

多层级语言脑解码技术的研究前瞻与发展路径

程凯文¹, 杜圆圆¹, 赵宇浩²

(1. 四川外国语大学语言智能学院, 重庆 400031;
2. 广东外语外贸大学英语语言文化学院, 广东 广州 510420)

摘要: 语言脑解码技术为失语患者重建交流能力提供了关键路径, 也是脑机接口技术发展的重要驱动力之一。近年来, 基于侵入性与非侵入性脑信号的多层级语言脑解码技术取得了显著进展, 但不同语言层级在脑信号表征、任务复杂度与模型方法上存在明显差异, 导致相关研究在技术路径和问题侧重上比较分散, 尚缺乏系统的层级化梳理。鉴于此, 系统梳理了2008—2025年语言脑解码研究在音素、声调、单词、短语、句子及语篇各层级的主要进展, 分析了脑信号类型、任务设定与模型架构对解码精度的影响, 并就多层级解码、开放词汇解码及预训练语言模型的融合应用等问题进行了重点讨论。在此基础上, 提出了未来语言脑机接口研发与智能化语言康复的若干研究路径, 以期对相关交叉研究提供理论支撑与技术参考。

关键词: 大脑信号; 语言脑解码; 人工智能; 语言脑机接口; 多层级解码

中图分类号: TP391.9

文献标志码: A

doi: 10.11959/j.issn.2096-6652.ZNKX26030

Prospects and pathways for research on multi-level brain language decoding

Cheng Kaiwen¹, Du Yuanyuan¹, Zhao Yuhao²

1. College of Language Intelligence, Sichuan International Studies University, Chongqing 400031, China
2. Faculty of English Language and Culture, Guangdong University of Foreign Studies, Guangzhou 510420, China

Abstract: Brain language decoding provides a critical pathway for restoring communication in patients with aphasia and serves as an important driving force for the development of brain-computer interface (BCI) technology. In recent years, substantial progress has been made in multi-level brain language decoding based on both invasive and non-invasive brain signals. However, different linguistic levels vary markedly in brain signal representation, task complexity, and model selection, which has led to fragmented technical pathways and research emphases, while a systematic level-based synthesis is still lacking. Based on this, the major advances in brain language decoding research from 2008 to 2025 were reviewed across different linguistic levels, including phonemes, words, phrases, sentences, and discourse. The influence of brain signal modality, task setting, and model architecture on decoding accuracy was analyzed, and further key issues such as multi-level decoding, open-vocabulary decoding, and the integration of pre-trained language models were discussed. On this basis, several future research pathways were proposed for the development of language brain-computer interfaces and intelligent language rehabilitation, with the aim of providing theoretical support and technical reference for related interdisciplinary studies.

收稿日期: 2026-02-10; 修回日期: 2026-05-11

基金项目: 重庆市社科规划项目 (No.2023NDYB167); 西南交通大学美国中心年度项目 (No.ARC2023006); 教育部人文社会科学项目 (No.25YJA710007)

Foundation Items: Chongqing Social Science Planning Project (No.2023NDYB167), The Annual Project of the American Studies Center in Southwest Jiaotong University (No.ARC2023006), General Project of the Humanities and Social Sciences Research Fund of the Chinese Ministry of Education (No.25YJA710007)

Key words: brain signal, brain language decoding, artificial intelligence, language brain-computer interface, multi-level decoding

0 引言

大脑语言解码技术简称为语言脑解码技术,是将文本、图像、视频或语音等外部刺激所引发的脑神经活动转换为可读文字或语音的技术。该技术一方面为失语症患者恢复交流能力提供新的可能,另一方面也成为推动脑机接口研究的重要技术路径。然而,不同语言层级在脑信号表征、任务复杂度与模型方法上并不一致,现有研究在技术路径与关注重点上也比较分散,尚缺乏系统梳理。鉴于此,本文系统回顾了2008—2025年领域内在音素、声调、单词、短语、句子及语篇等各语言层级上的大脑语言解码进展,从语言认知与人工智能两个角度探讨了影响解码精度的内外部因素,并就未来发展方向进行了展望。

本文系统检索了Web of Science Core Collection、PubMed、Scopus、IEEE Xplore、中国知网(CNKI)和Google Scholar等数据库,检索时间为2008—2025年。检索关键词采用中英文组合形式,主要包括语言脑解码、脑语言解码、语言脑机接口、文本脑解码、brain language decoding、language brain-computer interface、brain-to-text、speech decoding from brain activity以及与EEG、ECoG、fMRI、MEG等与脑信号模态相关的组合检索词。文献筛选主要遵循以下标准:聚焦语言相关脑信号解码任务;涉及音素、单词、短语、句子或语篇等至少一个语言层级;优先纳入发表于期刊、重要会议论文集或具有代表性的前沿研究;重点关注在研究范式、模型方法、任务设计或神经机制解释方面具有代表性的成果。对于与主题关联较弱、信息不完整或重复发表的文献,本文未进行重点讨论。

1 大脑语言解码技术原理、路径和应用价值

1.1 原理与实现路径

大脑语言解码技术依据处理对象可分为语言内解码和跨语言解码。语言内解码指将特定语言刺激诱发的大脑信号转化为同语言的语义类别或文字描述^[1-2];跨语言解码则需将与源语言相关的神经活动模式转化为目标语言的对应语义表征^[3-4]。从技

术实现角度来看,当前研究主要沿两条路径展开:一是分类解码,通过识别脑信号特征与语义标签的映射关系,对视听等多模态语言刺激引发的脑活动进行语义分类^[5];二是生成式解码,以深度学习模型为基础^[6-7],通过捕捉脑信号与语义内容的分布关系,直接从神经活动中重构单词、短语乃至完整语句。近年来,Transformer架构^[8]的引入进一步提升了生成式解码的性能。上述两条路径的整体技术流程如图1所示,以文本、语音、图像或视频等多模态刺激及对应脑活动为输入,分别提取语义特征与神经特征,建立二者之间的映射关系,再通过机器学习训练模型,实现对新输入脑信号的语义预测,输出文本或语音。

1.2 数据来源与模型算法选择

在当前的大脑语言解码研究中,数据采集主要依赖于脑电图(electroencephalogram, EEG)、皮层脑电图(electrocorticography, ECoG)、功能磁共振成像(functional magnetic resonance imaging, fMRI)以及脑磁图(magnetoencephalography, MEG)等神经影像技术。各类技术在信号特性与采集方式上存在明显差异,对解码任务的选择与效果均有直接影响。EEG技术具有安全无创和时间分辨率高的优势,能够细致追踪语言加工的动态过程,在脑机接口与语言解码研究中应用最广泛。ECoG因电极直接贴附于大脑皮层表面,信号干扰小,能够提供比EEG更高空间分辨率的局部脑活动信息,但其属于侵入性技术,通常仅限于特定医学和科研场景中使用^[9-10]。相较之下,fMRI的空间分辨率高,能够精确定位与语言处理相关的脑区,而且具备非侵入性特点,是目前大脑语言解码研究中最常用的数据来源之一^[1-2,11]。MEG兼具高时间分辨率与良好的空间定位精度,而且对神经元活动的探测比较直接,与ECoG同属非侵入性技术,是研究大脑快速动态功能和临床应用的理想工具,但其设备成本高,对操作环境要求严苛,在语言解码领域的应用目前仍比较有限^[12]。

处理不同语言层级的刺激时,其语义复杂度和时序组织方式并不相同,因此选择合适的解码模型算法对于提升解码准确率至关重要。结合本文纳入并重点讨论的代表性研究文献统计结果(如图2所

