) supports context integration and interference suppression [64]. Cross-modal activation of the left middle occipital gyrus implies that orthography–phonology associations may facilitate automatic mapping from speech to visual representations [65]. In higher level comprehension (e.g., discourse level processing), bilateral frontal poles and middle frontal gyri support context maintenance and are implicated in controlled semantic retrieval [66,67].

PPI analyses showed that the right-hemisphere target seed region exhibited significant functional coupling with the paracingulate gyrus, anterior cingulate cortex (ACC), and insular cortex. From a network perspective, this model can be reasonably explained theoretically. The paracingulate gyrus is primarily involved in self-monitoring and reality monitoring processes [68], the anterior cingulate integrates negative emotions, pain and cognitive control [69], and the insula enriches semantic understanding by integrating abstract semantic information with interoceptive bodily states and socio-emotional experience [70]. This task-dependent functional coupling mechanism may collectively support the multi-level synergistic integration of semantic content with contextual and emotional cues during narrative comprehension [71,72].

The GLM contrast between emotional and neutral words suggests:

- Network reorganization and resource redistribution. Emotional words were associated with widespread changes across cortical networks. High social value words (e.g., praise) activate an integrated network of empathic, motor simulation and evaluation, while activity in general executive control regions is significantly reduced [57].
- Embodied simulation mechanism. In the joy condition, significant activation was observed in the left and right precentral gyrus, which may involve the recruitment of oral and facial motor representations. In the praise condition, notable activation was detected in the postcentral gyrus and the anterior supramarginal gyrus, reflecting the engagement of somatosensory and speech-related pathways. These findings support the notion that abstract emotions participate in sensory-motor representation mapping [56].
- The emotion-visual imagery coupling mechanism. Positive emotions (e.g., joy, admiration) and negative emotions (e.g., fear) were associated with activation in the inferior occipital cortex, cuneus and right higher-order visual areas, revealing the multi-level integration of vivid mental imagery with emotional processing [73].
- The self/others reference and value assessment mechanism. Emotional words (e.g., admiration) activated the dorsomedial prefrontal cortex and superior frontal gyrus, reflecting metacognitive simulation of self and others in complex social emotions [74].
- Suppression patterns of specific emotions. Emotional words (e.g., criticism) elicited large-scale deactivation in visual-semantic, sensorimotor and executive/metacognitive networks (e.g., middle frontal gyrus, frontal pole). This pattern may reflect down-regulation of DMN processing to concentrate cognitive resources on affective synesthesia and social reasoning network [75].

Fine-grained emotion decoding results indicate that the proposed method can distinguish 14 emotional states from brain activity. Notably, negative emotions (anger, disgust, fear, grief, sadness)

 doi:10.20944/preprints202508.1518.v1

exhibit greater decoding accuracy, which may be attributed to humans' preferential processing of negative stimuli. During disgust-related words processing, participants experienced embodied interoceptive imagery (e.g., nausea), activating the right somatosensory cortex (postcentral gyrus) [76]. Concurrently, stronger engagement of higher-order cognitive regions (e.g., right frontal pole) was required to evaluate and regulate this negative emotional response (e.g., suppressing the "gagging" impulse) [76]. Sadness-related words can cause significant suppression in right lateral occipital cortex, right precuneus and right frontal pole. It indicates that the subjects exhibited weakened imagery, spatial association, and metacognitive processing under sad emotions [77]. Humans allocate more attentional, perceptual, learning and memory resources to negative stimuli, resulting in stronger and more stable neural responses with a higher signal-to-noise ratio, which facilitates emotion decoding [78].

5. Conclusions

In this study, we achieved the textual reconstruction of semantic content by aligning fMRI signals induced by Chinese auditory narratives with semantic representations and further extracted multidimensional emotional features for classification. Our findings not only demonstrate the neural decoding of complex emotional components in the Chinese context but also highlight the immense potential of integrating large-scale language models with neuroimaging technology. Overall, this approach could advance the development of affective brain-computer interfaces (BCIs) and show promising applications in the treatment of depression and cognitive behavioral therapy.

