Chinese Traditional Culture 国学, 2026, 14(5), 1387-1394
Published Online October 2026 in Hans. https://www.hanspub.org/journal/cnc
https://doi.org/10.12677/cnc.2026.145195

Hans汉斯

大语言模型与语言学研究：演进脉络、应用图景与反思

宋金歌[1]，刘庆岭[1]，蔡丛聪[2]，宋晓冉[3]

[1]新疆师范大学国际文化交流学院，新疆 乌鲁木齐
[2]乌鲁木齐市第124中学，新疆 乌鲁木齐
[3]新疆师范大学数学科学学院，新疆 乌鲁木齐

收稿日期：2026年9月23日；录用日期：2026年10月16日；发布日期：2026年10月28日

摘　要

大语言模型的发展深刻重塑语言学研究范式。立足于语言学学科本位，本文梳理近十五年大语言模型由词嵌入、Transformer架构到生成式爆发的演进脉络，从语言本体与应用层面考察其在语义、句法、语用、语言教育、机器翻译、语料库语言学中的应用图景。研究表明，现有交叉研究多侧重技术测评，普遍将模型视作单纯工具，忽视二者双向互动。当前研究存在模型本身语言学缺陷突出、重工具轻理论、传统理论解释力遭遇挑战、方法论困惑与语言治理风险并存等问题。据此，本文提出推动理论双向融合、推进汉语及低资源语言本土化、构建人机协同研究范式、完善语言伦理治理的优化路径。语言学应当主动适配数据驱动范式，在技术与人文交叉地带实现理论与方法的创新发展。

关键词

大语言模型，语言学研究，范式反思，理论挑战，语言治理

Large Language Models and Linguistic Research: Evolutionary Trajectory, Application Landscape, and Reflections

Jinge Song[1], Qingling Liu[1], Congcong Cai[2], Xiaoran Song[3]

[1]School of International Cultural Exchange, Xinjiang Normal University, Urumqi Xinjiang
[2]Urumqi No.124 Middle School, Urumqi Xinjiang
[3]School of Mathematical Sciences, Xinjiang Normal University, Urumqi Xinjiang

Received: September 23, 2026; accepted: October 16, 2026; published: October 28, 2026

文章引用: 宋金歌, 刘庆岭, 蔡丛聪, 宋晓冉. 大语言模型与语言学研究: 演进脉络、应用图景与反思[J]. 国学, 2026, 14(5): 1387-1394. DOI: 10.12677/cnc.2026.145195

Abstract

The development of large language models has profoundly reshaped the research paradigms in linguistics. Grounded in the disciplinary framework of linguistics, this paper traces the evolutionary trajectory of large language models over the past fifteen years—spanning from word embeddings and the Transformer architecture to their generative surge—and examines their applications across the domains of linguistic ontology and practice, including semantics, syntax, pragmatics, language education, machine translation, and corpus linguistics. The study reveals that much of the current interdisciplinary research focuses primarily on technical evaluation, often treating language models as mere tools while overlooking the bidirectional interaction between models and linguistic theory. Current research faces several challenges: prominent linguistic limitations within the models themselves, an overemphasis on tools at the expense of theory, the declining explanatory power of traditional linguistic theories, and the coexistence of methodological uncertainties alongside risks associated with language governance. In response, this paper proposes an optimal pathway to foster bidirectional integration of theory and practice, advance the localization of research approaches for Chinese and low-resource languages, establish a human-machine collaborative research paradigm, and enhance language ethics governance. Linguistics should proactively adapt to data-driven paradigms and drive innovative theoretical and methodological developments at the intersection of technology and the humanities.