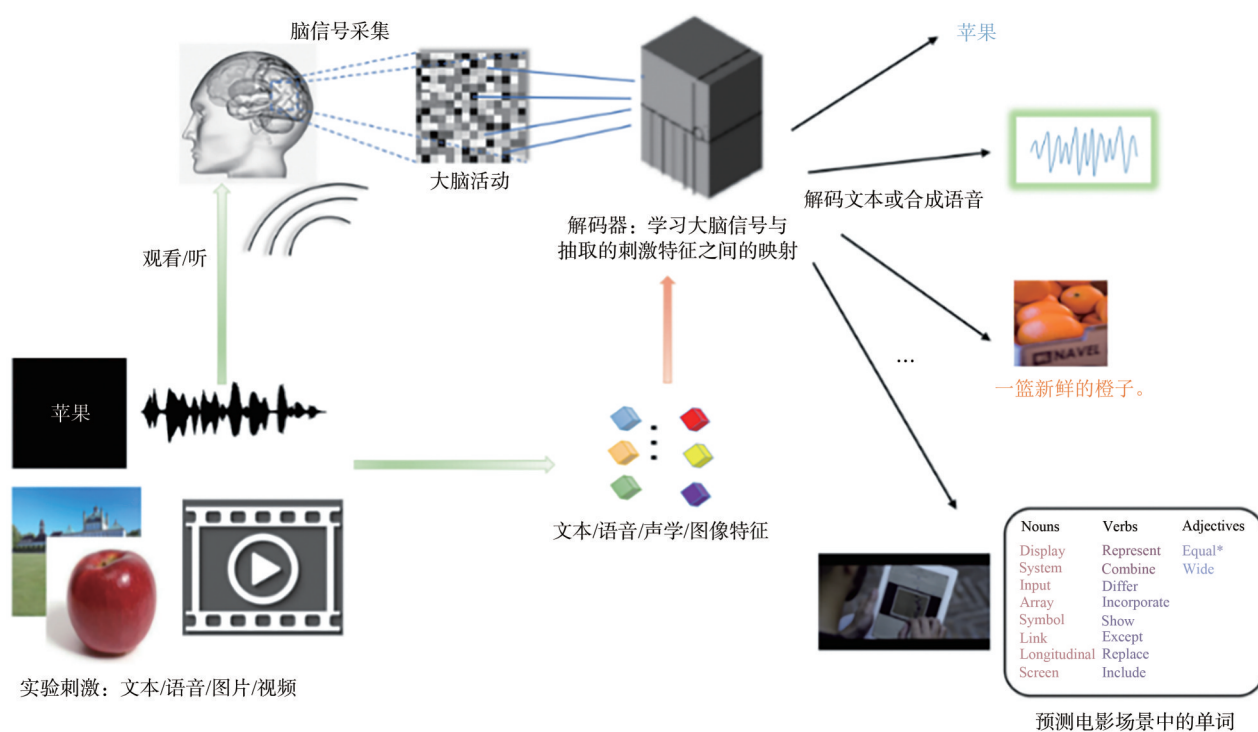


图1 大脑语言解码流程示意图

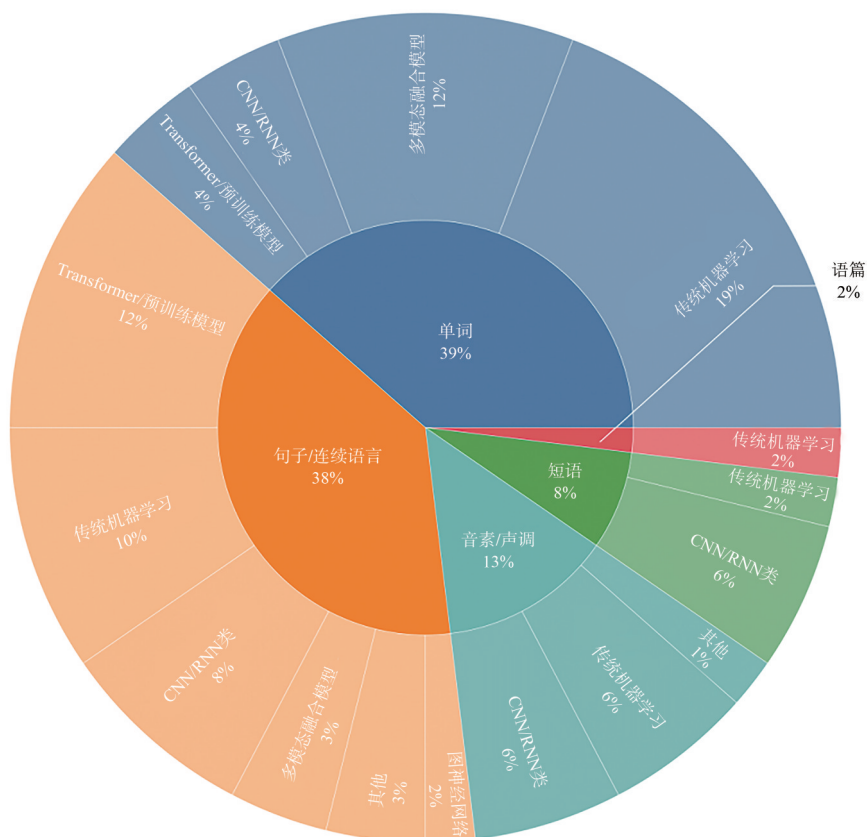


图2 多层次语言脑解码研究模型算法类型分布

示)可以发现,不同语言层级在模型方法上呈现出比较清晰的分布趋势。音素、声调及单词等相对基

础的语言单位,研究中常采用传统机器学习方法以及卷积神经网络(convolutional neural network,

CNN)、循环神经网络(recurrent neural network, RNN)等模型^[3,10,12],这类方法在处理局部时序特征、声学线索和相对清晰的类别边界时往往更直接有效。相比之下,短语、句子及连续语言等更高层级的语言解码任务往往涉及更长时程的语义整合与跨模态信息协同,因而更倾向于依赖长短期记忆网络(long short-term memory, LSTM)、门控循环单元(gated recurrent unit, GRU)、Transformer及其与预训练语言模型结合的深度学习框架^[13]。其中,句子及连续语言层级在模型使用上最多样,不仅广泛采用CNN、RNN及其变体,也逐渐形成了以Transformer、预训练语言模型和多模态融合框架为主流的技术范式。相较而言,图神经网络类方法在现有研究中的占比仍较小,主要应用于需要显式建模语义关系或结构依赖的任务,而语篇层级研究目前数量仍较少,长程上下文表征与复杂语义组织仍是制约该层级发展的主要难点。除了解码模型本身,不同刺激表征模型对脑信号解码的影响同样不可忽视。用于表征刺激内容的视觉模型、词嵌入模型或句子表征模型会改变输入特征的结构,从而影响后续解码器的训练效果与性能表现。例如,Holm等^[14]基于EEG记录被试快速视觉感知时的脑活动,分别采用视觉几何组网络-19(visual geometry group network-19, VGG-19)和视觉Transformer模型探究视觉皮层与深度神经网络之间的层级时间对应关系,发现早期脑电成分与网络浅层激活相关,晚期脑电成分与深层激活相关,表明不同架构的神经网络在视觉信息处理上可能收敛于相似的计算策略。此外,对于需要分词的语言,选择合适的分词工具同样重要。在抽取刺激的语义特征时,使用不同的分词工具可能会导致不同的分词结果,进而影响解码模型对刺激语义特征的训练输入。例如,中文分词常用的工具有结巴分词包^[15]和哈尔滨工业大学的分词包^[16],这些工具会导致文本分词存在一定差异,从而对模型训练产生影响。

1.3 脑机接口与临床应用

大脑语言解码技术通过建立语言音、形、义与神经活动的映射关系,在医学诊疗与脑机接口领域的应用已取得实质性进展,主要体现在以下3个方面。在神经功能障碍的辅助交流方面,该技术为肌萎缩侧索硬化症等言语障碍患者提供了不依赖肌肉运动的交流途径,通过语音合成系统将大脑语义信息转化为语音,帮助患者恢复沟通能力^[17]。Anu-

manchipalli等^[9]在癫痫患者研究中取得重大成果,开发了两阶段神经解码架构,将ECoG信号先解码为发音运动表征,再合成为可听语音,合成速度约为150词/分钟,首次实现了把神经活动直接合成完整句子的目标。在人机交互方面,语言脑解码技术也带来了新的可能。研究人员成功通过解码运动想象信号实现机械臂操控^[18]与智能轮椅导航^[19]。瘫痪患者可以通过想象书写动作达到90字符/分钟的文本输出速度^[20]。此外,“言语神经义肢”这一概念也逐渐受到关注,相关研究尝试通过解析与声道控制相关的脑电波,帮助中风患者重建语言交流能力。复旦大学附属华山医院吴劲松团队与中国科学院上海微系统与信息技术研究所周志涛团队报告的一项实时汉语全音节脑机接口解码研究,采用256通道ECoG植入系统对394个汉语音节进行实时神经解码,中位音节识别准确率达71.2%,单音节解码时延为65 ms,实时语句解码速率达49.6字/分钟。系统采用音节与声调并行的双流解码架构,专门针对汉语声调语言特点进行设计,在渐冻症、脑干卒中等失语患者群体中展现出较强的临床应用潜力^[21]。在认知障碍研究领域,fMRI研究显示,ADHD儿童顶叶下回与顶上小叶的功能连接性显著降低^[22];EEG证据表明,其视觉信息筛选存在神经编码紊乱^[23];Xu等^[24]发现,不同类型语音刺激引发的节律反应与意识障碍程度密切相关,并构建了机器学习模型作为诊断意识状态的客观量化工具。上述研究在深化病理机制理解的同时,也为基于语言刺激的临床诊断工具开发提供了新思路。未来,随着人工智能与神经工程的融合发展,大脑语言解码技术势必会重塑传统的医疗范式。