Author Contributions: Conceptualization, W.C. and L.M.; methodology, W.C. and Z.W.; software, W.C and Z.W.; validation, W.C and Z.W.; formal analysis, W.C and Z.W.; investigation, W.C. and L.M.; resources, L.M.; data curation, W.C and Z.W.; writing—original draft preparation, W.C.; writing—review and editing, W.C and Z.W.; visualization, W.C.; supervision, L.M.; project administration, L.M.; W.C. and Z.W. contributed equally to this work and should be considered co-first author. All authors have read and agreed to the published version of the manuscript.

Funding: Not applicable.

Institutional Review Board Statement: This study used publicly available dataset, and the original data had already obtained ethical approval and informed consent from the participants.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used and analyzed in the current study are available from https://openneuro.org/datasets/ds004078/versions/1.2.1.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

FP	Frontal Pole
PreCG	Precentral Gyrus
PoCG	Postcentral Gyrus
MFG	Middle Frontal Gyrus
SFG	Superior Frontal Gyrus
IFGpt	Inferior Frontal Gyrus, pars triangularis
IFGoper	Inferior Frontal Gyrus, pars opercularis
ACC	Anterior Cingulate Cortex
ParaCG	Paracingulate Gyrus
PCC	Posterior Cingulate Cortex
PCG	Posterior Cingulate Gyrus
SMA	Supplementary Motor Area

 doi:10.20944/preprints202508.1518.v1

FMC	Frontomedial Cortex
FOC	Frontal Orbital Cortex
COC	Central Opercular Cortex
SCC	Supracalcarine Cortex
LOCsup	Lateral Occipital Cortex, superior division
LOCid	Lateral Occipital Cortex, inferior division
LOCs	Lateral Occipital Cortex, superior division
OP	Occipital Pole
CunC	Cuneal Cortex
PCunC	Precuneous Cortex
MTG	Middle Temporal Gyrus
MTGto	Middle Temporal Gyrus, temporooccipital part
MTGpd	Middle Temporal Gyrus, posterior division
MTGad	Middle Temporal Gyrus, anterior division
STG	Superior Temporal Gyrus
STGpd	Superior Temporal Gyrus, posterior division
TP	Temporal Pole
ITGto	Inferior Temporal Gyrus, temporooccipital part
ITGpd	Inferior Temporal Gyrus, posterior division
AG	Angular Gyrus
SMGa	Supramarginal Gyrus, anterior division
SMGp	Supramarginal Gyrus, posterior division

References

1. Fernyhough, C.; Borghi, A.M. Inner speech as language process and cognitive tool. Trends Cogn. Sci. 2023, 27, 1180–1193. https://doi.org/10.1016/j.tics.2023.08.014.
