

非侵入性连续中文语言语义解码与重建

马磊¹, 崔文浩², 杨汶汶¹, 王朝欣¹

(1. 南通大学信息科学技术学院, 南通 226019; 2. 伦敦国王学院精神病学、心理学和神经科学研究所, 伦敦 WC2R 2LS)

摘要: 语言是沟通和认知的基础, 大脑多功能区域通过复杂神经网络共同参与语言的感知、理解与生成, 深入探索中文语义解码的神经机制对于中文脑机接口 (Brain-computer interface, BCI) 的研究意义重大。本研究旨在构建一种基于功能性磁共振成像 (Functional magnetic resonance imaging, fMRI) 的长序列中文连续语义解码方法, 称为中文长序列连续语义解码器 (Chinese long-sequence continuous semantic decoder, CLCSD), 通过信号处理流程和算法优化, 实现连续中文语义的高效解码。CLCSD 包含神经响应降维、编码模型、语速模型和束搜索解码模型 4 个部分。神经响应降维通过皮层重建、图像配准和脑区划定等方法, 将 4 维脑响应数据降为 2 维矩阵。编码模型采用 L2 正则化回归 (岭回归) 建立刺激特征与脑响应之间的关系, 通过自举法估计噪声协方差以增强泛化。语速模型采用与编码模型类似的思路, 将脑响应特征映射到预测的语速。束搜索解码模型利用语言模型的先验概率和编码模型的似然概率, 通过束搜索生成最可能的语义序列。CLCSD 在公开数据集 SMN4Lang 上取得了 0.674 的 BERTScore, 高于其他长序列中文连续语义解码模型。本研究提出一种高效的长序列中文连续语义解码方法, 为中文脑机接口技术的发展提供理论基础和方法参考。

关键词: 脑机接口; 中文语义解码; 岭回归; 语速模型; 束搜索

中图分类号: TP181; Q189

文献标志码: A

Non-invasive Continuous Chinese Language Semantic Decoding and Reconstruction

MA Lei¹, CUI Wenhao², YANG Wenwen¹, WANG Zhaoxin¹

(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)

Abstract: Language is an important tool for communication and cognition. Multiple functional areas of the brain, connected through complex neural networks, jointly participate in the perception, comprehension, and production of language. Exploring the neural mechanisms of Chinese semantic decoding is crucial for the development of Chinese brain-computer interface (BCI). This study aims to establish a long-sequence continuous semantic decoding method based on fMRI data, termed Chinese long-sequence continuous semantic decoder (CLCSD). Through signal processing workflows and algorithm optimization, it seeks to achieve efficient decoding of continuous Chinese semantics. The CLCSD framework is composed of four components: neural response dimensionality reduction, an encoding model, a word rate model, and a beam search decoding model. Neural response dimensionality reduction is performed through cortical

reconstruction, image registration, and brain region parcellation to reduce four-dimensional brain response data into a two-dimensional matrix. The encoding model is constructed using L2-regularized regression (ridge regression) to establish the relationship between stimulus features and brain responses, with noise covariance estimated via bootstrapping to enhance generalization. The word rate model follows a similar approach to the encoding model, where brain response features are mapped to predicted word rate. The beam search decoding model uses the prior probability of the language model and likelihood probabilities of the encoding model to generate the most probable semantic sequence through beam search. On publicly available dataset SMN4Lang, CLCSD achieves a mean BERTScore of 0.674, outperforming other long-sequence Chinese continuous semantic decoding models. The proposed method provides an efficient long-sequence continuous Chinese semantic decoding approach, offering both theoretical foundations and methodological references for the advancement of Chinese BCI technologies.

Key words: brain-computer interface (BCI); Chinese semantic decoding; ridge regression; speech rate model; beam search

引言

语言是人类沟通和认知的重要方式,是自我意识的表达和人类认知复杂性的体现^[1]。语言的感知、理解和生成依赖多个功能区域的复杂神经网络^[2]。不同个体的神经网络功能连接模式以及大脑表征方式存在差异,深入探究语言的解码方法对于揭示语言的神经机制、人类认知与自我意识的复杂性至关重要^[3]。

脑机接口(Brain-computer interface, BCI)是用于在大脑与外部环境之间建立信息交互通路的系统^[4]。语义解码BCI通过采集、分析语言处理相关的神经信号,实现语义信息的识别或重建^[5],在辅助语言功能障碍患者恢复沟通能力、促进语言机制研究等方面具有巨大潜力,逐渐成为BCI领域中的研究热点。

功能性磁共振成像(Functional magnetic resonance imaging, fMRI)是一种非侵入式的神经信号采集技术,具有较高的空间分辨率,在语义解码研究中具有重要地位。相比于脑电图(Electroencephalography, EEG)、脑磁图(Magnetoencephalography, MEG)和功能性近红外光谱(Functional near-infrared spectroscopy, fNIRS),fMRI对深层脑区及皮层较大范围内的血氧水平依赖(Blood oxygen level dependent, BOLD)信号更加敏感^[6]。

目前基于fMRI的长序列中文语义解码方法的构建存在以下几个难点。首先,中文相比于英语,展现出独特的语法结构、丰富的词汇和鲜明的表意特点,体现出不同于其他语言的语言加工机制^[7]。其次,研究人员对英文语义的神经表征方式已经有了一定理解,但在中文语义层面上,尤其是复杂、连续的中文语义方面的理解还不够^[8]。文化背景和语境差异显著影响神经表征结果,反映了人脑在更广泛的认知和情感处理过程中展现的复杂且动态的神经适应性。最后,fMRI时间分辨率不高,神经活动引起BOLD信号持续时间超过8 s,一个时间点的神经信号通常是多个语义的叠加,基于fMRI的大规模、长序列语义解码仍然是该领域的重要挑战^[9]。近些年来,大语言模型(Large language models, LLMs)和多模态数据融合技术的快速发展,为基于fMRI的长序列中文语义解码方法的设计与实现提供了技术可能性^[10]。

本研究旨在通过优化信号处理流程、改进算法,构建一种基于fMRI的中文连续语义解码方法,并探索中文语义解码技术的神经机制,为中文BCI技术的发展提供重要理论基础。

1 相关工作

1.1 语义感知和理解的神经基础

语义的感知和理解过程涉及额、颞、顶、枕这4大脑叶的协同工作^[11]。每个脑叶下又包含众多功能特化的脑区,这些脑区通过高度互联的神经网络相互配合,共同实现语义感知和理解^[12-14]。本研究涉及的脑区以及各脑区的功能描述如表1所示。

表1 语义解码相关脑区
Table 1 Brain regions associated with semantic decoding

脑区		功能描述	脑区		功能描述
额叶 ^[15]			梭状回		整合多模态信息
布洛卡区		负责语法分析	横颞回		负责声音初步处理
左侧额下回		参与语义获取和工作记忆	内嗅皮层		支持语义动态更新与整合
背外侧前额叶		支持工作记忆、注意调控及高层次语义加工	顶叶 ^[16]		
上额叶		协助高级认知与语义检索	角回及上缘回		整合多模态信息,构建情境化整体概念
中额皮层		处理复杂语境	角回及上缘回		整合多模态信息,构建情境化整体概念
眶额皮层		关联语义情感	内侧顶叶		辅助注意分配和信息筛选
中央旁皮层		整合参与语义表征与身体感知的交互	楔前叶		结合外界感知与既往经验,支持语义动态重构
额极		负责抽象思维、概念推理,对语义理解起顶层指导作用	上顶叶		提取空间与情境特征
额极		负责抽象思维、概念推理,对语义理解起顶层指导作用	中央后回		处理身体感觉信息,构建与对象相关的触感和运动特征
前扣带皮层		修正语义错误以确保语境适配	中央后回		处理身体感觉信息,构建与对象相关的触感和运动特征
右侧额上回		调节情感并整合复杂语义,辅助中文双关语理解	后扣带皮层		整合外界信息与内在经验,支持自我反思与记忆检索
颞叶 ^[14]			扣带峡部		调控语义处理中的情绪与注意因素
左侧上颞沟		参与词句含义解析	枕叶 ^[17]		
颞上回后部		参与词句含义解析	视觉皮层		提取形状、颜色、边缘和运动等特征
左侧颞中回		储存具体概念表征	楔形回		处理低级视觉信息,为高级视觉语义加工提供基础输入
左侧颞下回		储存具体概念表征	外侧枕叶		编码形状、颜色等特征,是构建物体语义表征的基础
上颞沟边缘区		支持环境变化语义感知			

1.2 侵入性语义解码技术

侵入性语义解码技术通过深部脑电极、皮层内微电极阵列及皮层表面电极等方式获取高精度神经信号,在语义解码精度方面展现出显著优势^[18]。

Moses等^[19]在受试者左脑上颞回和腹侧感觉运动皮层植入皮层电图(Electrocorticography, ECoG),结合多种机器学习方法,实现了感知语言的实时解码,取得了0.251的准确率。Anumanchipalli等^[20]通过在腹侧感觉运动皮层、左脑上颞回和下额回植入ECoG,利用双向长短期记忆网络,实现了语音信号的解码,取得了0.21的解码准确率。

Jamali等^[21]在语言优势半球的背侧前额叶皮层使用微电极阵列(Microelectrode array, MEA)记录数百个单神经元活动,结合语义向量聚类分析,实现了0.31的语义类别解码精度。Willett等^[22]采用MEA采集腹侧前运动皮层神经信息,结合递归神经网络与LLMs,对肌萎缩侧索硬化症患者实现了较高精度实时语义解码,词错误率为0.238。Card等^[23]将MEA植入与语音生成高度相关的左侧中央前回,仅需较少临床数据即可实现高精度语义解码,便于临床部署和日常使用。Metzger等^[24]将MEA植入与语言生成和语义加工相关的脑区,帮助患者实现了每分钟29.4个字符的语言表达,字符错误率(Character error rate, CER)为0.061。

虽然侵入性语义解码技术在解码精度上具有优势,但其需要进行脑部手术,难以长期维持,涉及较高的实施风险和伦理问题,很难在非医疗领域推广^[25-26]。

1.3 非侵入性语义解码技术

非侵入性语义解码技术相比于侵入性语义解码技术,安全性更高、适用范围更广、使用和维护更为便捷,但其语义解码精度不如侵入性方法^[27]。传统的非侵入性神经信号采集方式,如EEG、MEG和fNIRS等,空间分辨率较低,语义解码精度有限^[28]。