2 多层级大脑语言解码研究进展

经筛选,本文最终纳入与多层级语言脑解码直接相关的代表性文献59篇,其中英文文献51篇、中文文献8篇。从文献语言类型来看,现有研究以英文文献为主,中文撰写的公开成果相对较少,而且其中相当一部分为硕博学位论文。需要说明的是,中文文献数量较少并不意味着中文语境下的语言脑解码研究缺乏关注。相反,已有不少国内学者围绕汉语声调、中文词汇、句子乃至连续语言解码等方向开展了系统性研究,此类成果主要通过英文期刊与国际会议论文进行传播与交流。从文献研究主题来看,本文通过系统性梳理与多维度对比分

析,发现语言脑解码研究在语言层级、信号模态及任务范式等维度呈现出显著差异,见表1。总体看来,音素与单词层级研究仍以分类任务为主:早期多依托支持向量机、线性回归等传统机器学习方法,近年来逐渐扩展至声调语言建模、大词汇量特征提取及开放词汇解码。单词层级因研究积累最丰富,正从封闭集分类走向跨语言语义对齐与多模态融合。相比之下,短语、句子及连续语言解码更早引入深度神经网络架构,卷积神经网络、循环神经网络及其变体被广泛采用;当前,Transformer、预训练语言模型和多模态融合框架已逐渐确立为主流范式,任务范畴也从分类与生成拓展至连续语言重构和零样本匹配。现阶段,句子层级是该领域发展最活跃的前沿方向,而语篇层级研究尚处起步阶段,受限于长程上下文建模的难度与解码精度瓶颈,相关探索相对匮乏。后文将基于上述语言层

级,详细梳理近10年来语言脑解码研究在脑信号表征、任务复杂度与模型方法上的进展及突破。

2.1 音素和声调解码进展

在音素和声调解码领域,相关研究已逐步形成从信号预处理、特征提取到分类建模的相对完整的技术路径^[25-26]。杨淑淇等^[27]从多层感知机、循环神经网络、卷积神经网络及Transformer等架构出发,梳理了从EEG、MEG和ECoG信号中解码语音信息的技术进展,指出当前该领域的核心挑战是如何提高非侵入式方法的信噪比与解码准确率。首先,音素与声调作为解码的最小语音单元,其神经编解码研究已建立标准化实验范式。研究者通常采集被试在听觉或视听联合条件下的EEG、ECoG信号,利用机器学习模型构建脑信号特征与语音单元间的映射关系,揭示语音在大脑中的编码模式。例如,Ramsey等^[10]通过记录5名癫痫患者发音时感觉运

表1 多层次语言脑解码代表性研究归纳

语言层级	代表性文献	常用脑信号模态	主要任务类型	常用模型/方法	研究特点与发展趋势
音素、声调	Ramsey等 ^[10] Liu等 ^[28] Guo等 ^[72] Zhang等 ^[30] Chen等 ^[31] Littlejohn等 ^[32]	• EEG • ECoG • fNIRS	• 音素分类 • 声调分类 • 语音特征分层解码 • 实时语音合成	• SVM • 传统分类方法 • CNN、RNN类模型 • 深度学习+语音合成 • 双向LSTM	以分类任务为主,强调对低层语言单位的快速识别,近年来开始向声调语言、大词汇量语音特征建模和实时解码拓展
单词	Mitchell等 ^[1] Buchweitz等 ^[38] Correia等 ^[3,39] Fukuma等 ^[33] Du等 ^[40] Wang等 ^[43] Bruera等 ^[42] Jia等 ^[45] 刘晓龙等 ^[44] 邹宇童等 ^[37] 任子琪 ^[41]	• fMRI • EEG • ECoG • MEG	• 单词语义预测 • 跨语言语义分类 • 词性解码 • 开放词汇解码 • 中文词汇解码 • 专有名词语义解码	• 线性回归 • 朴素贝叶斯 • 分类模型 • 多模态模型 • 序列到序列模型 • 对比学习 • 搜索光解码 • CAM-Net • VAE+ECA • BLM	是研究最集中的层级之一,早期以传统机器学习为主,后续逐渐发展出多模态融合和开放词汇解码,跨语言语义共性成为重要议题,非侵入式中文词汇解码也开始起步
短语	Dash等 ^[12] Matsuo等 ^[47] Huang等 ^[13] Zhao等 ^[46]	• MEG • EEG • fMRI	• 短语分类 • 短语生成 • 自然阅读短语解码 • 短语边界检测	• CNN • VGG-LSTM • DC-LDM • 滑动窗口解码器	位于单词到句子的过渡层级,研究数量相对较少,从分类逐步转向生成式建模,并开始关注句法结构的神经表征机制
句子、连续语言	Yang等 ^[48] Li等 ^[49] Anumanchipalli等 ^[9] Tang等 ^[2] Wang等 ^[53,55] Zhang等 ^[51] Lévy等 ^[57] Défossez等 ^[56] 马磊等 ^[52]	• fMRI • EEG • MEG	• 跨语言句子解码 • 句子级语音合成 • 连续语言重构 • 脑文本解码 • 零样本匹配 • 中文连续语义重构	• 双向LSTM • 图注意力网络 • Transformer • BART • GPT类模型 • 预训练框架 • Brain2Qwerty • 对比学习 • CLCSD	是当前发展最快的层级,逐渐从封闭分类走向开放式生成与连续语言重构;Transformer、预训练模型和多模态融合成为主流方向
语篇	Dehghani等 ^[58]	fMRI	故事/语篇语义分类	doc2vec与分类模型	研究仍较少,语篇级解码对上下文建模要求高,整体准确率仍有限,长程上下文建模与解码精度仍是主要瓶颈

动皮层的 ECoG 信号, 采用支持向量机模型成功实现 /p/、/k/、/u/、/a/ 这 4 个口语音素的分类解码, 首次证实运动皮层在语音生成中的特异性编码机制。其次, 针对汉语等声调语言的特殊性, 研究者也开发了针对性的解码策略。Liu 等^[28]通过分离音节与声调的解码任务, 选取最小对立音节对 “mā/má/mǎ/mà”, 在精确控制阴平 55、阳平 35 等调域与调型特征后, 将汉字声调解码精度提升至 91.4%。

声调作为超音段特征, 其神经编码在喉部运动皮层中表现出不同于其他发音特征的独特模式。研究发现, 双侧喉部运动皮层 (bilateral laryngeal motor cortex, LMC) 编码的是声调产生过程中的发音运动动力学信息, 而非离散的声调类别, 而且音高升高与音高降低分别由两种不同的神经群体活动模式控制^[29], 这意味着, 声调解码不仅涉及类别判别, 还需要处理连续发音运动中的动态变化。围绕这一特点, 后续研究一方面尝试通过优化刺激输入增强脑区激活的可分离性, 另一方面尝试借助更强的模型框架提升语音重建能力。Zhang 等^[30]发现, 相较于单一拼音视觉刺激, 拼音加语音的视听联合刺激使声调分类准确率从 67.7% 提升至 80.1%, 提升幅度达 18.32%, 表明多通道输入会引发更强且更可分离的脑区激活, 从而提升声调信息的解码效果。Chen 等^[31]则提出结合深度学习与语音合成的神经语音解码框架, 通过构建可解释的低维度声学空间, 如音高、响度等韵律特征, 实现了高精度语音重建、听觉重复、视觉命名、句子生成等多任务范式, 整合了多模态信息以增强语音特征表征能力, 在连续语音重建任务中显著提升了语音质量与可理解性。此外, 相关研究也开始从受控实验条件逐步转向更接近真实交流需求的应用场景。Littlejohn 等^[32]采用高密度 ECoG 技术, 从一名因脑干中风失语 18 年的患者的无声说话意图中实时解码英语音素序列并合成语音, 解码速度在流式语音合成中约为 46.5 词/分钟, 文本输出可达约 70~75 词/分钟, 而且成功还原了患者发病前的个性化声音特征。与前述基于视听联合刺激的被动解码范式不同, 该研究验证了主动发声意图驱动的音素级实时解码, 并能泛化至未训练的新词。这表明, 音素解码在真实临床场景中具备实用可行性。