2. Nummenmaa, L.; Saarimäki, H.; Glerean, E.; et al. Emotional speech synchronizes brains across listeners and engages large-scale dynamic brain networks. Neuroimage 2014, 102, 498–509. https://doi.org/10.1016/j.neuroimage.2014.07.063.
3. Etkin, A.; Egner, T.; Kalisch, R. Emotional processing in anterior cingulate and medial prefrontal cortex. Trends Cogn. Sci. 2011, 15, 85–93. https://doi.org/10.1016/j.tics.2010.11.004.
4. Devinsky, O.; Morrell, M.J.; Vogt, B.A. Contributions of anterior cingulate cortex to behaviour. Brain 1995, 118, 279–306. https://doi.org/10.1093/brain/118.1.279.
5. Binder, J.R.; Conant, L.L.; Humphries, C.J.; et al. Toward a brain-based componential semantic representation. Cogn. Neuropsychol. 2016, 33, 130–174. https://doi.org/10.1080/02643294.2016.1147426.
6. Lenci, A.; Lebani, G.E.; Passaro, L.C. The emotions of abstract words: A distributional semantic analysis. Top. Cogn. Sci. 2018, 10, 550–572. https://doi.org/10.1111/tops.12335.
7. Satpute, A.B.; Lindquist, K.A. At the neural intersection between language and emotion. Affective Sci. 2021, 2, 207–220. https://doi.org/10.1007/s42761-021-00032-2.
8. Gaillard, R.; Del Cul, A.; Naccache, L.; et al. Nonconscious semantic processing of emotional words modulates conscious access. Proc. Natl. Acad. Sci. USA 2006, 103, 7524–7529. https://doi.org/10.1073/pnas.0600584103.
9. Kuperberg, G.R.; Deckersbach, T.; Holt, D.J.; et al. Increased temporal and prefrontal activity in response to semantic associations in schizophrenia. Arch. Gen. Psychiatry 2007, 64, 138–151. https://doi.org/10.1001/archpsyc.64.2.138.
10. Zhu, X.; Guo, C.; Feng, H.; et al. A review of key technologies for emotion analysis using multimodal information. Cogn. Comput. 2024, 16, 1504–1530. https://doi.org/10.1007/s12559-024-10287-z.
11. Zhang, Y.; Li, Y.; Yu, Z.; et al. Decoding the flow: Causemotion for emotional causality analysis in long-form conversations. arXiv 2025, arXiv:2501.00778.
12. Chaudhary, U. Non-invasive brain signal acquisition techniques: Exploring EEG, EOG, fNIRS, fMRI, MEG, and fUS. In Expanding Senses Using Neurotechnology: Volume 1—Foundation of Brain-Computer Interface Technology; Springer Nature: Cham, Switzerland, 2025; pp. 25–80. https://doi.org/10.1007/978-3-031-76081-5_2.