Foster等^[29]使用岭回归模型和Skip-Gram词向量模型,通过EEG实现符号语义映射的解码,证明了新语言学习的早期阶段,新语言和母语共享一部分语义表征,预测准确率达到了69.15%。Défossez等^[30]使用CLIP模型,将语境表征与大脑活动表征对齐。结果表明,MEG的语义解码表现远超EEG,能提供更多大脑活动信息,取得了0.8的准确率。Lévy^[31]结合了卷积神经网络(Convolutional neural networks, CNN)和Transformer,从两种非侵入性神经成像(MEG和EEG)提取特征,构建语义解码器,实现了0.19的CER。

Zinszer等^[32]将fNIRS功能性反应与高级语义表征相联系,证明fNIRS可用于语义解码,实现了0.70的解码准确率。Cao等^[33]通过多通道fNIRS信号,整合多体素模式分析与全局向量词嵌入技术,构建了将神经活动映射到语义概念的语义解码器,实现了0.52的组间匹配精度。Cooney等^[34]利用EEG和fNIRS双模态神经信息构建语义解码BCI,但是其在显性语言和想象语言的分类任务上准确率不高,实现了0.463的解码准确度。

1.4 基于fMRI的语义解码技术

fMRI是一种空间分辨率较高的非侵入性脑成像技术,能够精确定位大脑不同区域的神经活动,已被广泛应用于神经科学研究^[35]。

Huth等^[13]使用fMRI绘制脑语义地图并探究了大脑各区域与特定语义类别之间的关联。Tang等^[36]探索了抽象概念在大脑中的表征方式,当处理具体概念时,大脑呈现出强烈的视觉和运动区域激活,而在处理抽象概念时,前额叶皮层、颞叶和内侧顶叶等高级认知区域的激活更为显著。

Ferrante等^[37]将fMRI数据映射到CNN的潜在语义空间,用于预测语义相关类别,最后通过扩散生成图像的语义与刺激高度一致,实现0.81的语义解码准确率。Liu等^[38]利用CLIP模型的视觉编码器和文本编码器,结合对比损失函数,将fMRI信号映射到CLIP的嵌入空间构建语义解码器,能够完成高语义保真度的图像生成任务,实现了0.948的识别准确率。

Tang等^[39]利用回归模型和LLMs,构建了一种基于fMRI的英文连续语义解码器,实现了大脑活动连续语言的重建,取得了0.82的BERTScore,并探究了大脑语言功能区域的冗余编码机制。该研究通过对LLMs语义刺激向量进行拼接,改善了fMRI低时间分辨率导致的语义刺激脑反应叠加问题。Zhang等^[40]利用CNN和LLMs实现了连续中文语义解码,取得了0.536的BERTScore。Huang等^[41]使用Transformer捕获fMRI数据中的全局特征,使用CNN捕获fMRI数据的局部特征,构建了脑对话接口

模型,实现了大脑语义信息的动态解码与通信,解码准确率为0.241。

2 研究方法

2.1 架构概述

本研究提出的中文长序列连续语义解码器(Chinese long-sequence continuous semantic decoder, CLCSD)的总体架构如图1所示。首先,通过皮层重建、图像配准和皮层平面化等方法,将4维脑响应数据降维成2维矩阵。然后,采用L2正则化回归(岭回归)建立刺激特征与脑响应之间的关系,通过自举法估计噪声协方差以增强泛化。接着,采用与编码模型类似的方法构建语速模型,估计脑响应特征与语义序列在时间上的映射关系。最后,利用语言模型的先验概率和编码模型的似然概率,通过束搜索生成最佳的语义序列。

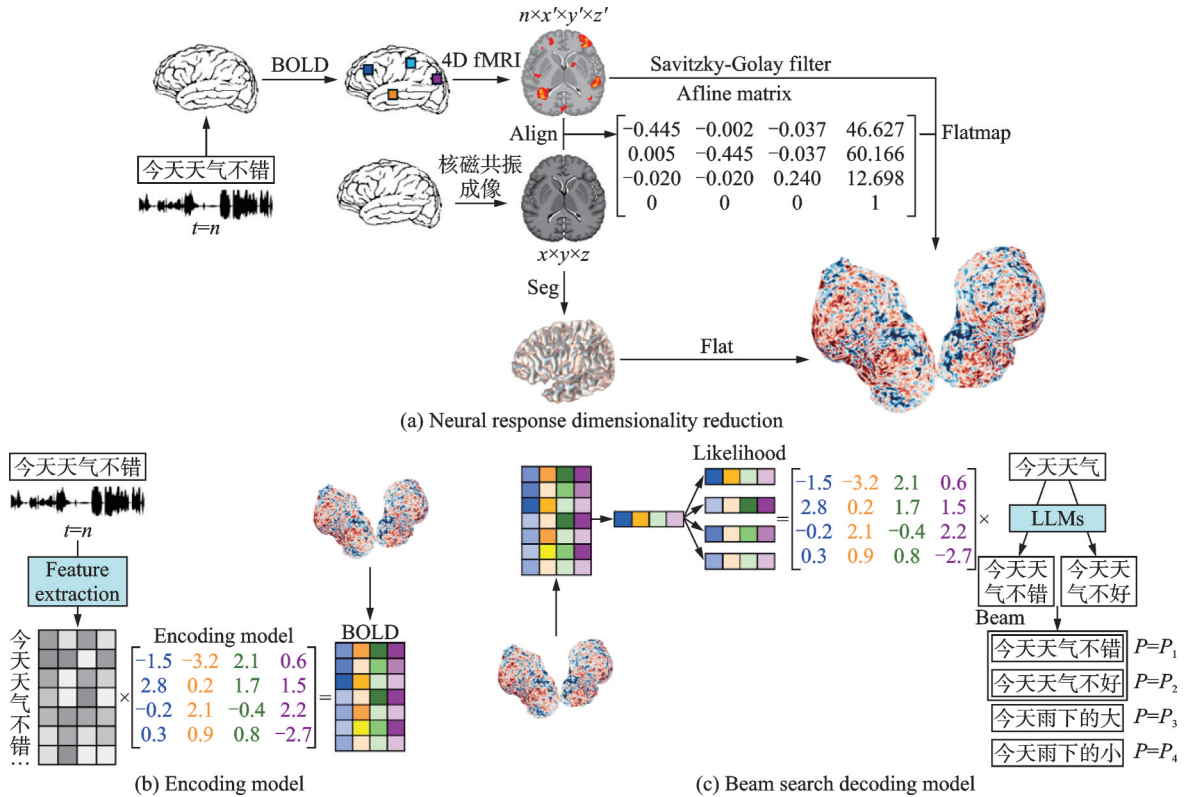


图1 CLCSD的整体架构

Fig.1 Overall architecture of CLCSD

2.2 数据集

本研究使用公开数据集SMN4Lang对模型进行了评估^[42]。SMN4Lang共收集了12名被试的结构性核磁共振成像(Magnetic resonance imaging, MRI)数据和fMRI数据,每位被试的fMRI数据时长为6 h。

实验开始,屏幕显示“等待扫描”的指令,随后呈现8 s的空白画面。然后,指令变为“此音频即将开始,请仔细聆听”,持续2.65 s后开始播放音频刺激材料。刺激材料包括60个长度为4~7 min的《人民日报》故事音频,涵盖教育、文化等多种主题,所有音频均由同一男性播音员朗读。时间戳经过手动校正,以确保音频与文本的精确对齐。

该数据集结构性MRI数据使用Siemens Prisma 3T扫描仪与64通道接收线圈获取,T1加权图像采用3D MPAGE序列,空间分辨率为0.8 mm体素,TR为2.4 s, TI为1 s, TE为2.22 ms, flip angle为8°, 视场(Field of view, FOV)为256 mm×256 mm。fMRI使用BOLD敏感的T2加权GE-EPI序列,空间分辨率为2 mm体素,TR为710 ms, TE为30 ms, flip angle为54°, FOV为212 mm×212 mm。

其中,4名被试的结构性MRI在预处理之后出现枕叶大范围缺损,因此最终选取了8名符合要求的被试用于后续分析。

2.3 神经响应降维

在语义感知和处理过程中,大脑皮层是神经信息的主要载体^[43]。因此,本研究仅对大脑皮层的神经响应数据进行降维^[34]。神经响应降维过程如图1(a)所示,将包含时间的4D fMRI数据降维成2D矩阵。2D矩阵每一列代表大脑皮层上的一个特定体素位置,每一行代表4D fMRI的一个时间点。

结构性MRI预处理包括3部分:(1)使用Freesurfer 7.4.1进行皮层分割与重建,从结构性MRI图像中提取大脑表面模型并获取皮质分区标签,皮层分割与重建的主要操作包含颅骨去除、偏置场校正、脑组织分割、皮质表面建模、皮层分区等,为后续皮层平面化提供准确的几何信息;(2)使用Blender去除胼胝体并标记皮层展平所需的接缝线,实现大脑皮层的切割和平面化;(3)根据语义感知的相关脑区和皮质分区标签,使用Freesurfer划定感兴趣的脑区(Regions of interest, ROIs),生成ROIs掩膜^[44]。

功能性MRI预处理包括2部分:(1)使用HCP的fMRIVolume和fMRISurface进行数据预处理(包括空间畸变修正、头动校正、偏置场校正、强度归一化和掩膜处理);(2)使用二阶Savitzky-Golay滤波器去除高频噪声^[45]。

最后,使用Pycortex进行结构性MRI和功能性MRI的自动配准和手动检测,确保结构性图像和功能性图像在标准空间中的一致性。采用“nearest”方案将预处理后的fMRI数据投影到大脑皮层展平图上,生成神经响应矩阵^[46]。

2.4 刺激语义特征提取策略

由于中文和英文在书写系统和词语分隔上的差异,两者在LLMs训练中采用不同的词元化策略。英文单词间存在明显的分隔标志(空格),词汇形态变化丰富,基于单词的词元化有助于捕捉语义和语法信息。中文缺乏明显的词间分隔标志,每个汉字具备独立意义,单字词元化有助于灵活组合形成完整词义^[35]。因此,英文的LLMs通常使用单词作为词元进行模型训练,而中文的LLMs通常使用字作为词元进行模型训练^[47]。