Keywords

Large Language Models, Linguistic Research, Paradigm Reflection, Theoretical Challenges, Language Governance

Copyright © 2026 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

1. 引言

1.1. 研究背景

语言是人类核心的交际媒介，技术迭代不断重塑语言的存在形态，也持续拓展语言学的研究边界。从文字、印刷术到广播电视、互联网，每一次媒介变革都推动语言载体、语体样态发生变迁，也促使语言学不断更新研究工具与分析视角。人工智能自 1956 年被正式提出，历经七十年发展，大语言模型的出现成为新的发展节点。大语言模型将人类语言转化为可运算的数据，实现语言的建模、压缩与生成，语言由此不再局限于人际交流，更成为人机、机机交互的重要媒介。

大语言模型依托语境嵌入形成三类语言学相关能力：向量驱动的语义相关性表征、基于概率的语言序列预测，以及依靠困惑度完成的语言质量计量。但模型依靠统计学习完成语言输出，并不等同于真正理解语言本身。这种全新的语言处理范式，既为语言学提供大规模量化分析的技术手段，也对传统语言学的认识论基础、研究范式形成冲击，引发学界对语言本质、语言习得机制的重新讨论。

当前大量研究多出自计算机与自然语言处理领域，侧重模型技术性能的评测，而立足于语言学学科本位的系统性梳理与批判性反思仍相对匮乏。不少研究简单将大语言模型视作语言学的辅助工具，忽视二者之间双向互动的内在关联：大语言模型既可以服务语言学研究，其涌现出的语言行为，同样可以成

为检验、反思语言学理论的实验对象。基于这一现实，有必要系统厘清大语言模型与语言学之间的复杂关系。

1.2. 大语言模型在语言学研究方面面临的主要问题

大语言模型的迅猛发展为语言学研究提供新工具的同时，也引发了一系列亟待回应的学术问题[1]。其一，模型存在固有的语言学局限，它仅模仿人类语言行为，不具备真正的语言理解能力。从规则建模转向概率化表征，不仅带来语义推理、常识判断的缺陷，也对经典语言学的认识论构成挑战。其二，研究取向重工具应用，轻理论对话，多数成果仅把模型当作分析手段，缺少语言学理论层面的对话与反思，亟须树立大数据时代下跨学科的大语言学观念。其三，传统语言学理论遭遇解释力危机，生成语法等经典理论既难以解释大模型的语言表现，也难以从模型行为中获得理论启发。其四，范式转型衍生方法论困惑与语言治理风险，研究从规则驱动转向数据驱动，大语言模型实验结果能否等同于人类语言认知尚存疑问；同时还面临生成内容失真、数据版权、低资源语言被技术忽视等现实困境。

1.3. 本文的主要贡献

第一，梳理近十五年大语言模型从词嵌入奠基到生成式爆发的演进脉络。立足语言学学科视角，系统考察其在语义、句法、语用、语料库、语言教育等领域的应用图景，弥补现有综述偏重技术介绍的不足。第二，揭示大语言模型与语言学理论的双向互动关系，不只关注模型的工具效用，更探讨其带给语言学的理论启示，呼应数据驱动范式下语言学的学科定位问题。第三，梳理当前交叉研究的争议与短板，从理论融合、汉语本土化实践、人机协同方法论、语言治理四个维度提出优化路径，为后续相关研究提供参照。

2. 近 15 年大语言模型发展演进历程

大语言模型的发展大致经历了三个紧密衔接的阶段，其演进脉络清晰地反映了大语言模型从技术奠基走向能力跃升的过程。

2.1. 词嵌入奠基期(2011~2016, Word-2-Vec)

2013 年，Mikolov 团队提出 Word-2-Vec 模型，将词语映射为连续向量，实现词语向量化[2]。此阶段核心在于使机器具备基本的语义理解能力。以 Word-2-Vec 为代表的词嵌入技术将词语转换为稠密的数值向量，使计算机得以捕捉词与词之间的语义关联，为后续神经网络语言模型奠定了基础。此外 CBOW (Continuous Bag of Words)和 Skip-gram 模型使用神经网络成功地建模了词与上下文之间的关系，前者用上下文预测目标词，后者则用目标词预测上下文。

与此同时，循环神经网络及其变体 LSTM 被广泛用于序列数据处理，但在长文本建模和并行计算效率方面仍存在明显瓶颈。直到 2017 年，预训练的 Transformer 架构被提出并逐渐成为主流，其自注意力机制不仅有效解决了长距离依赖问题，更支持并行训练，大幅提升了模型效率。

2.2. 预训练双向编码器时代(2017~2019, Transformer/BERT)