2.2 单词语义解码进展

单词语义解码通过解析文本、语音、图像或视频刺激引发的神经活动来重构单词语义信息,

其研究演进可分为 3 个阶段。早期奠基性工作由 Mitchell 等^[1]开启, 他们首次运用线性回归模型解码 fMRI 信号, 成功预测 “celery” 和 “airplane” 等 60 个英语具体单词的语义特征, 证实了大脑活动与词汇语义的映射关系。这一发现为后续研究提供了方法论基础与理论范式。在此之后, 相关研究开始进一步扩展单词层级解码的对象与任务边界。在单词词性解码方面, Fukuma 等^[33]提出的视频分割、语义向量建模、脑信号映射三阶段框架, 成功解码出视频场景中的名词、动词和形容词。然而, Nishida 等^[34]发现不同被试解码出的单词语义信息存在个体差异, 这可能与个体间脑激活模式不尽相同有关, 说明解码模型需要针对个体进行校准, 以维持模型的泛化能力。Lin 等^[35]进一步发现语言生成与感知过程诱发高度相似的神经活动模式, 这表明, 语义网络在产出与理解两个方向上具有统一性。最近, 词汇产出机制研究^[36]揭示, 从上下文到词汇、音节再到字母的自上而下级激活序列, 依赖于多层级表征的并行协调。另外, 在模型架构方面, 邹宇童等^[37]提出集成高效通道注意力机制与变分自编码器的深度学习解码模型, 在 Pereira 等^[11]的 180 词数据集上成对准确率高达 83.4%, 优于 Ridge 回归、CNN 等基线模型。

在证明单词语义可被有效解码之后, 研究者进一步将关注点转向跨语言条件下语义表征的脑神经解码问题。跨语言解码研究表明, 语义表征在神经层面具有跨语言的共性。Buchweitz 等^[38]采用高斯朴素贝叶斯模型, 通过分析英葡双语者观看 14 类名词图片 (如 “hammer” 和 “house”) 的脑信号, 实现了跨语言语义分类, 准确率显著超越机会水平, 这一发现表明不同语言在语义加工层面共享神经基质。在听觉词汇解码方面, Correia 等^[39]发现双语者处理名词语音时, 语言内解码和跨语言解码之间存在显著的时程差异: 前者在词 onset 后约 50 ms 时出现, 并持续至 620 ms 前后, 后者则在约 550 ms 才出现。这个发现说明, 语音特征的自动化处理早于语义概念的跨语言整合, 为优化语言脑解码提供了时间窗口选择的理论依据。

近年来, 多模态融合与词汇规模扩大逐渐成为不同脑信号源中单词解码的重要进展方向。Du 等^[40]提出 BraVL (brain-visual-linguistic) 模型, 整合图片视觉特征、维基百科文本语义特征与 fMRI 脑信号, 建立视觉、语言与神经活动之间的联合映射,

实现了对新视觉类别的零样本神经解码。实验结果显示,融合视觉与语言特征的解码效果明显优于仅使用单一模态的效果,加入文本特征后平均脑解码精度提升约6%。该研究验证了从脑活动中解码训练数据中未见类别的可行性,为从有限训练数据向开放词汇泛化提供了方法参考。任子琪^[41]提出脑语言模型(brain language model, BLM),通过脑编码器、脑文转换器与文本解码器三部分架构,基于稀疏掩码建模的自监督预训练从fMRI信号中捕获视觉与语义信息并生成文本描述,在NSD(natural scene dataset)数据集上,该方法在文本解码任务中的表现优于现有方法,且在METEOR、Rouge、CLIP-T等多项指标上均表现更优。此外,Bruera等^[42]采用EEG时间分辨与搜索光解码分析,发现人物与地点专有名词的语义类别可在词汇语义加工的典型时间窗口内被准确解码,而且人物熟悉度的神经表征比地点更易区分。空间分析进一步揭示,左侧后外侧脑区对两类专有名词均有响应,而右侧额颞区仅对人物表现出特异性激活。如果说前述研究更多是在既有语义空间内提升表征能力,那么另一条重要的发展方向则是扩大可解码词汇的规模。Wang等^[43]利用EEG脑电信号,实现了从封闭词汇到开放词汇解码的突破,将可解码词汇规模从传统的几百个词汇扩展至50 265个词汇。而且,该研究采用序列到序列的神经机器翻译框架,将连续的EEG特征向量直接映射到离散的英语词汇,实现了从神经信号到符号表征的端到端转换。在中文语境下,相关研究也逐步从方法验证走向任务拓展。刘晓龙等^[44]针对中文词语设计了三分类实验范式,选取“帮帮我”“坐起来”“翻身”3个具有临床实用价值的词语作为想象任务,提出融合多尺度时间卷积、不对称空间卷积、通道注意力机制与双向LSTM的CAM-Net模型,在15名健康受试者的64通道EEG数据上实现了48.54%的三分类准确率,二分类准确率最高达64.17%,优于多个基准模型。Jia等^[45]基于脑磁图(MEG)构建了首个面向中文语音脑机接口的数据集Chinese-MEG48,该数据集涵盖48个高频中文词汇。该研究提出多模态辅助言语解码(multimodal assisted speech decoding, MASD)算法,通过对比学习将MEG信号与文本语义特征和合成语音声学特征对齐,在48个词分类任务中实现了46.21%的解码准确率,显著高于随机水平,验证了非侵入式MEG技术解码中文词汇神

经表征的可行性。

2.3 词组与短语解码进展

词组与短语层级的神经解码研究沿分类式与生成式两条路径推进。短语既不同于单个词汇的局部语义识别,又未达到完整句子要求的长时程语义整合,因此这一层级的研究往往同时涉及“短语内容能否被区分”和“短语结构如何被组织”两个方面。一方面,在分类式解码领域,Dash等^[12]通过设计感知、想象与发音三任务范式,采用卷积神经网络分析MEG信号,发现在发音任务中(如“Do you understand me”和“How are you”等日常短语)的解码准确率达96.65%,该现象源于运动皮层调控发音动作与听觉皮层处理语音反馈的协同激活,其耦合信号携带的时空特征优于单一任务模式。另一方面,基于向心结构理论,Zhao等^[46]训练EEG词性滑动窗口解码器,区分动词与副词的神经信号,并将其应用于短语边界检测。研究通过汉语语法助词“得”与“了”操纵动词短语边界:使用“得”时目标副词位于短语末尾,使用“了”时副词位于短语中间,两条件词序与实词完全相同。结果显示,当副词出现在短语边界时,约600 ms后动词解码概率显著升高,而且该效应不受概念联想强度或词汇可预测性调节,同一时间窗口还出现负波成分及 β - δ 频段试次间相位一致性增加,这表明大脑在构建短语表征时会重新激活中心语信息。

相较于分类式研究聚焦短语内容或结构的判别,生成式解码则更进一步,尝试直接从脑活动中恢复短语级语言表述。Huang等^[13]提出双通道语言解码模型(dual-channel language decoding model, DC-LDM),并采用渐进迁移训练策略,通过CNN特征提取、双向GRU时序建模与Transformer语义融合的三级架构,显著提升了自然场景图片的短语解码性能。由此可见,短语层级的研究虽然数量不多,但已经清楚地呈现出从局部分类识别向更高语义粒度生成建模过渡的趋势。

2.4 句子和语篇解码进展

相较于词汇层级,句子加工涉及更长时间尺度的语义整合。早期研究受限于浅层模型架构,Matsuo等^[47]采用VGGN(VGG network)-LSTM混合模型解码图片诱发脑活动,生成描述的语义一致性仅58%,例如一张撑着伞的女性图片被误解为“A man is in the back of an umbrella”。随后,Yang等^[48]通过英语、葡萄牙语和汉语三语种的句子理解任

务,采用fMRI分析跨语言语义表征,研究首先筛选出3种语言加工过程中均稳定激活的脑区集合,并将其定义为语言通用簇,这些区域主要涉及额叶、颞叶和顶叶等脑区。结果显示,尽管汉语和葡萄牙语在表层结构上存在显著差异(前者为孤立语,后者为屈折语),但语言通用簇中的神经激活模式仍表现出较高的跨语言一致性。定量结果显示,双语训练模型的分类准确率为66.8%,高于单语训练模型的62.4%,且这种双语优势具有明显的概念域选择性,主要体现在社交互动、心理活动等抽象概念域,从而将Buchweitz等^[38]在词汇层面的发现进一步延伸至句子层级。Li等^[49]将研究焦点从fMRI拓展至EEG,探索双语者中英文句子的默读信号解码。结果显示,虽然两种语言的复杂语义在大脑中激活相似的神经网络模式,但解码表现具有明显差异,中文解码率为44.55%,英文解码率为54.70%。这个结果表明,语言类型学特征仍然影响句子神经表征的复杂程度。