人脑在语言理解过程中综合处理词汇、句法及语境等多重信息,通过分布式语义网络激活并映射相应的概念表征,在语义层面完成信息处理与推理。在此过程中,大脑会将新输入的符号与已有的语义网络进行匹配、检索和比较,激活相关概念或语义域^[48]。这种基于语义信息处理方式与词编码的LLMs一致。

本研究深入探究了两种刺激语义特征提取策略对于中文语义解码的影响。第1种策略是根据“字”索引语义特征向量,使用经过CLUECorpus2020数据集中的新闻语料微调的GPT2-Chinese模型构建刺激矩阵^[49];第2种策略是根据“词”索引语义特征向量,使用GPT4模型批量将中文时间戳转化为英文时间戳,使用经过DeepMind Q&A数据集中的新闻语料微调的GPT2模型构建刺激矩阵^[50]。

时间戳批量转化的LLMs提示信息为:“将text中的中文换成最贴切的英文,可以拆分intervals以及intervals对应的时间,确保每个interval只能有一个英文单词,每句话的最后一个xmax要一致”。

对于每个时间戳中的词-时间对 (s_i, t_i) ,将单词序列 $(s_{i-5}, s_{i-4}, \dots, s_i)$ 输入至GPT语言模型,并从该模型的第9层中提取目标单词 s_i 的语义特征向量 M_i ,其中 M_i 为768维语义嵌入^[51]。获得 (M_i, t_i) 列表

后,使用 Three-lobe Lanczos 滤波器将这些嵌入向量在时间轴上重采样到 fMRI 数据采集对应的时间点^[52]。

2.5 岭回归编码模型

编码模型旨在建立刺激特征与大脑体素的 BOLD 信号之间的联系,如图 1(b)所示。人脑中绝大多数体素的 BOLD 信号对于瞬时刺激反应的变化符合血氧动力学响应函数(Hemodynamic response function, HRF)^[53]。该信号通常在刺激开始后 2 s 左右上升,并在约 4~6 s 后达到峰值,然后缓慢恢复到基线水平^[54-55]。HRF 反映了神经元放电与随后的血流、血容量和血氧水平变化之间的耦合关系,这一机制为利用 BOLD 信号分析并推断脑区在特定认知过程中的功能提供了理论依据。本研究使用的数据集 fMRI 的 TR 为 0.71 s,特定语义刺激后,6TR、7TR、8TR 和 9TR 时 BOLD 信号较强,即刺激呈现后 4.26 s、4.97 s、5.68 s 和 6.39 s 接近信号峰值。因此,通过这 4 个时间点的刺激向量 $S(t) \in \mathbf{R}^{768}$ 拼接成延迟特征矩阵 $X(t) \in \mathbf{R}^{768 \times 4}$,即

$$X(t) = [S_i(t-6), S_i(t-7), S_i(t-8), S_i(t-9)] \quad (1)$$

式中: t 为当前时间点,刺激向量 $S(t)$ 是从 GPT 模型第 9 层提取的 768 维语义嵌入特征, $i \in [1, 768]$ 是特征维度的索引。

将延迟特征矩阵中的每个特征通道进行 Z 得分标准化得到 $\tilde{X}(t)$,将该延迟特征矩阵与一组权重做点积,在功能上等价于将原始刺激向量与在 6、7、8 和 9 个时间点延迟处的线性时间核进行卷积,即

$$\hat{R}(t) = \tilde{X}^T(t)W \quad (2)$$

式中: $\hat{R}(t)$ 为特定体素 t 时间点的预测反应, W 为需要估计的权重向量,形状为 $\mathbf{R}^{768 \times 4}$ 。

为了防止编码模型过度拟合,并提高其在未见过的数据上的预测性能和稳定性,选择使用 L2 正则化回归(岭回归)^[52],L2 正则化的目标函数为

$$\min_W \left\{ \sum_{t=1}^T [\mathbf{R}(t) - \tilde{X}^T(t)W]^2 + \lambda \sum_{j=1}^{3072} W_j^2 \right\} \quad (3)$$

通过多次自举实验,求得最大的 λ^* 作为最佳正则化参数,即

$$\lambda^* = \arg \max_{\lambda} E_{\text{bootstrap}}[\mathbf{R}^2(\lambda)] \quad (4)$$

将 λ^* 代入岭回归的闭式解,得到最终的回归权重 W_{λ^*} 。

$$W_{\lambda^*} = (\tilde{X}^T(t)\tilde{X}(t) + \lambda^*I)^{-1} \tilde{X}^T(t)R \quad (5)$$

式中: R 为对应的脑响应向量, I 为同维度的单位矩阵。

编码模型估计了一个从语义特征 $X(t)$ 到预测的大脑反应 $\hat{R}(t)$ 的映射函数。fMRI 信号受到多种独立噪声源的影响(如仪器热噪声、生理波动、环境干扰等),叠加的总体噪声近似服从高斯分布。在高斯加性噪声假设下,给定 $X(t)$ 时大脑反应 $\hat{R}(t)$ 可似然建模为多元高斯分布^[56]

$$P(R(t)|X(t)) = N(\mu = \hat{R}(t), \Sigma_n) \quad (6)$$

式中 $\Sigma_n = (R(t) - \hat{R}(t))^T (R(t) - \hat{R}(t))$ 为噪声协方差矩阵。

以往研究通常直接使用训练数据集上的预测残差来估计 Σ_n ,这种方法会导致噪声被低估,因为模型在训练中已部分拟合了数据中的噪声。为克服这一问题,本研究采用自举法,将每个故事从训练集中移除,用剩余数据训练编码模型,然后利用被移除故事的数据计算预测残差,从而获得该故事对应的引导噪声协方差矩阵。之后,对所有被移除故事的噪声协方差进行平均,从而得到更准确的 Σ_n ^[53]。

2.6 语速模型

本研究提出一种语速模型用于预测被试听到刺激语义的具体时间,而刺激语义的具体时间可以由一个TR内出现的刺激语义的数量估计得出^[57]。

类似于编码模型,首先为每个时间点 t 构建脑响应延迟特征向量 $\mathbf{R}_d(t) \in \mathbf{R}^{4V}$

$$\mathbf{R}_d(t) = \text{concat}(\mathbf{R}(t+6), \mathbf{R}(t+7), \mathbf{R}(t+8), \mathbf{R}(t+9)) \quad (7)$$

式中: $\mathbf{R}(t) \in \mathbf{R}^V$ 为时间 t 的脑响应, V 为选定脑区(包含额叶、颞叶、顶叶和枕叶)的体素数。

然后,使用L2正则化线性回归将脑响应映射到预测的语速 $\mathbf{WR}(t)$

$$\widehat{\mathbf{WR}}(t) = \mathbf{R}_d^T(t) \mathbf{W} \quad (8)$$

在确定最佳正则化系数并估计最终权重 $\mathbf{W}$ 后,根据预测的语速将一个TR等分为若干时段,以获得单词出现时间点。

2.7 束搜索解码模型

解码模型可以视为编码模型的一个逆向过程,从脑响应数据 $\mathbf{R}_{\text{test}}$ 中解码出最可能的语义序列 $\mathbf{S}$,如图1(c)所示。

利用贝叶斯定理将后验分布分解为先验概率和似然概率

$$P(\mathbf{S}|\mathbf{R}_{\text{test}}) \propto P(\mathbf{S})P(\mathbf{R}_{\text{test}}|\mathbf{S}) \quad (9)$$

式中: $P(\mathbf{S})$ 为语言模型给定的先验概率,反映了在没有脑响应数据的情况下某个词序列本身出现的可能性,越符合语言习惯的序列,先验概率就越高; $P(\mathbf{R}_{\text{test}}|\mathbf{S})$ 为编码模型给定的似然概率,表示脑响应数据 $\mathbf{R}_{\text{test}}$ 由词序列 $\mathbf{S}$ 触发的概率。

由于在语义解码任务中,随着文本的加长,词序列数量呈指数级增长,导致计算量过大。因此,本研究使用束搜索算法,利用语言模型提供先验概率 $P(\mathbf{S})$ 和编码模型提供的似然概率 $P(\mathbf{R}_{\text{test}}|\mathbf{S})$ 将搜索空间限制在大小为 k 的候选序列 $\mathcal{B}$ 中,避免对所有可能的句子进行枚举^[45]。

在每个预测时刻,首先使用语言模型为候选序列 $\mathbf{S} \in \mathcal{B}$ 生成下一个词的概率分布为 $P(s_{t+1}|s_{t-i}, s_{t-i+1}, \dots, s_t)$,然后使用核采样筛选出概率较高的词,并过滤掉已在上下文中出现的词。将剩余可选词依次扩展到候选序列,得到新序列 $\mathbf{S}'$ 。对于每个新序列 $\mathbf{S}'$,利用编码模型计算其似然概率 $P(\mathbf{R}_{\text{test}}|\mathbf{S}')$ 。在对数空间下,可结合语言模型先验 $\ln P(\mathbf{S}')$ 求和得到总分数

$$\text{Score}(\mathbf{S}') = \ln P(\mathbf{R}_{\text{test}}|\mathbf{S}') + \alpha \ln P(\mathbf{S}') \quad (10)$$

式中 α 为权重参数, $\alpha = \frac{2}{3}$,用于调制似然概率和先验概率的平衡。

对每个新序列计算编码概率,选取得分最高的 k 个保留在新束 $\mathcal{B}_{\text{new}}$ 中($k=200$)。重复上述过程直至处理完所有预测词时刻。最后,在所有预测语义时间点迭代完成后,束中具有最高后验概率的候选序列即为最可能的解码结果,即

$$\mathbf{S}_{\text{best}} = \arg \max_{\mathbf{S} \in \mathcal{B}} [P(\mathbf{R}_{\text{test}}|\mathbf{S})P(\mathbf{S})] \quad (11)$$

2.8 脑功能连接分析

为了探究不同脑区在中文语义处理过程中的互动模式,以及中文语义加工与英文语义加工的神经网络差异。本研究使用两阶广义线性模型(Generalized linear model, GLM)研究中英文大脑的激活差异,并使用心理生理交互作用分析(Psychophysiological interaction, PPI)研究中文语义加工过程中的神经通路。本研究从中文语义感知数据集^[58]和英文语义感知数据集^[59]中分别抽样30个长序列fMRI文件^[60-61]。所有fMRI数据均使用fMRIPrep进行批量预处理。