2017 年，Transformer 架构的提出成为大语言模型发展史上的里程碑[3]。其自注意力机制有效突破了传统递归神经网络在长距离依赖建模和并行计算方面的瓶颈，为后续大规模语言模型的涌现奠定了统一的架构基础。以此为契机，预训练与微调相结合的范式迅速成为主流。2018 年，GPT-1 与 BERT 相继问世，二者在海量无标注文本上进行预训练，再针对下游任务微调，效果显著优于传统方法。在技术路线上，GPT 系列采用 Transformer 解码器，以自回归方式进行单向语言建模，天然擅长文本生成；而 BERT 则利用 Transformer 编码器，通过掩码语言模型实现双向上下文理解，在自然语言理解类任务上表现卓越。

2019 年，GPT-2 以 15 亿参数量展现了强大的零样本学习能力，展示了规模扩展的巨大潜力[4]。

然而，大模型的快速发展也引发了学界的深层反思与激烈争议。2021 年，Bender 等学者提出“随机鹦鹉论”[5]，指出大模型仅通过统计规律模仿语言的表面形式，并未真正获得语义理解和逻辑推理能力。长期以来以乔姆斯基理论为代表的规则导向语言观，即语言能力根植于先天的普遍语法，在大语言模型凭借海量数据训练所涌现出的惊人语言生成能力面前，其解释能力受到前所未有的质疑。

计算语言学领域由此出现了统计模型与规则系统之间的路线之争，语言学研究从规则驱动逐步转向数据驱动。面对这一范式转移，语言学家被迫重新审视自身的角色定位：或回归语言本体理论的深耕，或在规则与数据的结合路径中寻求新的突破。

2.3. 生成式大模型爆发期(2020~2026, ChatGPT/国产大模型)

2020 年，GPT-3 以 1750 亿参数量震惊学界，其核心贡献在于验证了“规模定律”，即参数、数据和算力的同步提升能够在少样本乃至零样本场景下显著增强模型的泛化能力，大语言模型由此从学术研究走向大规模产业应用。2022 年 11 月，ChatGPT 的横空出世则成为这一阶段的标志性转折点，该模型基于 GPT-3.5，借助基于人类反馈的强化学习技术实现了模型输出与人类意图的有效对齐，极大地提升了对话的自然度与实用性，在全球范围内掀起生成式人工智能的热潮。笔者以“大语言模型”为关键词进行检索，结果显示以大语言模型为主题的相关研究自 2022 年以来急剧飙升。

纵观大语言模型的发展历程，从词向量表征到 Transformer 架构革新，从预训练范式确立到大规模生成式模型落地应用，每一次技术跃升都建立在前期积累之上。值得注意的是，这一演进过程始终伴随着语言学理论的参与和反思，从 Word-2-Vec 对分布式语义学原理的实践，到 Transformer 对句法结构建模能力的突破，再到 ChatGPT 对语言能力是否必须预设先天模块的拷问，语言学理论既是模型发展的参照系，也是模型冲击的承受者。

3. 大语言模型在语言学研究中的应用图景

传统语言学研究依赖田野调查和有限规模的语料分析，而大语言模型的出现使研究者得以在大规模自然语言数据上实现对语言结构、使用与认知的系统性量化分析。本文主要从语言本体及语言应用两个层面进行分析。

3.1. 语言本体层面

3.1.1. 语义研究

语义研究关注语言的意义表征与理解机制，大语言模型为此提供了从静态到动态、从分布式到语境化的分析范式。在词义表征方面，早期 Word-2-Vec 等静态词向量模型实现了词义的分布式数值表达，使得语义相似度可计算化。而基于 Transformer 架构的动态词向量则进一步使词义表征随具体语境动态变化，为语境义研究提供了直接的技术支持。在语义消歧方面，大语言模型凭借深层上下文建模能力，在多义词义项判别、指称关系分辨等任务中表现优异。隐喻识别是另一活跃方向，大语言模型能够捕捉源域与目标域之间的概念映射关系，在隐喻表达识别与隐喻类型分类中取得较传统方法更优的效果。