相较于前述以句子意义识别和跨语言比较为主的研究,生成式句子解码更强调直接从脑信号中重构自然语言输出^[50]。Anumanchipalli等^[9]在5名癫痫患者中植入高密度ECoG电极阵列,开发了两阶段解码架构:第一阶段通过双向LSTM将神经信号解码为发音运动表征,第二阶段将发音运动转换为可听语音,合成速度可达150词/分钟。这表明基于发音相关脑信号的LSTM模型能够促成句子级语音合成,在自然度上相较传统模型有很大提升。在此基础上,Tang等^[2]采用GPT-1构建语义解码器,利用被试累计16h的叙事故事聆听fMRI数据进行训练。测试阶段,解码器对全新故事的语义重构采用BERTScore评估,相似度范围为0.807 7~0.811 6,覆盖1 200类语义主题,并能解码想象语音和无声视频的语义内容,这表明连续语言的语义表征可在多个皮层区域中被独立解码,而且具有一致性与冗余性。在中文语境下,Liu等^[28]和Zhang等^[51]分别在声调语言语音解码与自然声调句子的脑文本解码方面取得重要进展,进一步说明了中文连续语言脑解码的可行性。在连续语义重构方向,马磊等^[52]提出中文长序列连续语义解码器CLCSD,通过岭回归编码模型建立GPT语义嵌入与fMRI脑响应之间的映射,并引入语速模型估计刺激语义的时间点,在SMN4Lang公开数据集上,CLCSD采用束搜索生成语义序列,BERTScore F1为0.673,颞叶脑区

的解码效果最佳。总体看来,围绕连续语言解码,不同脑成像技术在语言信息提取中各有侧重,MEG对快速变化的时序与声学特征更敏感^[53],fMRI则更适合刻画高层语义信息^[2]。

Wang等^[43]在文本脑解码研究中不仅在单词层级实现了开放词汇解码,也将BART模型引入脑文本句子生成任务,这一由Lewis等^[54]提出的预训练序列到序列语言模型兼具BERT的双向编码和GPT的自回归解码优势,能够通过去噪自编码学习上下文语义表示。研究者将多层Transformer编码器提取的EEG特征映射到BART嵌入空间,使模型不仅能生成语法正确的句子,还能输出“appointed”与“elected”等语义相近的表达,表明大脑语言理解更依赖抽象语义而非具体词形。随后,Wang等^[55]进一步融合跨模态预训练模型与BART,通过联合学习词级与句级EEG特征,成功重构自然句子,并发现提高训练语料的多样性可有效增强模型性能,为高精度脑语言接口的发展提供了有力支持。这一系列研究从语音到文本、从单词到句子,逐步构建起一条完整的生成式语言解码技术链条,系统展现了该领域从基础单元识别到复杂语言结构生成的渐进式演进路径。

在生成式句子解码逐步成熟之后,相关研究开始进一步突破固定句集、封闭类别和受控实验范式的限制,推动脑语言解码向更开放、更自然的应用形态发展。Défossez等^[56]提出的零样本解码模型通过对比学习将MEG信号与wav2vec 2.0提取的语音表征对齐,构建神经活动与语义特征之间的隐式映射,在上千候选句中平均top-1匹配准确率达41%,最佳被试可达80%,其主要依赖高层语义表征而非低层声学特征。与以英语连续语言为主的开放式重构研究相比,中文句子层级的探索则更多体现为从较低层级语言单位向完整句子解码的延伸。Zhang等^[51]在前期解码单音节声调语音^[28]的基础上,进一步搭建了汉语语句的解码算法架构,将汉语脑解码从“字”的层面拓展至“句”的层面,首次实现了从“意念”到汉语语句文本的解码。后续研究进一步将解码推进到连续自然语言的重构,Wang等^[53]基于MEG信号提出连续语言语义重构框架,将连续词嵌入重构模型与GPT语言模型相结合,在被试听取《福尔摩斯探案集》过程中生成开放词汇的自然连续文本,解码器能100%准确地从候选集中识别出受试者所想的对应故事,重建文本的BERTScore达

0.816, 其整体表现与 Tang 等^[2]基于 fMRI 的连续语言语义重构研究处于相近水平, 证实 MEG 更高的时间分辨率使其在实时解码场景中更具优势。同样, Lévy 等^[57]提出 Brain2Qwerty 模型, 记录被试打字任务的脑信号并对其进行解码, MEG 平均字符错误率为 32%, 显著低于 EEG 的 67%, 且解码错误模式与键盘空间布局高度相关, 这也表明语言产出涉及将抽象符号映射至空间位置与运动序列的认知过程。

语篇层级的脑解码目前仍处于探索阶段, 仅有少数研究尝试对语篇级脑信号进行语义分类。Dehghani 等^[58]对来自 3 个国家的各 30 名被试阅读 40 篇英语、汉语普通话和波斯语语篇时的脑信号进行了分析, 结果显示, 被试阅读故事时的默认模式网络激活具有跨语言一致性, 基于故事级语义向量的解码模型在语言内和跨语言条件下均实现了显著高于随机水平的分类准确率, 平均为 55%~58%, 最高可达 64%。这一结果表明神经语义编码发生在高于单词和句子的层级, 而且在个体和语言间具有系统性, 但与真正解码语篇层级的复杂语义结构相比仍有较大差距。

3 影响大脑语言解码效果的因素

现有研究已在英语、汉语、葡萄牙语等多种语言中开展了大脑语言解码, 涵盖文本、语音、图像和视频等多模态刺激, 解码层级从单词、短语延伸至句子。不同层级的大脑语言解码效果和效率均存在差异。下面从语言内和语言外两个维度来探讨影响语言脑解码效能的各种因素。

第一, 语言词汇形态特征显著调节大脑语言解码效能。Hakala 等^[59]在分析芬兰语多语素名词的 MEG 数据时发现, 词长与词频是预测解码准确率的关键变量, 词长范围为 6~15 字符, 词频范围为每百万词 0.02~60 次。研究显示, 含 2~5 个语素的多语素词汇解码表现优于单语素词汇, 这表明解码过程高度依赖于对形态结构信息的有效整合。此外, 该研究还发现, 训练语料频率的差异并未引起多语素词解码准确率的大幅变化, 这一结果与常见的词频效应^[60]有所出入, Hakala 等推测这可能出于 3 个方面的原因: 首先是基于语素的语义向量解码算法本身对词频相对不敏感; 其次是在芬兰语这类形态丰富的语言中, 多语素构词结构对解码贡献可能强于词频; 此外, MEG 的高时间分辨率使该研究能精准定位 350~500 ms 的关键解码窗口, 而

词频效应或许在其他时间窗口才有所体现。由此可见, 词频对语言脑解码效能的调节作用在不同语言和模态下可能呈现异质性。

第二, 词汇的语义特征(具体与抽象、动态与静态等)比词类等语法特征更能影响大脑语言解码的性能。Wang 等^[61]研究表明, 相较于英语名词, 英语动词的解码效果最佳, 可能因为动词激活了与动作相关的广泛脑区, 包括运动前区和辅助运动区; 具体词比抽象词更容易被解码, 其中具体动词的解码效果最佳; 相比具体名词, 抽象名词更依赖与语言相关的脑区和默认模式网络的激活; 具体名词和具体动词之间的神经表征差异明显大于抽象名词和抽象动词之间的差异。这一发现与双编码理论一致, 即具体词在大脑中同时有视觉编码和语言编码, 而抽象词则更多依赖语言编码^[62]。另外, Huang 等^[63]的最新研究表明, 相比于传统的词类划分, 词汇承载的内在语义特征更能预测神经加工模式, 并且是触发大脑视觉双通路(背侧通路和腹侧通路)功能分离与融合的重要因素。研究发现, 动词以及具有运动属性的名词在解码中表现尤为突出。具体而言, 背侧通路在解码动词以及“Flying”和“Standing”等动态特征名词时的准确率显著高于腹侧通路, 腹侧通路只能有效解码“Field”和“Snow”等与静态场景相关的名词。这一发现表明, 词汇包含的动作语义和动态特征会制约其依赖的神经加工通路: 当词汇涉及动作表征与时序信息叠加时, 背侧通路及相关额顶运动网络的参与程度更高; 当词汇主要对应静态物体或场景表征时, 则更依赖腹侧通路支持的视觉语义加工。