在任务态与静息态激活比较中,个体水平分析阶段,使用基于HRF和自回归噪声模型建立一阶GLM模型。群体水平分析阶段,通过二阶GLM模型构建群体水平激活统计图,经体素水平校正(显著性水平 $p < 0.001$,标准正态分数 $Z > 3.09$)与最小簇体积50体素过滤后,获得显著激活区^[62]。

在中英文激活差异研究中,首先,对经百分比信号变化处理的中英文任务态数据,使用基于HRF和自回归噪声模型建立一阶GLM模型,生成个体 β 系数图;然后,将 β 图重采样,通过二阶GLM进行组间比较,计算全脑 Z 统计参数图;最后,采用体素水平Bonferroni校正($p < 0.05$)确定显著激活脑区^[63]。

在中文语义加工神经通路研究中,提取特定感兴趣区域(Region of Interest, ROI)内体素的时间序列,对生理项、心理项及原始交互项进行HRF卷积,利用一阶GLM拟合和二阶GLM群组比较,生成 Z 统计参数图,并使用错误发现率(False discovery rate, FDR) ($p < 0.001$)校正,获得与特定ROI功能显著耦合的脑区^[64]。

3 实验结果

3.1 模型评价指标

本研究运用BERTScore、METEOR、BLEU₁、余弦相似度(Cosine similarity, CS)和词错误率(Word error rate, WER)等自然语言处理(Natural language processing, NLP)评价指标,从多个角度评价CLCSD。

BERTScore采用BERT模型计算文本之间语义相似度,衡量了候选文本与参考文本的深层语义向量的匹配程度。其中Precision反映了两段文本的匹配质量,Recall反映了两段文本的匹配覆盖度, F_1 则是平衡了“匹配质量”与“匹配覆盖度”两方面^[65]。

$$\text{Precision} = \frac{1}{M} \sum_{i=1}^M \max_{j \in [1, 2, \dots, N]} \text{CS}(a_i, b_j) \quad (12)$$

$$\text{Recall} = \frac{1}{N} \sum_{j=1}^N \max_{i \in [1, 2, \dots, M]} \text{CS}(b_j, a_i) \quad (13)$$

$$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (14)$$

式中: $A = \{a_1, a_2, \dots, a_M\}$ 为候选文本的向量集合, $B = \{b_1, b_2, \dots, b_N\}$ 为参考文本的向量集合, M 和 N 分别为候选文本和参考文本的token数量,CS为余弦相似度的计算函数。

BLEU₁评估了解码文本和参考文本在1-gram水平上的重叠程度、关注词表层的匹配^[66]

$$\text{BLEU}_1 = \text{BP} \times \frac{\sum_{w \in Y_{1\text{-gram}}} \min(\text{count}_{\text{cand}}(w), \text{count}_{\text{ref}}(w))}{\sum_{w \in Y_{1\text{-gram}}} \text{count}_{\text{cand}}(w)} \quad (15)$$

式中:BP为长度惩罚项, $\text{count}_{\text{cand}}(w)$ 表示候选文本中1-gram的 w 出现的次数, $\text{count}_{\text{ref}}(w)$ 表示参考文本中1-gram的 w 出现的次数。

METEOR计算两段文本匹配时不仅考虑精确匹配,还考虑了词形变化和同义词匹配^[67]

$$\text{METEOR} = \left(1 - \gamma \left(\frac{\text{ch}}{m} \right)^\theta \right) \times \frac{P \times R_u}{\delta \cdot P + (1 - \delta) \cdot R_u} \quad (16)$$

式中: P 为候选文本和参考文本unigram的匹配精确率, R_u 为候选文本和参考文本unigram的召回率, ch 为连续匹配块的数目, m 为候选文本和参考文本之间匹配到的unigram对总数, δ, γ, θ 为超参数。

余弦相似度通过计算向量化后的候选文本和参考文本的向量夹角来衡量两段文本语义的相似度^[68]

$$\text{CS}(c, r) = \frac{c \cdot r}{\|c\| \|r\|} \quad (17)$$

式中: $c=\{c_1, c_2, \dots, c_n\}$ 为候选文本的向量集合, $r=\{r_1, r_2, \dots, r_n\}$ 为参考文本的向量集合。

WER常用于评估候选文本与参考文本在单词层面的差异,衡量了两段文本中替换、删除和插入错误的比例,反映了候选文本与参考文本在词级别上的匹配程度^[69]

$$\text{WER} = \frac{S + D + I}{N} \quad (18)$$

式中: S 表示候选文本中被错误替换的单词数, D 表示参考文本中缺失的单词数, I 表示候选文本中多余的单词数, N 表示参考文本中单词的总数。

3.2 特征提取策略对比实验

本文对比了2.4节提出的CLCSD的两种刺激语义特征提取策略与随机文本和其他现有模型的效果,如表2和图2所示,GPT2C为使用GPT2-Chinese模型构建刺激矩阵的CLCSD模型。

表2 多特征提取策略对比实验结果

Table 2 Comparative experiment results with feature extraction strategies

Model	BERTScore			BLEU ₁	METEOR	CS	WER
	Precision	Recall	F_1				
Random	0.504±0.027	0.501±0.019	0.502±0.018	0.200±0.014	0.104±0.021	0.615±0.035	0.937±0.005
3dC-IB ^[40]	—	—	0.536±0.003	0.217±0.008	0.125±0.006	—	—
GPT2C	0.634±0.007	0.631±0.010	0.632±0.008	0.300±0.014	0.164±0.021	0.735±0.055	0.925±0.027
GPT2	0.670±0.007	0.676±0.002	0.673±0.003	0.427±0.037	0.226±0.017	0.869±0.011	0.919±0.021

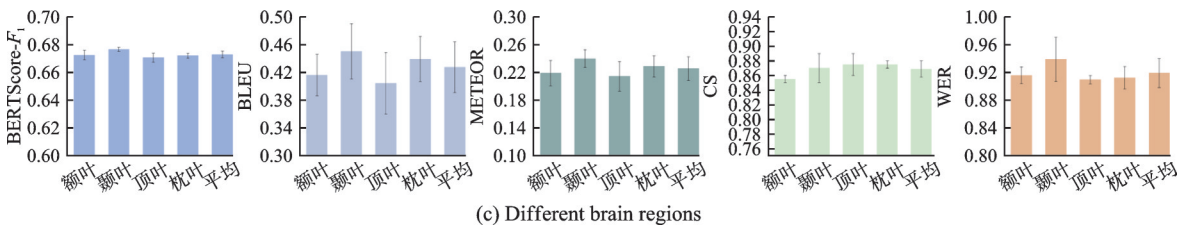
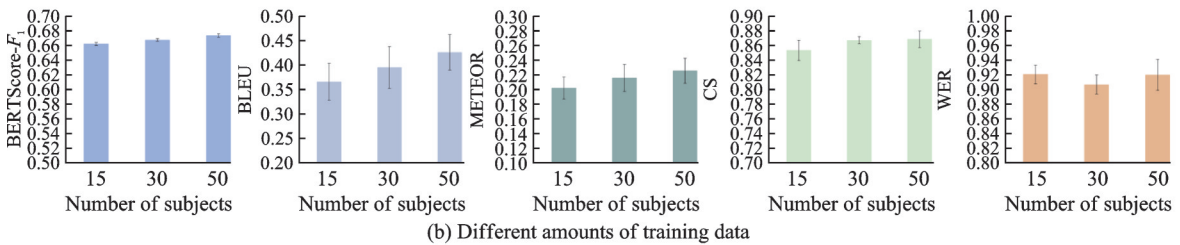
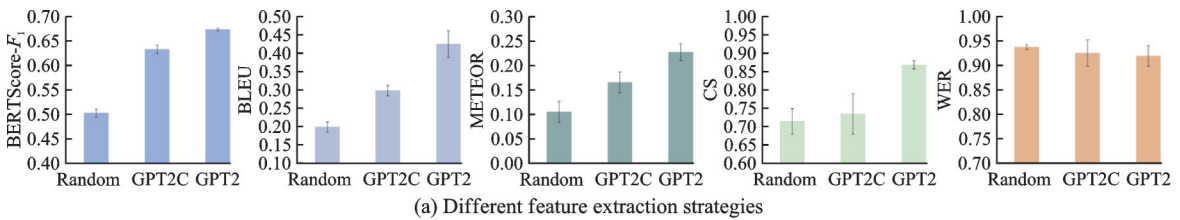


图2 对比实验结果

Fig.2 Comparison of experimental results

GPT2为使用GPT2模型构建刺激矩阵的CLCSD模型;Random为使用10对由GPT2-Chinese生成的不相关随机文本计算的平均实验结果。3dC-IB采用私有数据集,因此仅列出了部分实验结果。

对于BERTScore、BLEU₁和METEOR指标,随机模型、GPT2-Chinese和GPT2三者之间均存在显著差异($p<0.05$);在CS指标上,GPT2与GPT2-Chinese之间以及GPT2与随机模型之间均存在显著差异($p<0.05$);在WER指标上,GPT2-Chinese与随机模型之间以及GPT2与随机模型之间也均存在显著差异($p<0.05$)。

实验结果表明,基于两种特征提取策略构建的CLCSD模型均显著优于随机概率和3dC-IB,且基于GPT2(词编码)构建的CLCSD模型优于基于GPT2-Chinese(字编码)构建的CLCSD模型。这个结果可以由大脑的词边界效应来解释。

当人脑感知语言时,注意力更倾向于关注词的边界,词边界内的处理会更强烈地激活与语义相关的脑区,反应速度更快,且词边界编码会提升语速模型的性能^[70-73]。相较于单字表征,词语更完整地承载了上下文和语义信息,从而在一定程度上降低了潜在的歧义性^[74]。这种现象也在一定程度上解释了人脑在语言理解过程中主要依赖语义层面进行信息处理与推理的原因。

中文是一种高度依赖文字组合的语言,一个汉字本身可能具有多种含义,而多个汉字组合在一起形成的词语通常具有更明确的语义^[75]。GPT2-Chinese的单字切分策略在映射语义时可能会“割裂”词汇的整体含义,影响对神经信号的有效捕捉。通过将中文文本先转化为更有利于“词”级别分析的英文,使用GPT2进行词元化,这种语义嵌入使得编码模型能够减少词语内的歧义并理解词语之间的细微差别,从而提高语义解码效果。