现有语义研究的贡献在于将语义分析从理论推演推向可计算、可验证的实证层面，分布式语义学原理在大语言模型中的成功应用为语义研究提供了新的方法论基础。然而多数研究聚焦于大语言模型在特定语义任务上的表现评估，而缺乏对评估结果的理论解释[6]。此外，对汉语特殊语义现象，如量词搭配、主观性副词、离合词用法等研究尚处于起步阶段，现有评估框架多基于英语设计，是否适用于汉语的类型学特征有待检验。最后，大语言模型在语义理解和常识推理方面表现优秀，但经典的“图灵测试”难以识别欺骗和回避等作弊手段。这表明现有语义能力评估方法本身存在局限性，亟需建立更严谨的语言

学评估框架。

3.1.2. 句法研究

句法研究领域呈现语言学理论与大语言模型双向互动格局。一方面，理论语言学为考察模型句法能力提供概念参照框架。胡旭辉(2024)立足生成语言学的语言生物官能理论构建修正版语言模型，探究在普遍语法作用下，句法特征与功能结构实现句法系统内化的内在机制[7]。另一方面，依存语法具备较强普适性与计算适配优势，本就广泛应用于自然语言处理，大语言模型进一步挖掘了依存语法在计算分析层面的应用潜力。

从工具价值看，大语言模型为句法探测、依存解析、语法合法性判别开辟新研究路径。Hewitt (2019)等学者借助探针任务，检验模型内部是否建构层级化句法树表征，测试其对主谓一致、时态呼应、嵌套结构等句法规则的掌握水平[8]。已有研究证实，大语言模型参数空间存储大量句法信息，句法感知能力伴随模型规模扩张不断提升。同时，模型能够输出语句语法可接受度量化评分，为低频合法边缘句式、语法异常结构提供量化研判依据[9]。

但当前研究仍存在显著局限。第一，探针任务的设计高度依托研究者的句法理论认知，不同研究的实验方案差异显著，造成研究结论难以横向比对。第二，尽管学界普遍认可模型内部编码句法信息，但模型究竟习得了抽象句法规则，抑或只是捕捉表层词语统计共现关系，这一核心争议尚未形成共识。第三，汉语句法专题研究较为薄弱，“把”字句、“被”字句、兼语句等构式在模型内部的表征逻辑与处理机制，仍缺少系统性考察。

3.1.3. 语用与话语研究

语用学关注语言在具体语境中的使用与理解，人工智能与大语言模型的技术浪潮，为聚焦语言运用能力探索的修辞学、语用学等相关学科构建了异于传统研究范式的新维度。

在语用推理方面，研究者评估大语言模型对 Grice 合作原则的遵循程度、对会话隐含的识别能力以及对预设触发语的敏感程度。研究表明，新一代大语言模型在语用能力方面已有显著提升，但在处理反语和幽默等语用现象时，依然依赖于字面意义的直接理解，对复杂语境的推理能力有限。在言语行为方面，大语言模型可对请求、承诺、威胁、道歉等不同语力类型进行自动分类与强度量化[10]。在篇章层面，大语言模型在指代消解、话语标记识别、篇章修辞结构分析等任务中表现良好。耿雯雯、李冰芸(2026)从网络语用学视角对大语言模型的礼貌策略进行了专门研究，大语言模型在礼貌评价方面呈现出较高一致性，主要依托话轮转换规则、礼貌原则及语境要素对交际行为进行综合判断，但在话语组织与解释路径上体现出一定的语篇模式差异。在言语回应中，大语言模型以积极礼貌与消极礼貌策略为主，综合运用间接表达、情态动词及模糊限制语等手段，有助于构建话语的积极情感取向，但对复杂语境中的非言语信息关注相对不足[11]。与此同时，学界对大语言模型的语用能力保持审慎态度。此外，杨敏、王亚文(2023)等学者从语言学的形式主义、功能主义和批评话语研究三个理论层次，系统探讨了 ChatGPT 在理解自然语言意义方面所暴露的不足[12]。

