大语言模型在儿童语言障碍领域的应用

徐津津

池州学院　安徽池州

【摘要】儿童语言障碍主要包括发展性语言障碍（DLD）、特殊型语言障碍(SLI)、自闭症谱系障碍（ASD）相关的语言障碍等类型。语言障碍是儿童语言发展中的常见问题，具体表现为语言理解或表达能力的显著落后，交流过程中出现障碍。根据世界卫生组织统计，语言发育迟缓儿童发生率约为 6—8%。我国语言障碍儿童占儿童总数的比例估计达到 10%-20%。儿童语言障碍已经成为一个急需解决的问题。传统干预方法依赖人工评估与一对一训练，效率低且难以实现个性化支持，导致治疗效果不理想。近年来，以 DeepSeek-R1 为代表的大语言模型技术的快速发展为语言障碍儿童的诊断、干预及教育支持提供了全新路径。本文结合文献研究与应用案例，探讨大语言模型在语言障碍领域的应用现状、技术优势及未来挑战，旨在更好地服务语言障碍儿童。

【关键词】大语言模型；儿童；语言障碍

【基金项目】池州学院校级人文重点项目“汉语特殊型语言障碍儿童与高功能自闭症儿童量词习得对比研究”（项目编号：CZ2020RWZ10）

【收稿日期】2025 年 3 月 21 日　　【出刊日期】2025 年 4 月 22 日　　【DOI】10.12208/j.ssr.20250164

Application of large language models in the field of children’s language disorders

Jinjin Xu

Chizhou University, Chizhou, Anhui

【**Abstract**】Children’s language disorders mainly include developmental language disorders (DLD), specific language disorders (SLI), and language disorders related to autism spectrum disorder (ASD), etc. Language disorders are a common problem in children’s language development, manifested as a significant lag in language comprehension or expression abilities, as well as obstacles in communication processes. According to the World Health Organization, about 6% to 8% of children worldwide have language disorders. It is estimated that the proportion of children with language disorders in China accounts for 10% to 20% of the total number of children. Children’s language disorders have become an urgent issue that needs to be addressed. Traditional intervention methods rely on manual assessment and one-on-one training, which are inefficient and difficult to achieve personalized support, resulting in unsatisfactory treatment outcomes. In recent years, the rapid development of large language modeling technology represented by DeepSeek-R1 has provided a new path for the diagnosis, intervention, and educational support of children with language disorders. This article combines literature review and application cases to explore the current application, technological advantages, and future challenges of large language models in the field of language disorders, aiming to better serve children with language disorders.

【**Keywords**】Large Language Model (LLM); Children; Language disorders

1 引言

儿童是国家的希望，儿童的健康关系到民族昌盛和国家强盛。党中央和国务院高度重视儿童健康工作，近年来连续出台《健康儿童行动提升计划（2021—2025 年）》、《中国儿童发展纲要（2021—2030 年）》和《关于推进儿童医疗卫生服务高质量发展的意见》等系列重要文件，着力构建起从生命起点到成长全周期的健康保障体系，推动我国儿童健康事业迈向高质量发展新阶段。儿童时期是语言能力发展的重要阶段，语言能力的初步发展不仅为日后的交流沟通奠定基础，还在很大程度上影响着个体在学业成就及职业发展等方面的长远表现。我国语言障碍儿童占儿童总数的比

例估计达到 10%-20%，庞大人口基数带来的社会影响巨大[1]。儿童语言障碍主要包括发展性语言障碍（DLD）、特殊型语言障碍(SLI)、自闭症谱系障碍（ASD）相关的语言障碍等类型。这些障碍影响患者的沟通能力，由于部分类型语言障碍儿童不会出现身体或者智力残疾，因此原发性语言障碍的认识相对落后，还可能对其心理健康和社交功能造成严重影响。

近年来，人工智能技术在我国快速发展，其中自然语言处理（NLP）领域基于深度学习框架（如 Transformer 架构）训练的大语言模型，已突破语义理解、文本生成及多模态分析的技术瓶颈，推动人机交互向认知智能阶段演进。特别是以 DeepSeek-R1 为代表的大模型不仅能够精准解析复杂语言结构，还可生成符合人类逻辑的创造性内容，并在医疗、教育等场景实现智能决策支持。2023 年 DeepMinds 团队发布全球第一个医疗健康大语言模型 Med-PaLM 及其升级版 Med-PaLM2。该模型通过联邦学习整合全球 150 万份电子病历，在美国医师执照考试（USMLE）的 MedQA 数据集上取得 92.6%的准确率。我国医疗行业大模型发展态势迅猛，目前国内已诞生超 200 个医疗大模型。涵盖自然语言处理、机器视觉等多项技术。如阿里巴巴的“阿里健康 AI”，科大讯飞的“智医助理”，华为研发的“数智本草大模型”以及 DeepSeek-R1 医疗定制版。随着社会经济的快速发展和社会对儿童语言能力需求的提升，儿童语言健康问题日益受到重视。