3.3 训练样本数量对比实验

本节对比了不同训练样本数量对CLCSD性能的影响,如表3所示。随着训练样本数量的增加,各项NLP指标均有增加。对各项指标进行Shapiro-Wilk正态性检验,结果显示服从正态分布($p>0.05$)。采用单因素RM-ANOVA分析,15个训练量和50个训练量各项指标性能差异均达到统计显著性($p<0.05$),对于BERTScore- F_1 和CS,30个训练量和50个训练量性能差异也达到统计显著性($p<0.05$)。

表3 不同训练数量对比实验结果

Table 3 Comparative experiment results with different training amounts

训练数量	BERTScore			BLEU ₁	METEOR	CS	WER
	Precision	Recall	F_1				
15	0.632±0.005	0.634±0.001	0.633±0.003	0.389±0.027	0.199±0.019	0.801±0.004	0.920±0.013
30	0.660±0.005	0.659±0.001	0.660±0.004	0.401±0.030	0.214±0.019	0.853±0.005	0.919±0.021
50	0.670±0.007	0.676±0.002	0.673±0.003	0.427±0.037	0.226±0.017	0.869±0.011	0.906±0.013

此外,随着训练集规模的增加,生成语句的连贯性显著提升,本研究选取了额叶解码结果的一个实例进行对比分析。原始文本为:最近,一架已经在滑行道上准备起飞的飞机突然紧急返回登机口,这一事件引发了广泛的关注。原因是有一对老年夫妇突然接到亲属去世的消息,他们希望更改行程。机组人员因此决定将飞机滑回,导致航班延误了几分钟。

如表4所示,当训练样本规模较小时,生成的文本在句子之间仍存在不连贯的情况。当训练样本数量为15时,第1句与第2句、第2句与第3句之间在语义衔接与逻辑连续性上出现了较为明显的脱节;当训练样本数量为30时,第1句与第2句之间也存在一定程度的断裂。当训练样本规模进一步增大至50时,生成结果在话题衔接与语义延展方面表现出更高的一致性和连贯度。

随着训练样本数量的增加,模型能够学习到更全面的语义特征与上下文依赖关系,从而在生成阶段更好地保证段落之间主题一致与逻辑连贯。相反,训练样本过少时,模型对部分语言模式和衔接方

表 4 不同训练样本数量解码效果对比

Table 4 Comparison of decoding performance with different numbers of training sets

训练样本 数量	解码效果
15	你太忙无法登上飞机去乘坐私人飞机飞往最近的机场,那就是另一回事了。有一位刚因癌症去世的老妇人,她的邻居刚得知这起意外,想把车开回机场,用几个小时完成这趟旅程。这趟飞行时间不过几分钟,但还没结束,旅程就已经匆匆而过。
31	我不认识你,但曾见过你去机场的路上。当时飞机停在跑道上以避免撞上一辆卡车,结果发生了一起小事故,一对年长夫妇因此受伤。他们有一位家庭成员受了伤,后来决定换一辆新车,他们想开车回到机场。这次航班并不长,飞机刚刚降落,但时间已经过去了。
50	她在路上,遇到一位老年妇女,她刚得知一位家庭好友去世的消息,并希望尽快处理好这件事。因此,她决定开车回家,然后再赶回机场。那次航班非常短暂,飞机只降落了大约几分钟,但时间已经过去了。

式的学习不足,导致在合成文本中出现突兀或“跳脱”的句子。当训练样本数量达到 30 时,模型已能学习到基本的语义特征和上下文依赖关系,生成效果趋于稳定;而进一步增加训练量至 50 则主要优化了句间连贯性和语义流畅度。这表明,30 个样本已能使模型具备较好的语言生成能力,而更大规模的训练则进一步提升了文本的衔接质量。

3.4 脑区对比实验

本节对比了不同脑区对 CLCSD 性能的影响,结果如表 5 和表 6 所示。本节所使用的实例与 3.3 节相同。

表 5 不同脑叶对比实验结果

Table 5 Comparative experiment results on different brain regions

脑区	BERTScore			BLEU ₁	METEOR	CS	WER
	Precision	Recall	F ₁				
额叶	0.671±0.007	0.674±0.001	0.672±0.003	0.416±0.030	0.219±0.019	0.855±0.005	0.916±0.012
颞叶	0.673±0.006	0.680±0.003	0.676±0.001	0.450±0.040	0.240±0.013	0.870±0.020	0.939±0.032
顶叶	0.669±0.008	0.672±0.001	0.670±0.003	0.404±0.044	0.215±0.022	0.875±0.015	0.909±0.006
枕叶	0.668±0.007	0.676±0.004	0.672±0.002	0.439±0.033	0.229±0.015	0.875±0.005	0.912±0.016
平均	0.670±0.007	0.676±0.002	0.673±0.003	0.427±0.037	0.226±0.017	0.869±0.011	0.919±0.021

表 6 不同脑区解码效果对比

Table 6 Comparison of decoding performance of different brain regions

脑区	解码效果
额叶	她在路上,遇到一位老年妇女,她刚得知一位家庭好友去世的消息,并希望尽快处理好这件事。因此,她决定开车回家,然后再赶回机场。那次航班非常短暂,飞机只降落了大约几分钟,但时间已经过去了。
颞叶	你会告诉她,你希望她开车到机场,再乘飞机返回家中。飞机降落后滑行到一个等候区,她抵达时发现一位身患癌症、濒临去世的老人,于是决定迅速恢复状态并再次赶回机场乘机回家。这趟航班非常短暂,飞机刚刚落地,几分钟就过去了。
顶叶	我曾搭乘一架航班赶往机场,另一架航班正在登机,准备飞回我的家乡。机上的乘客是一对刚刚因心脏病去世的老年夫妇,他们决定开车回家,然后再赶去机场。这次飞行很短,飞机刚刚降落,然而头等舱的登机已经开始,所以我别无选择,只能跟随自己的直觉。
枕叶	我曾乘坐一架飞机,前往一架私人飞机转机的机场,准备飞回我父母那里。航班因一起不幸的事故而延误,一对刚刚失去孩子的年长夫妇决定先回家一趟,然后再驱车返回机场。这趟飞行很短,飞机刚刚降落,时间却悄然流逝了几分钟。

对各项指标进行 Shapiro-Wilk 正态性检验,结果显示服从正态分布($p>0.05$)。采用单因素 RM-ANOVA 分析,对于 BERTScore-Precision、BLEU₁、METEOR、CS 和 WER,4 大脑叶解码器性能差异达到统计显著性($p<0.05$)。

脑区对比实验结果表明,颞叶的解码结果取得了相对最高的 NLP 指标结果,额叶与枕叶的解码效果接近,略低于颞叶,顶叶的解码效果最差。实验结果的差异可由不同脑区在语义感知过程中的功能定位来解释^[76-77]。

不同脑区在语义处理过程中通过白质纤维束及功能网络相互协作、动态交互实现了语义的高效理解^[71]。颞叶在语义解码任务中表现最好,取得了相对最高的 BERTScore、BLEU₁ 和 METEOR,反映了其在深层语义和潜层语义的解码均具备优越性^[78-81]。额叶在语义解码任务中表现相对较差,主要原因在于额叶主要负责语义检索和认知控制,更多关注信息的选择和调控,而非细致语义内容的构建^[82]。顶叶和枕叶取得了相对较高的 CS 和 WER,反映了其在字面上的精确匹配和顺序一致性,这可能是由于顶叶主要负责信息的整合,而枕叶主要负责信息的感知^[83-86]。

3.5 脑功能连接分析结果

3.5.1 GLM 分析结果

中文语义感知任务态与静息态激活比较结果和中英文语义感知脑区激活差异结果如图 3、4 和表 7、8 所示。

在中文语义感知任务中,左额极、左颞上回后部、右颞上回后部及右额极表现出显著激活。左额极和右内侧额极的协同激活,可能与语义感知过程中自上而下的注意调控与跨模态语义整合的层级化加工机制有关^[87]。双侧颞上回后部的激活符合语音-语义初级映射机制。双侧颞上回后部可能负责声学特征解析及语义基元提取,而双侧颞极的参与反映了抽象语义的表征与社会情境信息的耦合加工特征^[88]。语义感知任务期间,默认模式网络及非任务相关脑区的同步抑制,体现了大脑认知资源向语言核心网络的定向优化和配置的过程^[89]。

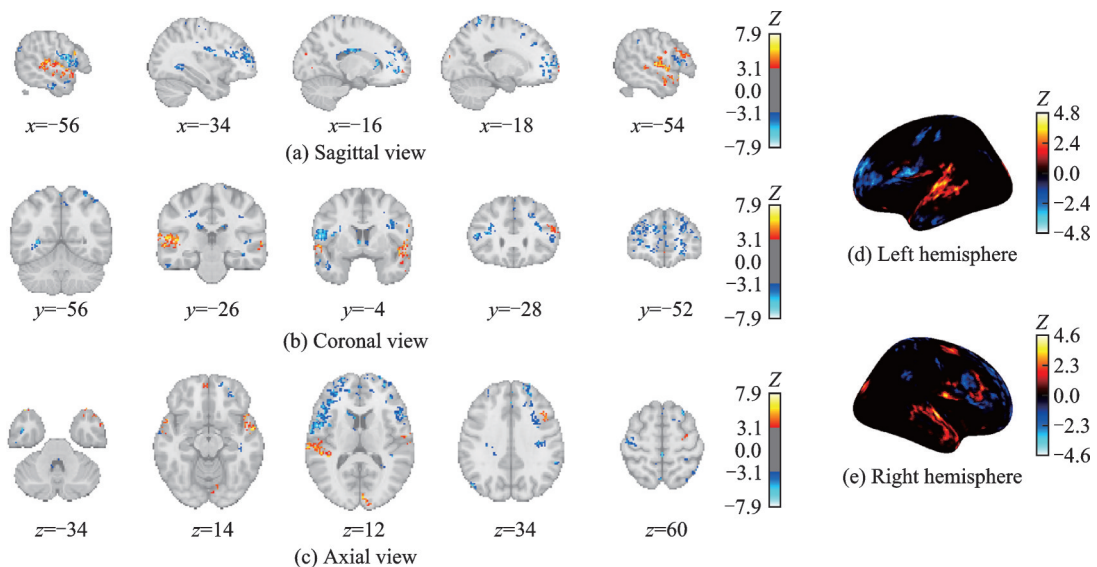


图3 中文语义感知任务态与静息态激活比较

Fig.3 Activation comparison between Chinese semantic perception task state and resting state