第三, 在跨语言层面, 语言类型学特征和第二语言熟练度是影响跨语言解码效果的重要因素^[4,64]。Yang 等^[48]的 fMRI 实验显示, 句级语内解码准确率显著高于跨语言条件, 表明不同语言的神经表征存在特异性。第二语言和母语刺激激发的大脑活动存在差异^[48,65], 直接影响了解码模型的训练数据质量, 进而降低跨语言解码准确率。不过, 这种差异并不意味着不同语言之间完全缺乏可对齐的神经语义基础。Zinszer 等^[66]采用结构对齐方法发现, 中英两种语言的神经语义空间存在可对齐的共享结构。此外, 语言类型学差异也会影响跨语言的神经相似度, 例如, 正字法透明度相近的英语与葡萄牙语在梭状回的神经相似度高于类型学差异大的汉语与葡萄牙语^[67]。除语言类型学因素外, 个体的第

二语言熟练度也会影响跨语言神经表征的重叠程度。Liberto等^[68]从编码角度提供了直接证据,通过EEG结合时间响应函数分析发现,第二语言熟练度与大脑对声学、音位、音位组合规则及语义等多层级语言学特征的神经编码强度显著相关;更重要的是,基于这些神经信号可以较高精度地解码个体的第二语言熟练度水平,高熟练度者的神经表征模式明显更接近母语者。这说明,更高的熟练度通常意味着更强的跨语言神经对齐。然而,正因为跨语言解码的同时受到共享结构与语言特异性两方面因素的共同制约,其实际表现并不总是简单呈现为“母语优于第二语言”的单一模式,也有不少研究发现双语者在某些条件下母语的解码准确率反而低于第二语言,其具体原因值得深究^[49,69]。

第四,大脑对刺激的加工深度也会影响语言解码效果。Sheikh等通过掩蔽单词范式^[70]、阅读与心理模拟对比范式^[71]发现,语言内解码可在无意识或浅层加工条件下实现,而跨语言语义泛化则需要深层有意识的语义加工,且这一效应在顶下小叶、外侧颞叶、背内侧前额叶及额下回等多个语义脑区均有体现。Guo等^[72]基于功能性近红外光谱(functional near-infrared spectroscopy, fNIRS)技术探究言语想象生成速率对解码性能的影响,发现当被试以更快速率想象说“yes”或“no”时,言语相关脑区的神经活动更强,二分类准确率从62.81%提升至67.50%,进一步表明加工强度的提升有助于改善解码效果。另外,有意识加工比部分意识加工或无意识加工解码准确率更高。例如,西班牙语与巴斯克语双语者fMRI实验表明,当被试进行有意识的动物名词分类时,语言内解码准确率在所有测试脑区均显著高于基线水平,而跨语言解码准确率在无意识状态或部分意识状态下始终处于随机水平^[70],这表明提高解码准确率需要更多的注意力资源。再者,刺激模态的生态效度同样关键。有学者发现,与词云范式相比,图片范式呈现刺激下的语言解码准确率更高^[61]。这可能是因为图像表达语义信息比文字描述更直观具体,大脑对图片的加工深度也更高。大量研究证据一致表明,从刺激输入到模型构建,多模态信息的整合均展现出显著优势。在刺激层面,相比单一模态刺激,多模态刺激(如视听结合)能引发更丰富的神经活动,从而获得更高的解码准确率^[30]。在模型层面,融合视觉、听觉等不同来源的神经信号特征,可进一步提升解

码性能^[9,31]。

第五,不同刺激类型引发的神经响应具有显著的脑区特异性,需针对性选择信号源区域以提升解码效果。例如,对于语音刺激解码,感觉运动皮层在语音生成中显示出强烈的神经激活^[10],将发音运动学特征作为解码脑信号声学特征的中间步骤则有明显的解码优势^[9],这是由于语音产生与发音器官的控制密切相关,感觉运动皮层对解码语音脑信号尤为重要。在解码图像和视频信息时,需要利用高级视觉皮层的大脑体素信息,因为此类刺激主要激活高级视觉皮层。Nishida等^[34]在对电视广告视频片段激发的大脑活动进行解码时,发现信息体素主要位于V5区和颞中区等高级视觉区域。同时,Huang等^[73]也发现高级视觉皮层大脑信号的解码效果优于低级视觉皮层。

4 大脑语言解码研究的挑战与未来展望

大脑语言解码领域仍面临若干关键挑战。首先,当前研究的数据规模普遍较小,多数研究仅限于特定词汇集、句子类型或封闭语料中的语篇。其次,现有的解码器性能相对有限,训练出的解码器在处理超出初始训练集的内容时表现不佳,在更广泛的语义空间和多样化的语言环境中,解码器的表现仍然不尽如人意,这必然也限制了其在实际应用中的潜力。鉴于此,当前的大脑语言解码准确率仍有待提高,产学研成果转化仍有较大提升空间。未来可以从以下4个方向推进。

第一,多模态协同与多任务优化。视听结合运动的多模态实验范式有助于验证梭状回面孔区与视觉词形区的层级特征映射机制,Huang等^[50]提出的视觉语言解码模型(visual language decoding model, VLDM)通过并行解码图像类别、语义标签及文本描述,例如同时解码风景画的图像类别、落日鸟儿湖水云彩等语义标签,以及“落霞与孤鹜齐飞,秋水共长天一色”的文本描述,实现任务间的信息互补,有助于捕获大脑信号中的隐藏信息。

第二,语言神经机制的深入探索。计算模型的改进固然重要,但对大脑语言加工机制本身的深入理解或许才是解码效果突破的前提^[74-75]。针对截瘫患者的研究发现,缘上回可解码内在语言(语法)与外在语言(发音)的神经表征差异,未来需构建实时内在语言脑机接口系统,进一步验证缘上回区单神经元和神经元群体语义选择性的临床应用潜力^[76]。

第三,模型与算法革新。随着技术的进步,深度神经网络正逐步取代传统机器学习模型,成为推动脑机接口领域发展的核心技术。特别是随着深度学习、Transformer、预训练语言模型^[77]及GPT-4、DeepSeek等大语言模型的发展,开发适应多种语言和跨被试者的通用解码器,已成为技术创新的关键^[74]。多头交叉注意力机制的引入可提升全脑交互特征提取效率,进而提高句子解码与生成的准确率,在算法层面为通用型脑机接口提供支持,从而降低脑机接口的模型设计与训练成本。从智能科学视角看,多模态机器学习的发展为脑信号、视觉信息和语言表征的联合建模提供了重要方法支撑^[78]。此外,大规模神经数据库(如NeuroLang)的建立整合了多语系、多模态数据资源,促进了自监督预训练框架的发展,这类数据库使模型能够在更大规模的数据集上训练,将会增强对不同语言和个体差异的适应能力。

第四,临床转化与生态化应用。未来的研究亟须从受控的实验室环境向真实、自然的交互场景延伸。新一代系统需在解码精度、响应时延与生态效度之间实现协同优化,提升在日常场景中的稳定性与可用性,为失语症患者、听障人群等提供实用的辅助交流工具。例如,对手语与书写等非口语表达系统的研究应结合感知与执行双通道范式,解析动作规划与语义编码之间的神经耦合机制^[79]。这一方向的推进有望深化对多模态语言生成神经机制的理解,也为构建更自然的人机交互系统提供依据。

5 结束语

作为未来新赛道,语言脑机接口是前瞻性、战略性新兴产业和新质生产力的典型代表。大脑语言解码研究在过去十余年间取得了显著进展,研究者已能够从神经活动中解析词汇、句法乃至语篇层级的语言信息。但随着解码对象由低层语言单位扩展到更复杂的语义结构,现有研究在解码准确率保持和动态表征建模方面仍面临明显挑战。未来研究需在多模态融合、语言神经机制探索、算法改进与临床转化等方面持续推进,期望为更多失语患者提供长期稳定的语言恢复解决方案,进一步拓宽脑机接口在神经疾病诊疗领域的边界。

参考文献:

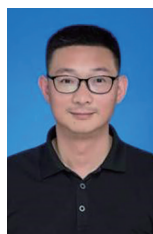
- [1] Mitchell T M, Shinkareva S V, Carlson A, et al. Predicting human brain activity associated with the meanings of nouns[J]. *Science*, 2008, 320(5880): 1191-1195.
- [2] Tang J, LeBel A, Jain S, et al. Semantic reconstruction of continuous language from non-invasive brain recordings[J]. *Nature Neuroscience*, 2023, 26(5): 858-866.
- [3] Correia J, Formisano E, Valente G, et al. Brain-based translation: fMRI decoding of spoken words in bilinguals reveals language-independent semantic representations in anterior temporal lobe[J]. *The Journal of Neuroscience*, 2014, 34(1): 332-338.
- [4] Xu M, Li D, Li P. Brain decoding in multiple languages: Can cross-language brain decoding work? [J]. *Brain and Language*, 2021, 215: 104922.
- [5] Hu Y, Mohsenzadeh Y. Neural processing of naturalistic audiovisual events in space and time[J]. *Communications Biology*, 2025, 8: 110.
- [6] Goodfellow I, Bengio Y, Courville A. *Deep learning*[M]. Cambridge: MIT Press, 2016.
- [7] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. *Nature*, 2015, 521(7553): 436-444.
- [8] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]// *Advances in Neural Information Processing Systems*. Red Hook: Curran Associates, 2017: 5998-6008.
- [9] Anumanchipalli G K, Chartier J, Chang E F. Speech synthesis from neural decoding of spoken sentences[J]. *Nature*, 2019, 568(7753): 493-498.
- [10] Ramsey N F, Salari E, Aarnoutse E J, et al. Decoding spoken phonemes from sensorimotor cortex with high-density ECoG grids[J]. *NeuroImage*, 2018, 180: 301-311.
- [11] Pereira F, Lou B, Pritchett B, et al. Toward a universal decoder of linguistic meaning from brain activation[J]. *Nature Communications*, 2018, 9: 963.
- [12] Dash D, Ferrari P, Wang J. Decoding imagined and spoken phrases from non-invasive neural (MEG) signals[J]. *Frontiers in Neuroscience*, 2020, 14: 290.
- [13] Huang W, Yan H M, Cheng K W, et al. A dual-channel language decoding from brain activity with progressive transfer training[J]. *Human Brain Mapping*, 2021, 42(15): 5089-5100.
- [14] Holm E, Marraffini G F G, Slezak D F, et al. Hierarchical temporal correspondence between brain activity and different deep neural network architectures[C]// *Advances in Neural Information Processing Systems*. [S.l.:s.n.], 2025: 1-8.
- [15] 曾小芹. 基于Python的中文结巴分词技术实现[J]. *信息与电脑(理论版)*, 2019, 31(18): 38-39, 42.
Zeng X Q. Technology implementation of Chinese jieba segmentation based on Python[J]. *China Computer & Communication*, 2019, 31(18): 38-39, 42.
- [16] Che W X, Feng Y L, Qin L B, et al. N-LTP: an open-source neural language technology platform for Chinese[C]// *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. Stroudsburg: ACL, 2021: 42-49.
- [17] Angrick M, Luo S Y, Rabbani Q, et al. Online speech synthesis using a chronically implanted brain-computer interface in an individual with ALS[J]. *Scientific Reports*, 2024, 14: 9617.
- [18] Vilela M, Hochberg L R. Applications of brain-computer interfaces to the control of robotic and prosthetic arms[M]// *Brain-Computer Interfaces*. Amsterdam: Elsevier, 2020: 87-99.
- [19] Ciabattini L, Ferracuti F, Freddi A, et al. Human-in-the-loop approach to safe navigation of a smart wheelchair via brain computer interface[M]// *Ambient Assisted Living*. Cham: Springer, 2021: 197-209.
- [20] Willett F R, Avansino D T, Hochberg L R, et al. High-performance brain-to-

[1] Mitchell T M, Shinkareva S V, Carlson A, et al. Predicting human

- text communication via handwriting[J]. *Nature*, 2021, 593(7858): 249-254.
- [21] Qian Y K, Liu C J, Yu P X, et al. Real-time decoding of full-spectrum Chinese using brain-computer interface[J]. *Science Advances*, 2025, 11(45): eadz9968.
- [22] Redwan S M, Uddin M P, Sharif M I, et al. A network theory investigation into the altered resting state functional connectivity in attention-deficit hyperactivity disorder[PP]. V1. arXiv (2022-11-23) [2026-03-10]. arXiv:2212.02402.
- [23] Li D W, Luo X S, Guo J L, et al. Information-based multivariate decoding reveals imprecise neural encoding in children with attention deficit hyperactivity disorder during visual selective attention[J]. *Human Brain Mapping*, 2023, 44(3): 937-947.
- [24] Xu C, Zou J J, He F P, et al. Neural tracking of sound rhythms correlates with diagnosis, severity, and prognosis of disorders of consciousness[J]. *Frontiers in Neuroscience*, 2021, 15: 646543.
- [25] Rahman N, Khan D M, Masroor K, et al. Advances in brain-computer interface for decoding speech imagery from EEG signals: a systematic review[J]. *Cognitive Neurodynamics*, 2024, 18(6): 3565-3583.
- [26] 陈霏, 潘昌杰. 基于发音想象的脑机接口的研究综述[J]. *信号处理*, 2020, 36(6): 816-830.
Chen F, Pan C J. A review on brain-computer interface technologies for speech imagery[J]. *Journal of Signal Processing*, 2020, 36(6): 816-830.
- [27] 杨淑淇, 谭颖. 深度学习在解码大脑语音信息中的应用[J]. *西南民族大学学报(自然科学版)*, 2025, 51(3): 315-322.
Yang S Q, Tan Y. Deep learning in decoding brain speech information[J]. *Journal of Southwest University for Nationalities (Natural Science Edition)*, 2025, 51(3): 315-322.
- [28] Liu Y, Zhao Z H, Xu M P, et al. Decoding and synthesizing tonal language speech from brain activity[J]. *Science Advances*, 2023, 9(23): eadh0478.
- [29] Lu J F, Li Y N, Zhao Z H, et al. Neural control of lexical tone production in human laryngeal motor cortex[J]. *Nature Communications*, 2023, 14: 6917.
- [30] Zhang X Y, Li H, Chen F. EEG-based classification of imaginary mandarin tones[C]//Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). Piscataway: IEEE Press, 2020: 3889-3892.
- [31] Chen X P, Wang R, Khalilian-Gourtani A, et al. A neural speech decoding framework leveraging deep learning and speech synthesis[J]. *Nature Machine Intelligence*, 2024, 6(4): 467-480.
- [32] Littlejohn K T, Cho C J, Liu J R, et al. A streaming brain-to-voice neuroprosthesis to restore naturalistic communication[J]. *Nature Neuroscience*, 2025, 28(4): 902-912.
- [33] Fukuma R, Yanagisawa T, Nishimoto S, et al. Decoding visual stimulus in semantic space from electrocorticography signals[C]//Proceedings of the 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Piscataway: IEEE Press, 2018: 102-104.
- [34] Nishida S, Nishimoto S. Decoding naturalistic experiences from human brain activity via distributed representations of words[J]. *NeuroImage*, 2018, 180: 232-242.
- [35] Lin Y, Hsieh P J. Neural decoding of speech with semantic-based classification[J]. *Cortex*, 2022, 154: 231-240.
- [36] Zhang M F, Lévy J, D'Ascoli S, et al. From thought to action: how a hierarchy of neural dynamics supports language production[PP]. V2. arXiv (2025-02-18) [2026-03-10]. arXiv:2502.07429.
- [37] 邹宇童, 李波, 徐晓翠. 面向分布式语义空间的fMRI解码研究[J]. *计算机技术与发展*, 2026: 2026.0076.
Zou Y T, Li B, Xu X C. fMRI decoding research for distributed semantic spaces[J]. *Computer Technology and Development*, 2026: 2026.0076.
- [38] Buchweitz A, Shinkareva S V, Mason R A, et al. Identifying bilingual semantic neural representations across languages[J]. *Brain and Language*, 2012, 120(3): 282-289.
- [39] Correia J M, Jansma B, Hausfeld L, et al. EEG decoding of spoken words in bilingual listeners: from words to language invariant semantic-conceptual representations[J]. *Frontiers in Psychology*, 2015, 6: 71.
- [40] Du C D, Fu K C, Li J P, et al. Decoding visual neural representations by multimodal learning of brain-visual-linguistic features[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(9): 10760-10777.
- [41] 任子琪. 基于深度生成模型的视觉神经信息解码与重构[D]. 西安: 西安电子科技大学, 2024.
Ren Z Q. Deep generative model-based neural decoding and visual reconstruction[D]. Xi'an: Xidian University, 2024.
- [42] Bruera A, Poesio M. Electroencephalography searchlight decoding reveals person- and place-specific responses for semantic category and familiarity[J]. *Journal of Cognitive Neuroscience*, 2025, 37(1): 135-154.
- [43] Wang Z, Ji H. Open vocabulary electroencephalography-to-text decoding and zero-shot sentiment classification[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022, 36(5): 5350-5358.
- [44] 刘晓龙, 杨帮华, 甘安安, 等. 基于CAM-Net模型的中文词语言语想象脑电解码研究[J]. *生物医学工程学杂志*, 2025, 42(3): 473-479.
Liu X L, Yang B H, Gan A A, et al. Study on speech imagery electroencephalography decoding of Chinese words based on the CAM-Net model[J]. *Journal of Biomedical Engineering*, 2025, 42(3): 473-479.
- [45] Jia Z H, Wang H B, Shen Y Z, et al. Magnetoencephalography (MEG) based non-invasive Chinese speech decoding[J]. *Journal of Neural Engineering*, 2025, 22(6): 066014.
- [46] Zhao J Y, Gao R M, Brennan J R. Decoding the neural dynamics of headed syntactic structure building[J]. *The Journal of Neuroscience*, 2025, 45(17): e2126242025.
- [47] Matsuo E, Kobayashi I, Nishimoto S, et al. Generating natural language descriptions for semantic representations of human brain activity[C]//Proceedings of the ACL 2016 Student Research Workshop. Stroudsburg: ACL, 2016: 22-29.
- [48] Yang Y, Wang J, Bailer C, et al. Commonalities and differences in the neural representations of English, Portuguese, and Mandarin sentences: When knowledge of the brain-language mappings for two languages is better than one[J]. *Brain and Language*, 2017, 175: 77-85.
- [49] Li C F, Liu Y, Li J L, et al. Decoding bilingual EEG signals with complex semantics using adaptive graph attention convolutional network[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2024, 32: 249-258.
- [50] Huang W, Yang P F, Tang Y, et al. From sight to insight: a multi-task approach with the visual language decoding model[J]. *Information Fusion*, 2024, 112: 102573.
- [51] Zhang D H, Wang Z J, Qian Y K, et al. A brain-to-text framework for decoding natural tonal sentences[J]. *Cell Reports*, 2024, 43(11): 114924.
- [52] 马磊, 崔文浩, 杨汶汶, 等. 非侵入性连续中文语言语义解码与重建[J]. *数据采集与处理*, 2025, 40(3): 616-636.
Ma L, Cui W H, Yang W W, et al. Non-invasive continuous Chinese language semantic decoding and reconstruction[J]. *Journal of Data Acquisition and Processing*, 2025, 40(3): 616-636.
- [53] Wang B, Xu X R, Zhang L X, et al. Semantic reconstruction of continuous language from MEG signals[PP]. V1. arXiv (2023-09-14) [2026-03-10]. arXiv:2309.07701.
- [54] Lewis M, Liu Y H, Goyal N, et al. BART: denoising sequence-to-sequence pre-training for natural language generation, translation, and