现有语用研究领域多集中于对反语、幽默等特定语用现象的评估，缺乏对语用能力的系统性、多维度评估框架。大语言模型在跨文化语用差异方面的表现研究极为薄弱，不同文化背景下的礼貌策略、面子需求等语用规范在模型中的表征方式尚不清楚。耿雯雯、李冰芸(2026)虽从网络语用学视角进行了研究，但整体而言，大语言模型语用研究的理论深度不足，多数研究停留于模型能否完成某任务的描述层面。

3.2. 语言应用层面

笔者以大语言模型为关键词进行检索，以中国语言文字为学科限定，进行可视化分析得出不同学者

的研究方向和研究趋势，其中研究方向最多的为：二语习得、语言教育、机器翻译、语料库建设等，故选取这四个方面进行分析。

3.2.1. 大语言模型与语言教育

在二语习得层面，大语言模型依托对话生成能力，可模拟可理解性输入、提供即时纠错反馈，支持个性化会话练习、写作修改，有效降低学习者开口焦虑，拓展课堂之外的练习场景[13]。从理论对话来看，二语习得强调人类学习依托社会互动、内隐与外显认知协同，而模型仅依靠统计拟合完成语言输出，并不具备学习者真实的习得认知过程。现有实证研究多聚焦于写作、口语辅助练习，证实模型反馈能够提升文本句法准确度与词汇丰富度，但多数实验为短期小样本，缺少长期追踪证据。当前研究的突出局限在于：大语言模型完全依赖数据驱动的统计学习，而人类能将情感与现实世界背景融入学习(具身认知)，部分研究直接将大模型作为“虚拟二语学习者”或理想教学代理，缺乏人类的社会交互与认知灵活性[14]。此外，学习者过度依赖模型生成答案，也存在弱化自主输出能力的现实风险。

在语言测试与评估层面，大语言模型推动自动测评从浅层句法打分走向多维质性评价，在写作、口语的内容立意、语用得体性、篇章逻辑等维度展现出较强潜力，人机评分相关度已经接近人工评判水平[15]。马瑞祯(2026)则指出区别于传统自动作文评分系统，大模型可输出解释性评语，不再局限于简单分数输出。但该领域尚存显著不足：第一，模型评分稳定性易受提示词、输出随机性干扰，面对汉语学习者的偏误类型识别偶有偏差，幻觉问题会造成误判；第二，现有评估基准大多源自英语，适配 HSK 等汉语二语考试的专用评测框架尚不成熟；第三，不少研究只验证模型和人工分数的相关性，缺少从语言学视角对评分逻辑的剖析，模型内部评判标准不可解释，难以满足标准化考试的信效度要求[16]。

3.2.2. 大语言模型与机器翻译

机器翻译[17]已进入大语言模型驱动的新发展阶段。大语言模型凭借超长上下文窗口、交互式提示调控与多模态处理能力，整合全篇逻辑与专业背景信息，消解单句翻译中的词汇歧义。在多语言处理方面，大语言模型在跨语言信息检索、文本分类及语义相似度计算中展现出较强的零样本迁移能力。翻译作为融合多类非语言要素的高级人类智能活动，当前大语言模型解读复杂人文内容的能力依旧有限，对蕴含历史内涵与文化意象的内容常做浅层化处理。此外，低资源语种翻译准确率偏低且易生成文化刻板印象，译文知识产权、算法偏见与数据泄露等伦理问题亦不容忽视。

现有研究仍面临以下问题：其一，对翻译质量的评估多集中于流畅度和忠实度等表层指标，缺乏对文化传递、语体适切性等深层维度的评估；其二，浅层化处理问题虽然被识别，但调整训练数据、优化模型架构等问题尚缺乏系统研究；其三，低资源语言的翻译问题虽然被提及，但真正针对低资源语言的大语言模型翻译研究极为有限。

3.2.3. 大语言模型与语料库语言学

20 世纪后半叶兴起的语料库语言学，核心是依托计算机对大规模真实语料开展检索、统计与分析，以此归纳语言形式的分布与使用规律。大语言模型的介入，深刻改变了语料库语言学的技术路径与方法论基础。学者围绕翻译研究，论证了大语言模型在语义识别、深度分析、多模态处理方面的应用价值。