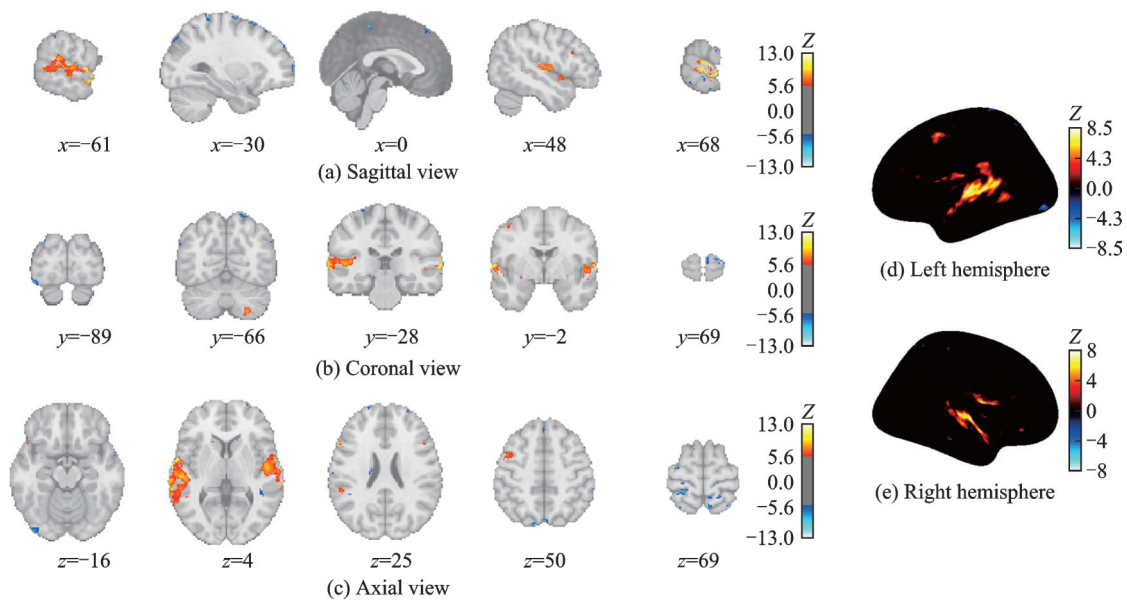


图4 中英文语义感知脑区激活差异

Fig.4 Comparative analysis of brain activation patterns in Chinese and English semantic decoding

表7 中文语义感知任务下的激活簇

Table 7 Activation clusters during Chinese semantic decoding tasks

脑区	MNI坐标			峰值Z	簇大小
	<i>x</i>	<i>y</i>	<i>z</i>		
正向激活簇					
左额极	−26	62	−8	6.15	2 225
左颞上回后部	−60	−10	8	7.26	990
右颞上回后部	56	4	−2	7.89	880
右额极	26	68	−2	5.26	520
右额中回	52	18	24	6.20	180
左颞极	−56	4	−2	6.41	112
负向激活簇					
右额极	36	10	34	−3.09	730
右顶上小叶	30	−40	78	−3.09	346
楔前叶皮层	0	−38	52	−3.09	280
左中央后回	−34	−48	76	−3.09	269
旁扣带回	0	44	14	−3.10	249
左中央前回	−52	−14	50	−3.09	236

相较于英文语义感知任务,中文任务在左颞中回诱发更强的神经激活,提示汉语语义加工可能更依赖该区域对语音-语义信息的整合解析^[90]。中央沟的显著激活表明中文处理需强化语音工作记忆以应对语音歧义消解需求。同时,工作记忆网络(左额上回-右顶上叶)与注意网络(右楔前叶)的协同激活,反映中文理解需要更高认知资源以维持上下文整合及干扰抑制^[91]。左枕中回的跨模态激活进一步支持汉语形-音关联特性可能促进语音信息向视觉表征的自动映射^[92]。

表8 中英文语义感知脑区激活簇的差异

Table 8 Differences in activation clusters of semantic perception brain regions between Chinese and English

脑区	MNI坐标			峰值 Z	簇大小
	x	y	z		
正向激活簇					
左颞中回	−57.96	−24.45	7.13	12.75	1 708
右中央操作皮层	58.12	−9.72	2.98	11.48	1 083
左额上回	−42.63	3.04	50.18	7.26	643
左枕中回	−42.74	−84.85	−12.66	8.39	588
右顶上叶	6.69	−61.89	65.89	7.88	454
右楔前叶	−26.04	−39.95	19.61	9.19	390
负向激活簇					
左中央前回	36	10	34	−3.09	730
右枕下回	30	−40	78	−3.09	346
右颞下回	0	−38	52	−3.09	280
左枕下回	−34	−48	76	−3.09	269
旁扣带回	0	44	14	−3.10	249
左中央前回	−52	−14	50	−3.09	236

3.5.2 PPI网络分析结果

本研究选取中文语义感知任务下的显著激活脑区(包含左额极、左颞上回后部、右颞上回后部和右额极)作为种子ROI,采用PPI分析方法探讨中文语义加工的神经通路机制,具体结果见图5和表9。图中A为左颞上回后部,B为右颞上回后部,C为左额极,D为右额极,横断位投射距离相近点归为同一簇。

PPI结果表明,左右颞上回后部与中颞回、颞极、前扣带皮层与眶额内侧皮层交界处均存在显著功能耦合。双侧颞上回后部参与了声学特征感知到语义概念的转换,在长序列自然语言理解任务中,通过动员更多的认知控制资源维持注意集中、解决语义冲突^[93]。双侧颞上回通过全脑范围的连接模式参与整体神经调控,这可能是颞叶在中文语义解码任务中解码效果表现最好的重要原因之一^[94]。左侧颞上回后部与中央后回网络、左侧额极具有显著功能耦合,说明其与语义加工中与感觉、运动表征的功能整合相关,有助于多词串联语义推理及句际整合的高阶控制^[95]。此外,右侧颞上回后部存在大量且分散的功能耦合簇,提示其在连续中文语义解码时承担更为分散的关联功能,促进语义感知任务中多模态感觉信息的整合^[96]。

左侧额极通过背外侧前额叶-后部颞叶通路,与颞下回后部、右侧中央后回、右侧额极和眶额皮层等语言与感觉运动核心区强耦合,表明其可能在目标导向的语义选择与社会认知整合中发挥关键调控^[97]。右侧额极主要与左侧额极、右侧中央后回、右侧顶缘回后部、枕极及眶额皮层等区域存在功能连接耦合,可能在大脑非语言性社会认知加工、情绪效价评估、复杂情境监控任务中承担重要角色^[98]。总

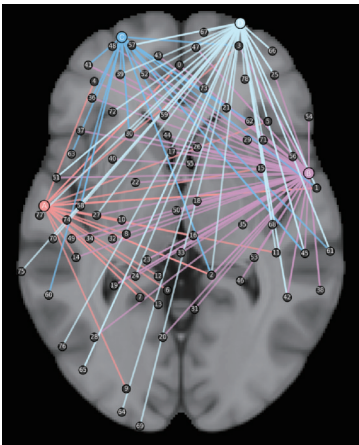


图5 PPI网络分析结果
Fig.5 PPI network analysis results

表9 PPI功能连接索引表

Table 9 PPI functional connectivity index table

编号	簇	编号	簇	编号	簇	编号	簇
0	额内皮层,旁扣带回	20	楔前叶皮层	40	额中回	60	颞中回颞枕部
1	颞中回,中央后回,颞上回	21	额极,额上回	41	额极	61	下颞回后部,颞中回后部
2	扣带回后部	22	额上回	42	角回	62	额中回
3	额极	23	中央后回	43	旁扣带回	63	下额回操作部
4	额极	24	中央后回	44	额上回	64	枕极
5	颞极	25	额极	45	下颞回后部,顶缘回后部	65	枕外侧皮层上部
6	中央后回	26	旁扣带回,额上回	46	顶上小叶	66	额极
7	顶上小叶	27	中央前回	47	额极	67	额极
8	中央前回	28	枕外侧皮层上部	48	额极	68	岛叶皮层,中央后回
9	枕外侧皮层上部	29	额中回	49	中央后回	69	枕极
10	中央前回	30	额眶皮层,额上回	50	扣带回前部	70	顶盖皮层
11	顶盖皮层,中央后回	31	楔前叶皮层	51	颞中回前部,中央前回	71	岛叶皮层
12	中央后回	32	中央前回	52	额极,旁扣带回	72	额眶皮层
13	楔前叶皮层	33	扣带回后部	53	中央后回	73	扣带回前部,旁扣带回
14	中央后回	34	中央后回	54	下额回三角部	74	中央盖区皮层
15	中央盖区皮层,颞极	35	中央前回	55	扣带回前部	75	颞中回后部
16	扣带回后部,中央前回	36	额眶皮层,额中回	56	下额回操作部	76	枕外侧皮层上部
17	扣带回前部,旁扣带回	37	下额回三角部	57	额极	77	中央后回
18	辅助运动皮层	38	下颞回颞枕部	58	中央盖区皮层,中央前回	78	额极
19	顶上小叶	39	额极	59	额眶皮层,额上回		

之,双侧颞上回后部与双侧额极在语义表征与认知控制任务中相互配合,为中文语义解码提供了丰富又稳定的神经基础。

4 结束语

本研究的创新点在于:(1)提出了一种基于fMRI的长序列中文连续语义解码方法,通过融合LLMs和机器学习技术,实现了高效的中文连续语义解码,为中文语义解码BCI技术的发展提供重要理论基础;(2)探究了大脑的词边界机制,对比了基于“词”和基于“字”两种刺激语义特征提取策略对于中文语义解码技术的影响;(3)评估了4大脑叶对于语义解码性能的贡献度以及神经活动模式差异,为脑语言功能研究提供更深入的见解。