- comprehension[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2020: 7871-7880.
- [55] Wang J Q, Song Z X, Ma Z Y, et al. Enhancing EEG-to-text decoding through transferable representations from pre-trained contrastive EEG-text masked autoencoder[C]//Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Stroudsburg: ACL, 2024: 7278-7292.
- [56] 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.
- [57] Lévy J, Zhang M F, Pinet S, et al. Brain-to-text decoding: a non-invasive approach via typing[PP]. V1. arXiv (2025-02-18)[2026-03-10]. arXiv:2502.17480.
- [58] Dehghani M, Boghrati R, Man K, et al. Decoding the neural representation of story meanings across languages[J]. *Human Brain Mapping*, 2017, 38(12): 6096-6106.
- [59] Hakala T, Lindh-Knuutila T, Hultén A, et al. Subword representations successfully decode brain responses to morphologically complex written words[J]. *Neurobiology of Language*, 2024, 5(4): 844-863.
- [60] Brysbaert M, Mandera P, Keuleers E. The word frequency effect in word processing: an updated review[J]. *Current Directions in Psychological Science*, 2018, 27(1): 45-50.
- [61] Wang S N, Zhang J J, Wang H Y, et al. Fine-grained neural decoding with distributed word representations[J]. *Information Sciences*, 2020, 507: 256-272.
- [62] Bi Y C. Dual coding of knowledge in the human brain[J]. *Trends in Cognitive Sciences*, 2021, 25(10): 883-895.
- [63] Huang W, Tang Y, Wang S Z, et al. Unraveling the differential efficiency of dorsal and ventral pathways in visual semantic decoding[J]. *International Journal of Neural Systems*, 2025, 35(3): 2550009.
- [64] Polczyńska M M, Bookheimer S Y. General principles governing the amount of neuroanatomical overlap between languages in bilinguals[J]. *Neuroscience & Biobehavioral Reviews*, 2021, 130: 1-14.
- [65] Liu X Y, Qu J, Li H L, et al. Similar activation patterns in the bilateral dorsal inferior frontal gyrus for monolingual and bilingual contexts in second language production[J]. *Neuropsychologia*, 2021, 156: 107857.
- [66] Zinszer B D, Anderson A J, Kang O, et al. Semantic structural alignment of neural representational spaces enables translation between English and Chinese words[J]. *Journal of Cognitive Neuroscience*, 2016, 28(11): 1749-1759.
- [67] Kim S Y, Qi T, Feng X X, et al. How does language distance between L1 and L2 affect the L2 brain network? An fMRI study of Korean-Chinese-English trilinguals[J]. *NeuroImage*, 2016, 129: 25-39.
- [68] Liberto G M D, Nie J P, Yeaton J, et al. Neural representation of linguistic feature hierarchy reflects second-language proficiency[J]. *NeuroImage*, 2021, 227: 117586.
- [69] Zhao Y H, Huang W, Cao S Q, et al. Revisiting the conceptual representation of Chinese-English Bilinguals: Evidence from AI-based brain language decoding[J]. *Journal of Neurolinguistics*, 2026, 79: 101341.
- [70] Sheikh U A, Carreiras M, Soto D. Decoding the meaning of unconsciously processed words using fMRI-based MVPA[J]. *NeuroImage*, 2019, 191: 430-440.
- [71] Sheikh U A, Carreiras M, Soto D. Neurocognitive mechanisms supporting the generalization of concepts across languages[J]. *Neuropsychologia*, 2021, 153: 107740.
- [72] Guo Z Z, Xu L S, Tan W J, et al. Impact of generation rate of speech imagery on neural activity and BCI decoding performance: a fNIRS study[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2025, 33: 1180-1190.
- [73] Huang W, Yan H M, Cheng K W, et al. A neural decoding algorithm that generates language from visual activity evoked by natural images[J]. *Neural Networks*, 2021, 144: 90-100.
- [74] Zhao Y H, Chen Y, Cheng K W, et al. Artificial intelligence based multimodal language decoding from brain activity: a review[J]. *Brain Research Bulletin*, 2023, 201: 110713.
- [75] Keller T A. The neural bases of sentence comprehension: a fMRI examination of syntactic and lexical processing[J]. *Cerebral Cortex*, 2001, 11(3): 223-237.
- [76] Wandelt S K, Bjånes D A, Pejsa K, et al. Representation of internal speech by single neurons in human supramarginal gyrus[J]. *Nature Human Behaviour*, 2024, 8(6): 1136-1149.
- [77] Devlin J, Chang M W, Lee K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1. Kerrville: Association for Computational Linguistics, 2019: 4171-4186.
- [78] Baltrušaitis T, Ahuja C, Morency L P. Multimodal machine learning: a survey and taxonomy[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(2): 423-443.
- [79] Wang P P, Zhou Y Y, Li Z N, et al. Neural decoding of Chinese sign language with machine learning for brain-computer interfaces[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2021, 29: 2721-2732.

[作者简介]



程凯文（1978—），男，博士，四川外国语大学语言智能学院副教授，主要研究方向为认知神经科学、心理语言学、双语和语言智能。



杜圆圆（2000—），女，四川外国语大学语言智能学院硕士生，主要研究方向为神经语言学、自然语言处理、语言智能。



赵宇浩（1999—），男，广东外语外贸大学英语语言文化学院博士生，主要研究方向为认知语言学、语言智能、认知神经科学。