 doi:10.20944/preprints202508.1518.v1

13. Winter, W.R.; Nunez, P.L.; Ding, J.; et al. Comparison of the effect of volume conduction on EEG coherence with the effect of field spread on MEG coherence. Stat. Med. 2007, 26, 3946–3957. https://doi.org/10.1002/sim.2978.
14. Wilcox, T.; Biondi, M. fNIRS in the developmental sciences. Wiley Interdiscip. Rev. Cogn. Sci. 2015, 6, 263–283. https://doi.org/10.1002/wcs.1343.
15. deCharms, C.R. Applications of real-time fMRI. Nat. Rev. Neurosci. 2008, 9, 720–729. https://doi.org/10.1038/nrn2414.
16. Kassam, K.S.; Markey, A.R.; Cherkassky, V.L.; et al. Identifying emotions on the basis of neural activation. PLoS ONE 2013, 8, e66032. https://doi.org/10.1371/journal.pone.0066032.
17. Saarimäki, H.; Gotsopoulos, A.; Jääskeläinen, I.P.; et al. Discrete neural signatures of basic emotions. Cereb. Cortex 2016, 26, 2563–2573. https://doi.org/10.1093/cercor/bhv125.
18. Kragel, P.A.; LaBar, K.S. Decoding the nature of emotion in the brain. Trends Cogn. Sci. 2016, 20, 444–455. https://doi.org/10.1016/j.tics.2016.03.011.
19. Wei, J.; Tay, Y.; Bommasani, R.; et al. Emergent abilities of large language models. arXiv 2022, arXiv:2206.07682.
20. Tang, J.; LeBel, A.; Jain, S.; et al. Semantic reconstruction of continuous language from non-invasive brain recordings. Nat. Neurosci. 2023, 26, 858–866. https://doi.org/10.1038/s41593-023-01304-9.
21. Ye, Z.; Ai, Q.; Liu, Y.; et al. Generative language reconstruction from brain recordings. Commun. Biol. 2025, 8, 346. https://doi.org/10.1038/s42003-025-07731-7.
22. Liu, P.; Dong, G.; Guo, D.; et al. A survey on fMRI-based brain decoding for reconstructing multimodal stimuli. arXiv 2025, arXiv:2503.15978.
23. Wang, S.; Zhang, X.; Zhang, J.; et al. A synchronized multimodal neuroimaging dataset for studying brain language processing. Sci. Data 2022, 9, 590. https://doi.org/10.1038/s41597-022-01559-8.
24. Pajula, J.; Tohka, J. How many is enough? Effect of sample size in inter-subject correlation analysis of fMRI. Comput. Intell. Neurosci. 2016, 2016, 2094601. https://doi.org/10.1155/2016/2094601.
25. Baker, D.H.; Vilidaite, G.; Lygo, F.A.; et al. Power contours: Optimising sample size and precision in experimental psychology and human neuroscience. Psychol. Methods 2021, 26, 295. https://doi.org/10.1037/met0000337.
26. Di, X.; Zhang, Z.; Biswal, B.B. Understanding psychophysiological interaction and its relations to beta series correlation. Brain Imaging Behav. 2021, 15, 958–973. https://doi.org/10.1007/s11682-020-00330-2.
27. Roiser, J.P.; Linden, D.E.; Gorno-Tempinin, M.L.; et al. Minimum statistical standards for submissions to Neuroimage: Clinical. Neuroimage Clin. 2016, 12, 1045. https://doi.org/10.1016/j.nicl.2016.08.004.
28. Xu, L.; Lin, H.; Pan, Y.; et al. Constructing the affective lexicon ontology. J. China Soc. Sci. Tech. Inf. 2008, 27, 180–185. https://doi.org/10.3969/j.issn.1000-0135.2008.02.004.
29. Ge, J.; Gao, J.H. A review of functional MRI application for brain research of Chinese language processing. Magn. Reson. Lett. 2023, 3, 1–13. https://doi.org/10.1016/j.mrl.2022.12.001.
30. Zhang, Q.; Wang, H.; Luo, C.; et al. The neural basis of semantic cognition in Mandarin Chinese: A combined fMRI and TMS study. Hum. Brain Mapp. 2019, 40, 5412–5423. https://doi.org/10.1002/hbm.24781.
31. Demszky, D.; Movshovitz-Attias, D.; Ko, J.; et al. GoEmotions: A dataset of fine-grained emotions. arXiv 2020, arXiv:2005.00547.
32. Si, C.; Zhang, Z.; Chen, Y.; et al. Sub-character tokenization for Chinese pretrained language models. Trans. Assoc. Comput. Linguist. 2023, 11, 469–487. https://doi.org/10.1162/tacl_a_00560.
33. Zhang, Z.; Han, X.; Zhou, H.; et al. CPM: A large-scale generative Chinese pre-trained language model. AI Open 2021, 2, 93–99. https://doi.org/10.1016/j.aiopen.2021.07.001.
34. Ma, L.; Cui, W.; Yang, W.; et al. Noninvasive decoding and reconstruction of continuous Chinese language semantics. J. Data Acquis. Process. 2025, 40. https://doi.org/10.16337/j.1004-9037.2025.03.005.
35. Xun, G.R.E. Word Boundary Information and Chinese Word Segmentation.
36. Binder, J.R.; Desai, R.H.; Graves, W.W.; et al. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb. Cortex 2009, 19, 2767–2796. https://doi.org/10.1093/cercor/bhp055.