本研究的不足:(1)由于满足要求的数据存在稀缺性,本研究仅在8个被试上进行实验,模型的泛化能力有待进一步验证,但是每个被试用于训练的功能性数据超过5 h,具有较好的代表性和数据规模,能够一定程度上缓解这个问题;(2)由于个体间大脑功能连接和语义表征存在差异,单纯依赖结构MRI进行ROIs划定存在局限,尽管结构MRI标准化和重复性较高,但其主要反映解剖特征,难以捕捉个体功能和语义加工的特异性,结合语义感知任务态与静息态fMRI的差异划定更多ROIs,有望提升解码器性能和模型泛化能力;(3)fMRI中的噪声并不完全符合高斯分布,采用高斯去噪可能存在噪声残留,后续可以引入更全面的噪声处理方法^[99];(4)中文语言具有复杂性,中文语法独特、词汇丰富、表意性强,语义加工机制复杂,在实际应用场景中,面对复杂多样的中文表达和语义情境,模型可能难以全面准确地捕捉和理解语义信息,后续可以引入中文能力更强的LLMs。

参考文献:

- [1] FEDORENKO E, PIANTADOSI S T, GIBSON E A F. Language is primarily a tool for communication rather than thought [J]. *Nature*, 2024, 630(8017): 575-586.
- [2] SUEOKA Y, PAUNOV A, TANNER A, et al. The language network reliably “tracks” naturalistic meaningful nonverbal stimuli[J]. *Neurobiology of Language*, 2024, 5(2): 385-408.
- [3] SCHRIMPF M, BLANK I, TUCKUTE G, et al. The neural architecture of language: Integrative reverse-engineering converges on a model for predictive processing[J]. *Proceedings of the National Academy of Sciences*, 2021, 118(45): e2105646118.
- [4] HE B, YUAN H, MENG J, et al. Brain-computer interfaces[EB/OL]. (2020-09-22). <http://doi.org/10.1007/978-3-030-43395-64>.
- [5] NICOLAS-ALONSO L F, GOMEZ-GIL J. Brain computer interfaces, a review[J]. *Sensors*, 2012, 12(2): 1211-1279.
- [6] CHIARELLI A M, ZAPPASODI F, DI POMPEO F, et al. Simultaneous functional near-infrared spectroscopy and electroencephalography for monitoring of human brain activity and oxygenation: A review[J]. *Neurophotonics*, 2017, 4(4): 041411.
- [7] HANSEN C. Chinese ideographs and western ideas[J]. *The Journal of Asian Studies*, 1993, 52(2): 373-399.
- [8] FENG S, QI R, YANG J, et al. Neural correlates for nouns and verbs in phrases during syntactic and semantic processing: An fMRI study[J]. *Journal of Neurolinguistics*, 2020, 53: 100860.
- [9] DALE A M, LIU A K, FISCHL B R, et al. Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity[J]. *Neuron*, 2000, 26(1): 55-67.
- [10] ZHAO Y, CHEN Y, CHENG K, et al. Artificial intelligence based multimodal language decoding from brain activity: A review[J]. *Brain Research Bulletin*, 2023, 201: 110713.
- [11] 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[J]. *Cerebral Cortex*, 2009, 19(12): 2767-2796.
- [12] BINDER J R, DESAI R H. The neurobiology of semantic memory[J]. *Trends in Cognitive Sciences*, 2011, 15(11): 527-536.
- [13] HUTH A G, DE HEER W A, GRIFFITHS T L, et al. Natural speech reveals the semantic maps that tile human cerebral cortex[J]. *Nature*, 2016, 532(7600): 453-458.
- [14] RALPH M A, JEFFERIES E, PATTERSON K, et al. The neural and computational bases of semantic cognition[J]. *Nature Reviews Neuroscience*, 2017, 18(1): 42-55.
- [15] THOMPSON-SCHILL S L, D’ESPOSITO M, AGUIRRE G K, et al. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation[J]. *Proceedings of the National Academy of Sciences*, 1997, 94(26): 14792-14797.
- [16] SEGHER M L. The angular gyrus: Multiple functions and multiple subdivisions[J]. *The Neuroscientist*, 2013, 19(1): 43-61.
- [17] CICHY R M, PANTAZIS D, OLIVA A. Resolving human object recognition in space and time[J]. *Nature Neuroscience*, 2014, 17(3): 455-462.
- [18] GHOSH S, MÁTHÉ D, HARISHITA P B, et al. Review of multimodal data acquisition approaches for brain-computer interfaces[J]. *BioMed*, 2024, 4(4): 548-587.
- [19] MOSES D A, LEONARD M K, MAKIN J G, et al. Real-time decoding of question-and-answer speech dialogue using human cortical activity[J]. *Nature Communications*, 2019, 10(1): 3096.
- [20] ANUMANCHIPALLI G K, CHARTIER J, CHANG E F. Speech synthesis from neural decoding of spoken sentences[J]. *Nature*, 2019, 568(7753): 493-498.
- [21] JAMALI M, GRANNAN B, CAI J, et al. Semantic encoding during language comprehension at single-cell resolution[J]. *Nature*, 2024, 631(8021): 610-616.
- [22] WILLETT F R, KUNZ E M, FAN C, et al. A high-performance speech neuroprosthesis[J]. *Nature*, 2023, 620(7976): 1031-1036.
- [23] CARD N S, WAIRAGKAR M, IACOBACCI C, et al. An accurate and rapidly calibrating speech neuroprosthesis[J]. *New England Journal of Medicine*, 2024, 391(7): 609-618.

- [24] METZGER S L, LIU J R, MOSES D A, et al. Generalizable spelling using a speech neuroprosthesis in an individual with severe limb and vocal paralysis[J]. *Nature Communications*, 2022, 13(1): 6510.
- [25] FARAH M J. Neuroethics: The ethical, legal, and societal impact of neuroscience[J]. *Annual Review of Psychology*, 2012, 63(1): 571-591.
- [26] LEUTHARDT E C, MORAN D W, MULLEN T R. Defining surgical terminology and risk for brain computer interface technologies[J]. *Frontiers in Neuroscience*, 2021, 15: 599549.
- [27] FIRUZI R, AHMADYANI H, ABDI M F, et al. Decoding neural signals with computational models: A systematic review of invasive BMI[EB/OL]. (2022-11-07). <https://arxiv.org/abs/2211.03324v2>.
- [28] LI R, YANG D, FANG F, et al. Concurrent fNIRS and EEG for brain function investigation: A systematic, methodology-focused review[J]. *Sensors*, 2022, 22(15): 5865.
- [29] FOSTER C, WILLIAMS C C, KRIGOLSON O E, et al. Using EEG to decode semantics during an artificial language learning task[J]. *Brain and Behavior*, 2021, 11(8): e2234.
- [30] DÉFOSSEZ A, CAUCHETEUX C, RAPIN J, et al. Decoding speech perception from non-invasive brain recordings[J]. *Nature Machine Intelligence*, 2023, 5(10): 1097-1107.
- [31] LÉVY J, ZHANG M, PINET S, et al. Brain-to-text decoding: A non-invasive approach via typing[EB/OL]. (2025-02-18). <https://arxiv.org/abs/2502.17480v1>.
- [32] ZINSZER B D, BAYET L, EMBERSON L L, et al. Decoding semantic representations from functional near-infrared spectroscopy signals[J]. *Neurophotonics*, 2018, 5(1): 011003.
- [33] CAO L, HUANG D, ZHANG Y, et al. Brain decoding using fnirs[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, 35(14): 12602-12611.
- [34] COONEY C, FOLLI R, COYLE D. A bimodal deep learning architecture for EEG-fNIRS decoding of overt and imagined speech[J]. *IEEE Transactions on Bio-Medical Engineering*, 2022, 69(6): 1983-1994.
- [35] FINN E S, POLDRACK R A, SHINE J M. Functional neuroimaging as a catalyst for integrated neuroscience[J]. *Nature*, 2023, 623(7986): 263-273.
- [36] TANG J, LEBEL A, HUTH A G. Cortical representations of concrete and abstract concepts in language combine visual and linguistic representations[J]. *Cold Spring Harbor Laboratory*, 2021. DOI: 10.1101/2021.05.19.444701.
- [37] FERRANTE M, BOCCATO T, TOSCHI N. Semantic brain decoding: From fMRI to conceptually similar image reconstruction of visual stimuli[EB/OL]. (2022-12-13). <https://arxiv.org/abs/2212.06726v2>.
- [38] LIU Y, MA Y, ZHOU W, et al. BrainCLIP: Bridging brain and visual-linguistic representation via CLIP for generic natural visual stimulus decoding[EB/OL]. (2023-02-25). <https://arxiv.org/abs/2302.12971v3>.
- [39] TANG J, LEBEL A, JAIN S, et al. Semantic reconstruction of continuous language from non-invasive brain recordings[J]. *Nature Neuroscience*, 2023, 26(5): 858-866.
- [40] ZHANG C, ZHENG X, YIN R, et al. Decoding continuous character-based language from non-invasive brain recordings[EB/OL]. (2024-03-17). <https://arxiv.org/abs/2403.11183>.
- [41] HUANG H, ZHAO L, WU Z, et al. Brain dialogue interface (BDI): A user-friendly fMRI model for interactive brain decoding[EB/OL]. (2024-06-17). <https://arxiv.org/abs/2407.09509v1>.
- [42] WANG S, ZHANG X, ZHANG J, et al. A synchronized multimodal neuroimaging dataset for studying brain language processing[J]. *Scientific Data*, 2022, 9: 590.
- [43] FAIRHALL S L, CARAMAZZA A. Brain regions that represent amodal conceptual knowledge[J]. *Journal of Neuroscience*, 2013, 33(25): 10552-10558.
- [44] FISCHL B. FreeSurfer[J]. *NeuroImage*, 2012, 62(2): 774-781.
- [45] BARRY R L, GORE J C. Enhanced phase regression with Savitzky-Golay filtering for high-resolution BOLD fMRI[J]. *Human Brain Mapping*, 2014, 35(8): 3832-3840.
- [46] GAO J S, HUTH A G, LESCROART M D, et al. Pycortex: An interactive surface visualizer for fMRI[J]. *Frontiers in Neuroinformatics*, 2015, 9: 00023.
- [47] ZHAO W X, ZHOU K, LI J, et al. A survey of large language models[EB/OL]. (2023-03-31). <https://arxiv.org/abs/>

2303.18223.