技术应用层面，大语言模型可完成词性、句法成分、语义角色乃至语用信息的多层自动标注，有效压缩人工成本，提升标注结果一致性，同时能够对海量实时文本完成智能筛选、分类与结构化存储，促使语料库由静态封闭模式转向动态开放。但新技术也带来范式层面的深层挑战。部分研究直接将大模型生成文本用作语料开展分析，混淆了人类真实语言事实与算法生成内容的边界。国内还面临语料库数量、质量不足的现实困境，难以适配大语言模型的训练要求[18]。

总体来看，大语言模型在语料库方面的应用极大提升了语料标注的效率与一致性，为大规模、多层语料库的建设提供了技术可能。但现有研究仍存在明显短板：一是将大模型用作语料分析工具时，缺少充分的方法论反思；二是现有实证多集中于翻译方向，大模型在通用语料库语言学中的相关成果较为零散，尚未形成完整体系；三是本土语料资源短板同时制约大模型迭代与语料库语言学的创新发展，是亟待破解的现实瓶颈[19]。

4. 现存问题及优化路径

4.1. 现存问题

尽管大语言模型在语言学研究的多个领域取得了显著进展，但仍面临一系列挑战，这些挑战限制了它在实际应用中的有效性和可持续发展性。

第一，模型存在固有的语言学短板，其本质依靠统计拟合模仿人类语言，并不具备真正的语言理解能力。模型易产生幻觉，在汉语特殊构式、复杂语用推理、常识逻辑推演上存在缺陷，黑箱特性致使难以厘清模型是习得抽象语言规则，抑或是仅捕捉表层统计共现关系。

第二，研究存在重工具、轻理论的偏向。多数研究仅将大模型视作分析工具，缺少与语言学理论的双向对话与反思。研究多聚焦任务性能测评，缺少从语言学学科本位出发的理论阐释；同时存在学科壁垒，理论语言学与计算语言学对话不足，研究成果较为零散，尚未形成统一分析框架。

第三，汉语本土化适配不足。现有评估范式多移植自英语研究，缺少面向汉语类型特征的专门设计；高质量汉语、方言及少数民族语言语料资源供给不足，制约本土大模型迭代与汉语语言学实证研究。

第四，方法论困惑与语言治理风险并存。部分研究混淆模型生成文本与真实自然语料边界；大语言模型实验结果能否等同于人类语言认知仍存争议。此外，训练数据版权、低资源语言被边缘化、算法偏见、生成内容失真等伦理问题，尚未形成语言学视角下的系统性治理方案。

4.2. 优化路径

针对以上问题，本文从理论、本土实践、研究方法、语言治理四方面，为未来大语言模型在语言学研究方面提出优化路径。

首先，推动语言学理论与大模型双向融合。一方面把构式语法、依存语法等理论转化为可计算约束与评测指标；另一方面将大模型作为实验平台检验语言学假说，搭建理论语言学与计算语言学常态化交流机制。

其次，强化汉语与低资源语言本土化建设。扩充汉语、方言、少数民族语言高质量语料库，立足汉语无形态变化的类型特征开展模型适配，摆脱对英语研究范式的简单照搬。

再次，构建人机协同的语言学研究范式。建立模型初加工和语言学家人工核验的研究流程，拒绝直接将模型输出等价于语言事实；构建适配汉语现象的语言学评测指标，重视研究的可重复性与可对比性。

最后，完善语言伦理与治理体系。推动语言学者参与大语言模型语言规范制定，围绕数据偏见检测、语言多样性保护开展实证研究；明确大模型在语言教育、语料分析等场景的使用边界，构建全链条语言伦理审查机制。

5. 总结

本文梳理了大语言模型由词嵌入奠基、Transformer 革新到生成式模型爆发的发展脉络，其实现了从语言表征到语言生成的技术跃迁。大语言模型为语义、句法、语用、语言教育等语言学分支提供全新研究工具，同时也对传统语言学理论形成反思与冲击。

综合文献分析可见，现有交叉研究仍存在明显局限：多侧重技术应用而轻视理论对话，缺少语言学

本位的批判性考察；汉语本土化适配研究薄弱，传统语言学理论解释力遭遇范式挑战；相关伦理与语言治理议题缺少语言学者充分参与。

面向未来，语言学应当打破学科壁垒，推动语言学理论与大模型双向融合，推进汉语及低资源语言本土化建设，建立人机协同的研究范式，完善语言伦理治理框架。语言学无需固守传统研究边界，应立足人文视角把握技术变革机遇，在技术与人文的交叉场域开拓学科新方向。

参考文献

[1] 袁毓林. ChatGPT 和语言学的梦醒时分: 语言大模型对语言学的挑战与启示[M]. 北京: 商务印书馆, 2026.