 doi:10.20944/preprints202508.1518.v1

37. Caucheteux, C.; Gramfort, A.; King, J.R. Disentangling syntax and semantics in the brain with deep networks. In Proceedings of the International Conference on Machine Learning; PMLR, 2021; pp. 1336–1348.
38. Hermann, K.M.; Kocisky, T.; Grefenstette, E.; et al. Teaching machines to read and comprehend. Adv. Neural Inf. Process. Syst. 2015, 28. https://doi.org/10.5555/2969239.2969428.
39. Jain, S.; Huth, A. Incorporating context into language encoding models for fMRI. Adv. Neural Inf. Process. Syst. 2018, 31. https://doi.org/10.5555/3327757.3327769.
40. Deniz, F.; Nunez-Elizalde, A.O.; Huth, A.G.; et al. The representation of semantic information across human cerebral cortex during listening versus reading is invariant to stimulus modality. J. Neurosci. 2019, 39, 7722–7736. https://doi.org/10.1523/JNEUROSCI.0675-19.2019.
41. Benara, V.; Singh, C.; Morris, J.X.; et al. Crafting interpretable embeddings for language neuroscience by asking LLMs questions. Adv. Neural Inf. Process. Syst. 2024, 37, 124137.
42. Lin, J.; Nogueira, R.; Yates, A. Pretrained Transformers for Text Ranking: Bert and Beyond; Springer Nature, 2022. https://doi.org/10.1007/978-3-031-02181-7.
43. Elmore, K.L.; Richman, M.B. Euclidean distance as a similarity metric for principal component analysis. Mon. Weather Rev. 2001, 129, 540–549. https://doi.org/10.1175/1520-0493(2001)129.
44. Ali, A.; Renals, S. Word error rate estimation for speech recognition: e-WER. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers); Association for Computational Linguistics (ACL), 2018; pp. 20–24. https://doi.org/10.18653/v1/P18-2004.
45. Xia, P.; Zhang, L.; Li, F. Learning similarity with cosine similarity ensemble. Inf. Sci. 2015, 307, 39–52. https://doi.org/10.1016/j.ins.2015.02.024.
46. Nielsen, F. On a generalization of the Jensen–Shannon divergence and the Jensen–Shannon centroid. Entropy 2020, 22, 221. https://doi.org/10.3390/e22020221.
47. Podani, J.; Ricotta, C.; Schmera, D. A general framework for analyzing beta diversity, nestedness and related community-level phenomena based on abundance data. Ecol. Complex. 2013, 15, 52–61.
48. Davey, J.; Thompson, H.E.; Hallam, G.; et al. Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. Neuroimage 2016, 137, 165–177.
49. Simony, E.; Honey, C.J.; Chen, J.; et al. Dynamic reconfiguration of the default mode network during narrative comprehension. Nat. Commun. 2016, 7, 12141.
50. Giacobbe, C.; Raimo, S.; Cropano, M.; et al. Neural correlates of embodied action language processing: A systematic review and meta-analytic study. Brain Imaging Behav. 2022, 16, 2353–2374. https://doi.org/10.1007/s11682-022-00680-3.
51. Piai, V.; Roelofs, A.; Acheson, D.J.; et al. Attention for speaking: Domain-general control from the anterior cingulate cortex in spoken word production. Front. Hum. Neurosci. 2013, 7, 832.
52. Epstein, R.A. Parahippocampal and retrosplenial contributions to human spatial navigation. Trends Cogn. Sci. 2008, 12, 388–396.
53. Jackson, R.L.; Hoffman, P.; Pobric, G.; et al. The semantic network at work and rest: differential connectivity of anterior temporal lobe subregions. J. Neurosci. 2016, 36, 1490–1501. https://doi.org/10.1523/JNEUROSCI.2999-15.2016.
54. Gennari, S.P.; Millman, R.E.; Hymers, M.; et al. Anterior paracingulate and cingulate cortex mediates the effects of cognitive load on speech sound discrimination. Neuroimage 2018, 178, 735–743.
55. Ballotta, D.; Maramotti, R.; Borelli, E.; et al. Neural correlates of emotional valence for faces and words. Front. Psychol. 2023, 14, 1055054.
56. Hauk, O.; Johnsrude, I.; Pulvermüller, F. Somatotopic representation of action words in human motor and premotor cortex. Neuron 2004, 41, 301–307. https://doi.org/10.1016/S0896-6273(03)00838-9.
57. Citron, F.M.M. Neural correlates of written emotion word processing: A review of recent electrophysiological and hemodynamic neuroimaging studies. Brain Lang. 2012, 122, 211–226.
58. Ritchey, M.; Dolcos, F.; Eddington, K.M.; et al. Neural correlates of emotional processing in depression: Changes with cognitive behavioral therapy and predictors of treatment response. J. Psychiatr. Res. 2011, 45, 577–587. https://doi.org/10.1016/j.jpsychires.2010.09.007.