- [48] GIBAUT W, PEREIRA L, GRASSIOTTO F, et al. Neurosymbolic AI and its taxonomy: A survey[EB/OL]. (2023-05-12). <https://arxiv.org/abs/2305.08876v2>.
- [49] XU L, ZHANG X, DONG Q. CLUECorpus2020: A large-scale Chinese corpus for pre-training language model[EB/OL]. (2020-03-03). <https://arxiv.org/abs/2003.01355v2>.
- [50] HERMANN K M, KOČISKÝ T, GREFFENSTETTE E, et al. Teaching machines to read and comprehend[EB/OL]. (2015-06-10). <https://arxiv.org/abs/1506.03340v3>.
- [51] WOLFE R, CALISKAN A. VAST: The valence-assessing semantics test for contextualizing language models[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022, 36(10): 11477-11485.
- [52] POPHAM S F, HUTH A G, BILENKO N Y, et al. Visual and linguistic semantic representations are aligned at the border of human visual cortex[J]. *Nature Neuroscience*, 2021, 24(11): 1628-1636.
- [53] CHEN G, TAYLOR P A, REYNOLDS R C, et al. BOLD response is more than just magnitude: Improving detection sensitivity through capturing hemodynamic profiles[J]. *NeuroImage*, 2023, 277: 120224.
- [54] BOYNTON G M, ENGEL S A, GLOVER G H, et al. Linear systems analysis of functional magnetic resonance imaging in human V1[J]. *Journal of Neuroscience*, 1996, 16(13): 4207-4221.
- [55] GLOVER G H. Deconvolution of impulse response in event-related BOLD fMRI[J]. *NeuroImage*, 1999, 9(4): 416-429.
- [56] NOZARI E, BERTOLERO M A, STISO J, et al. Macroscopic resting-state brain dynamics are best described by linear models[J]. *Nature Biomedical Engineering*, 2024, 8(1): 68-84.
- [57] ANDERSON A J, BINDER J R, FERNANDINO L, et al. Predicting neural activity patterns associated with sentences using a neurobiologically motivated model of semantic representation[J]. *Cerebral Cortex*, 2017, 27(9): 4379-4395.
- [58] LEBEL A, WAGNER L, JAIN S, et al. A natural language fMRI dataset for voxelwise encoding models[J]. *Scientific Data*, 2023, 10(1): 555.
- [59] FEDORENKO E, IVANOVA A A, REGEV T I. The language network as a natural kind within the broader landscape of the human brain[J]. *Nature Reviews Neuroscience*, 2024, 25(5): 289-312.
- [60] DESMOND J E, GLOVER G H. Estimating sample size in functional MRI (fMRI) neuroimaging studies: Statistical power analyses[J]. *Journal of Neuroscience Methods*, 2002, 118(2): 115-128.
- [61] MUMFORD J A, NICHOLS T E. Power calculation for group fMRI studies accounting for arbitrary design and temporal autocorrelation[J]. *NeuroImage*, 2008, 39(1): 261-268.
- [62] FRISTON K J, HOLMES A P, WORSLEY K J, et al. Statistical parametric maps in functional imaging: A general linear approach[J]. *Human Brain Mapping*, 1994, 2(4): 189-210.
- [63] MUMFORD J A, NICHOLS T. Simple group fMRI modeling and inference[J]. *NeuroImage*, 2009, 47(4): 1469-1475.
- [64] MCLAREN D G, RIES M L, XU G, et al. A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches[J]. *NeuroImage*, 2012, 61(4): 1277-1286.
- [65] ZHANG T, KISHORE V, WU F, et al. BERTScore: Evaluating text generation with BERT[EB/OL]. (2019-04-21). <https://arxiv.org/abs/1904.09675v3>.
- [66] PAPINENI K, ROUKOS S, WARD T, et al. BLEU: A method for automatic evaluation of machine translation[C]// *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*. [S.l.]: ACM, 2002: 311-318.
- [67] BANERJEE S, LAVIE A. METEOR: An automatic metric for MT evaluation with improved correlation with human judgments[C]// *Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*. [S.l.]: ACL, 2005: 65-72.
- [68] QURASHI A W, HOLMES V, JOHNSON A P. Document processing: Methods for semantic text similarity analysis[C]// *Proceedings of 2020 International Conference on Innovations in Intelligent Systems and Applications (INISTA)*. Novi Sad, Serbia: IEEE, 2020: 1-6.
- [69] HINTON G, DENG L, YU D, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups[J]. *IEEE Signal Processing Magazine*, 2012, 29(6): 82-97.
- [70] FRIEDERICI A D. The brain basis of language processing: From structure to function[J]. *Physiological Reviews*, 2011, 91(4):

1357-1392.

- [71] 顾俊娟, 石金富. 汉字位置加工和词边界效应的认知机制[J]. 心理科学进展, 2021, 29(2): 191-201.
GU Junjuan, SHI Jinfu. The cognitive mechanism of Chinese character position processing and word boundary effect[J]. Advances in Psychological Science, 2021, 29(2): 191-201.
- [72] LI X, ZHAO W, POLLATSEK A. Dividing lines at the word boundary position helps reading in Chinese[J]. Psychonomic Bulletin & Review, 2012, 19(5): 929-934.
- [73] LI X, MA G. Word boundaries affect visual attention in Chinese reading[J]. PLoS One, 2012, 7(11): e48905.
- [74] IYER V, CHEN P, BIRCH A. Towards effective disambiguation for machine translation with large language models[EB/OL]. (2023-09-20). <https://arxiv.org/abs/2309.11668v2>.
- [75] WONG K F, LI W, XU R, et al. Introduction to Chinese natural language processing[M]. [S.l.]: Springer International Publishing, 2010
- [76] RYBÁŘ M, DALY I. Neural decoding of semantic concepts: A systematic literature review[J]. Journal of Neural Engineering, 2022, 19(2): 021002.
- [77] FRIEDERICI A D. The brain basis of language processing: From structure to function[J]. Physiological Reviews, 2011, 91(4): 1357-1392.
- [78] COCQUYT E M, LANCKMANS E, VAN MIERLO P, et al. The white matter architecture underlying semantic processing: A systematic review[J]. Neuropsychologia, 2020, 136: 107182.
- [79] HERTRICH I, DIETRICH S, ACKERMANN H. The margins of the language network in the brain[J]. Frontiers in Communication, 2020, 5: 519955.
- [80] MIDDLEBROOKS E H, YAGMURLU K, SZAFLARSKI J P, et al. A contemporary framework of language processing in the human brain in the context of preoperative and intraoperative language mapping[J]. Neuroradiology, 2017, 59(1): 69-87.
- [81] BROWN M, KUPERBERG G R. A hierarchical generative framework of language processing: Linking language perception, interpretation, and production abnormalities in schizophrenia[J]. Frontiers in Human Neuroscience, 2015, 9: 643.
- [82] FRIEDMAN N P, ROBBINS T W. The role of prefrontal cortex in cognitive control and executive function[J]. Neuropsychopharmacology, 2022, 47(1): 72-89.
- [83] TAJADURA-JIMÉNEZ A. Embodied psychoacoustics: Spatial and multisensory determinants of auditory-induced emotion [M]. Gothenburg: Chalmers University of Technology, 2008.
- [84] ZHANG Y. Neuroimaging investigations of cortical specialisation for different types of semantic knowledge[D]. Edinburgh: Edinburgh Research Archive, 2024.
- [85] 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[J]. Cerebral Cortex, 2009, 19(12): 2767-2796.
- [86] PATTERSON K, NESTOR P J, ROGERS T T. Where do you know what you know? The representation of semantic knowledge in the human brain[J]. Nature Reviews Neuroscience, 2007, 8(12): 976-987.
- [87] THYE M, HOFFMAN P, MIRMAN D. The neural basis of naturalistic semantic and social cognition[J]. Scientific Reports, 2024, 14: 6796.
- [88] ZHU Z, YANG H, WEN H, et al. Innate network mechanisms of temporal pole for semantic cognition in neonatal and adult twin studies[J]. Nature Communications, 2025, 16: 3835.
- [89] LIU L, LI H, REN Z, et al. The “two-brain” approach reveals the active role of task-deactivated default mode network in speech comprehension[J]. Cerebral Cortex, 2022, 32(21): 4869-4884.
- [90] FENG S, QI R, YANG J, et al. Neural correlates for nouns and verbs in phrases during syntactic and semantic processing: An fMRI study[J]. Journal of Neurolinguistics, 2020, 53: 100860.
- [91] YU M, WU Z, LUAN M, et al. Neural correlates of semantic and phonological processing revealed by functional connectivity patterns in the language network[J]. Neuropsychologia, 2018, 121: 47-57.
- [92] ZOU L, PACKARD J L, XIA Z, et al. Neural correlates of morphological processing: Evidence from Chinese[J]. Frontiers in Human Neuroscience, 2016, 9: 714.
- [93] BLANK I A, FEDORENKO E. No evidence for differences among language regions in their temporal receptive windows[J].

- NeuroImage, 2020, 219: 116925.
- [94] SAUR D, SCHELTER B, SCHNELL S, et al. Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension[J]. NeuroImage, 2010, 49(4): 3187-3197.
- [95] THYE M, HOFFMAN P, MIRMAN D. The neural basis of naturalistic semantic and social cognition[J]. Scientific Reports, 2024, 14(1): 6796.
- [96] BRAMSON B, FOLLONI D, VERHAGEN L, et al. Human lateral frontal pole contributes to control over emotional approach-avoidance actions[J]. Journal of Neuroscience, 2020, 40(14): 2925-2934.
- [97] FEDORENKO E, BLANK I A. Broca's area is not a natural kind[J]. Trends in Cognitive Sciences, 2020, 24(4): 270-284.
- [98] ISAACS B R, FORSTMANN B U, TEMEL Y, et al. The connectivity fingerprint of the human frontal cortex, subthalamic nucleus, and striatum[J]. Frontiers in Neuroanatomy, 2018, 12: 60.
- [99] PHẠM D Đ, MCDONALD D J, DING L, et al. Less is more: Balancing noise reduction and data retention in fMRI with data-driven scrubbing[J]. NeuroImage, 2023, 270: 119972.

作者简介:



马磊(1983-),男,副教授,研究方向:人工智能、生物信息采集与处理,E-mail:ma-
lei@ntu.edu.cn。



崔文浩(2000-),男,硕士研究生,研究方向:人工智能、认知神经科学,E-mail:
1066459980@qq.com。



杨汶汶(1985-),男,教授,研究方向:人工智能、数字信号处理,E-mail:wwyang20
08@hotmail.com。



王朝欣(2001-),通信作者,男,硕士研究生,研究方向:人工神经网络、认知神经科学,E-mail:919874984@qq.
com。

(编辑:王婕,陈琨)