[2] Mikolov, T., Chen, K., Corrado, G., *et al*. (2013) Efficient Estimation of Word Representations in Vector Space. 1*st International Conference on Learning Representations*, Scottsdale, 2-4 May 2013, 1-12.

[3] Vaswani, A., Shazeer, N., Parmar, N., *et al*. (2017) Attention Is All You Need. 2017 *Advances in Neural Information Processing Systems* 30 (*NeurIPS* 2017), Long Beach, 4-9 December 2017, 5998-6008.

[4] Radford, A., Wu, J., Child, R., *et al*. (2019) Language Models Are Unsupervised Multitask Learners. OpenAI.

[5] Bender, E.M., Gebru, T., McMillan-Major, A. and Shmitchell, S. (2021) On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *Proceedings of the* 2021 *ACM Conference on Fairness*, *Accountability*, *and Transparency*, Online, 3-10 March 2021, 610-623. https://doi.org/10.1145/3442188.3445922

[6] 鲍贵. 当代语言学关键词·大语言模型[J]. 外语研究, 2026, 43(2): 8-17+90.

[7] 胡旭辉. 语言官能与外部世界的联结——基于生成语言学的语言模型初探[J]. 外语教学与研究, 2024, 56(3): 452-462+481.

[8] Hewitt, J. and Manning, C.D. (2019) A Structural Probe for Finding Syntax in Word Representations. *Proceedings of the* 2019 *Conference of the North American Chapter of the Association for Computational Linguistics*: *Human Language Technologies*, *Volume* 1 (*Long and Short Papers*), Minneapolis, June 2019, 4129-4138. https://doi.org/10.18653/v1/n19-1419

[9] 刘海涛. 依存语法和大语言模型[J]. 语言文字应用, 2025(4): 2-15.

[10] 毛延生, 王一航. 生成式预训练语言模型汉语语用能力的实证研究——以 ChatGPT 为例[J]. 外国语言文学, 2024, 41(6): 40-55+130+137.

[11] 耿雯雯, 李冰芸. 大语言模型的礼貌策略研究: 基于网络语用学视角[J]. 外语学刊, 2026(3): 37-47.

[12] 杨敏, 王亚文. ChatGPT 的“理解”与“意义”: 论其生成语言背后的形式、功能与立场[J]. 中国外语, 2023, 20(3): 24-32.

[13] 李书弘, 杨连瑞. 大语言模型何以赋能二语习得?——基于国内外文献的范围述评[J]. 中国外语, 2025, 22(3): 23-31.

[14] 王建勤. 广西师范大学——学术讲座实录[EB/OL]. https://wxy.gxu.edu.cn/info/1103/4471.htm, 2025-05-22.

[15] 冷静, 曹凌琳, 王思瑜, 等. 面向深层语言特征的智能评估系统构建与应用验证[J]. 远程教育杂志, 2025, 43(6): 91-100.

[16] 马瑞祯, 丰迪, 徐娟. 面向大语言模型的中文二语作文评分量规构建[J]. 西安外国语大学学报, 2026, 34(1): 38-44.

[17] 胡开宝, 李晓倩. 大语言模型背景下翻译研究的发展: 问题与前景[J]. 中国翻译, 2023, 44(6): 64-73+192.

[18] 李兴腾, 冯锋, 黄鹂强. 突破人工智能大模型的“数据瓶颈”——构建国家级语料库运营平台的思考[J]. 中国科学院院刊, 2025, 40(3): 522-529.

[19] 陈强, 邢窈窈. 面向生成式人工智能的语料库建设: 现状、挑战与未来策略[J]. 社会科学, 2025(6): 177-192.