 doi:10.20944/preprints202508.1518.v1

59. van ckeren, M.J.; Rueschemeyer, S.A. Cross-modal integration of lexical-semantic features during word processing: Evidence from oscillatory dynamics during EEG. PLoS ONE 2014, 9, e101042. https://doi.org/10.1371/journal.pone.0101042.
60. Davey, J.; Thompson, H.E.; Hallam, G.; et al. Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. Neuroimage 2016, 137, 165–177. https://doi.org/10.1016/j.neuroimage.2016.05.051.
61. Poeppel, D.; Idsardi, W.J.; Van Wassenhove, V. Speech perception at the interface of neurobiology and linguistics. Philos. Trans. R. Soc. B Biol. Sci. 2008, 363, 1071–1086. https://doi.org/10.1098/rstb.2007.2160.
62. Eysenck, M.W.; Moser, J.S.; Derakshan, N.; et al. A neurocognitive account of attentional control theory: How does trait anxiety affect the brain's attentional networks? Cogn. Emot. 2023, 37, 220–237. https://doi.org/10.1080/02699931.2022.2159936.
63. Xue, G.; Dong, Q.; Jin, Z.; et al. Mapping of verbal working memory in nonfluent Chinese–English bilinguals with functional MRI. Neuroimage 2004, 22, 1–10.
64. Wu, C.Y.; Ho, M.H.R.; Chen, S.H.A. A meta-analysis of fMRI studies on Chinese orthographic, phonological, and semantic processing. Neuroimage 2012, 63, 381–391.
65. Booth, J.R.; Burman, D.D.; Meyer, J.R.; et al. Development of brain mechanisms for processing orthographic and phonologic representations. J. Cogn. Neurosci. 2004, 16, 1234–1249.
66. Yang, X.; Zhang, X.; Yang, Y.; et al. How context features modulate the involvement of the working memory system during discourse comprehension. Neuropsychologia 2018, 111, 36–44.
67. Perfetti, C.A.; Frishkoff, G.A. The neural bases of text and discourse processing. Handb. Neurosci. Lang. 2008, 2, 165–174.
68. Lavallé, L.; Brunelin, J.; Jardri, R.; et al. The neural signature of reality-monitoring: A meta-analysis of functional neuroimaging studies. Hum. Brain Mapp. 2023, 44, 4372–4389. https://doi.org/10.1002/hbm.26408.
69. Shackman, A.J.; Salomons, T.V.; Slagter, H.A.; et al. The integration of negative affect, pain and cognitive control in the cingulate cortex. Nat. Rev. Neurosci. 2011, 12, 154–167.
70. Craig, A.D. How do you feel—now? The anterior insula and human awareness. Nat. Rev. Neurosci. 2009, 10, 59–70. https://doi.org/10.1038/nrn2555.
71. Simony, E.; Honey, C.J.; Chen, J.; et al. Dynamic reconfiguration of the default mode network during narrative comprehension. Nat. Commun. 2016, 7, 12141. https://doi.org/10.1038/ncomms12141.
72. Vaccaro, A.G.; Scott, B.; Gimbel, S.I.; et al. Functional brain connectivity during narrative processing relates to transportation and story influence. Front. Hum. Neurosci. 2021, 15, 665319. https://doi.org/10.3389/fnhum.2021.665319.
73. Sabatinelli, D.; Fortune, E.E.; Li, Q.; et al. Emotional perception: Meta-analyses of face and natural scene processing. Neuroimage 2011, 54, 2524–2533. https://doi.org/10.1016/j.neuroimage.2010.10.011.
74. Immordino-Yang, M.H.; McColl, A.; Damasio, H.; et al. Neural correlates of admiration and compassion. Proc. Natl. Acad. Sci. USA 2009, 106, 8021–8026. https://doi.org/10.1073/pnas.0810363106.
75. Anticevic, A.; Cole, M.W.; Murray, J.D.; et al. The role of default network deactivation in cognition and disease. Trends Cogn. Sci. 2012, 16, 584–592. https://doi.org/10.1016/j.tics.2012.10.008.
76. Wicker, B.; Keysers, C.; Plailly, J.; et al. Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. Neuron 2003, 40, 655–664. https://doi.org/10.1016/S0896-6273(03)00679-2.
77. Maddock, R.J. The retrosplenial cortex and emotion: New insights from functional neuroimaging of the human brain. Trends Neurosci. 1999, 22, 310–316. https://doi.org/10.1016/S0166-2236(98)01374-5.
78. Baumeister, R.F.; Bratslavsky, E.; Finkenauer, C.; et al. Bad is stronger than good. Rev. Gen. Psychol. 2001, 5, 323–370. https://doi.org/10.1037/1089-2680.5.4.323.