基于特有自然语言处理属性，本文将聚焦于探讨大语言模型在语言障碍儿童领域的运用，描述及分析其技术优势、未来挑战并提出相关建议。

2 大语言模型与语言障碍相关研究

2.1 语言障碍的成因与影响

语言障碍的类型多样，不同类型的语言障碍对患者的影响也各不相同，盛欐等（2022）指出语言障碍大致可以分为原发性语言障碍和伴随性语言障碍。由于原发性语言障碍主要表现在语言方面，又被称作特殊型语言障碍（SLI）或发展性语言障碍（DLD）。[2]。徐林荔等（2020）研究表明，汉语普通话特殊型语言障碍（SLI）儿童在量词习得方面发展迟缓，特别是形状量词的畴化和生成能力显著低于正常儿童[3]。周彦每(2023)研究表明，学龄前语言障碍儿童的病因复杂多样，特发性语言发育迟缓占比最大，而听力障碍和智力障碍患儿的语言发育迟缓程度较重且发现年龄较早[4]。陈炜等（2022）发现失语症则是一种严重的语言障碍，通常由脑部损伤引起，导致患者在语言表达、理解、阅读和书写等方面均存在困难。失语症患者由于脑部损伤导致语言能力受损，常常无法正常表达自己的思想和需求，这不仅影响他们的日常生活，还可能导致孤独感和抑郁症状。语言发展迟缓的儿童则可能在学习和社交方面面临困难，影响他们的整体发展[5]。

2.2 传统评估与康复方法的局限

石定栩与杨洋（2020）指出传统的语言康复方法依赖于人工治疗，虽然在一定程度上有效，但存在成本高、资源分配不均等问题[1]。罗薇等（2024）通过对全国儿童语言康复人员问卷调查发现传统语言康复方法的局限性还表现在治疗效率和个性化程度方面。人工治疗通常需要较长时间，且每次治疗的时间有限，难以满足所有患者的需求[6]。刘雪曼(2022)指出治疗的效果在很大程度上取决于治疗师的专业水平和经验，这可能导致不同患者之间的治疗效果差异较大。一些偏远地区或资源匮乏的地区可能缺乏专业的语言治疗师，导致患者无法及时获得有效的康复治疗。此外，每个患者的语言障碍类型和程度不同，需要个性化的治疗方案，而传统的治疗方法往往难以做到高度个性化，导致治疗效果不理想[7]。

2.3 医疗大语言模型相关研究

刘通等（2024）研究发现医疗健康大模型已应用于智能问诊、影像解读、智慧病案、康复管理等领域。从分诊前、诊中和诊后 3 个阶段分析医疗健康大模型的落地应用情况[8]。Gu（2021）提出，对于生物医学等具有大量未标记文本的领域，新的预训练语言模型可以比持续预训练通用领域语言模型取得更大的收益。利用大语言模型预测能力，为早期诊断和干预提供支持[9]。Xu（2024）利用大型语言模型 ChatGPT，提高对自闭症相关语言障碍的敏感性，并识别出独特的特征，从而克服传统诊断方法的不足，提高速度和精度[10]。闫温馨等（2025）探讨了大模型在基层医疗系统中的应用，提出应进一步深化大模型的应用，以辅助基层医生常见病诊疗，推动智能化健康教育和慢性病管理[11]。

3 大语言模型在语言康复服务中的应用

3.1 基于多模态高效语言评估能力

传统评估方式需要耗费专业人员大量时间和精力，大语言模型在评估儿童语言障碍方面展现出显著优势，基于自然语言处理（NLP）技术对儿童的语言表达进行分析，评估语言障碍儿童语法、词汇和语用能力。通过分析海量语言数据，建立客观评估模型，量化儿童的语言发展指标（如语速、语法错误率），减少传统评估的主观性。特别是利用机器学习算法对儿童的句子结构、

词汇使用和语言流畅性进行量化评估，帮助诊断特殊型语言障碍儿童。由于此类儿童没有智力与肢体残疾，非常容易与其他类型语言障碍混淆，大语言模型可以结合儿童语音、文本、面部表情和行为数据，能够高效地评估儿童的语言障碍情况，提升语言评估效率。

3.2 个性化康复方案与持续改进能力

大语言模型康复系统能够实时监测和分析儿童在康复训练中的语音、行为和环境数据，构建全面的语言康复模型。模型参数基于全球康复医生、儿童保健医生以及言语治疗师的丰富经验，确保为每位儿童量身定制精准的计划。系统通过人工智能算法实时监测语言儿童能力进步，识别改善程度和潜在问题，并提供个性化反馈报告。这些信息不仅支持医生或康复师调整策略，还能同步给家长，增强家庭参与度。随着儿童在训练中的进展，这些计划也会不断更新，以最大限度地提高训练效果，确保儿童的语言能力得到最优化。同时，言语治疗专业技术人员在此基础上，通过际观察以及家长的反馈进一步的微调和优化参数，使得大语言模型能够持续改进并准确地评估儿童的语言障碍情况。

3.3 提升家庭沟通与宣教能力

大语言模型在促进儿童情绪表达上具有广泛应用，虚拟情境生成技术创设沉浸式对话场景，引导儿童进行情感叙事。利用共情对话算法生成符合儿童认知特点的反馈语句，降低情绪表达的心理阻抗，不仅增强其语用能力（意图识别与情境化表达水平），更重要的是促进儿童向父母主动分享感受。由于语言干预的专业性要求与家长知识储备不足形成结构性矛盾，加之语言沟通障碍导致亲子互动质量下降，容易加剧家长心理压力。混合专家系统架构医疗健康大模型（如DeepSeek-R1 医疗版）可以通过情感计算模块开展共情对话缓解家长焦虑情绪，利用知识图谱宣传模块提供循证医学指导，使家长认知误区减少，积极配合干预工作，进而显著提升康复效果。

3.4 降低评估与康复成本，促进医疗资源均衡发展

通过建立标准化评估与康复系统，大语言模型的评估模块可提供 24 小时在线支持，实现优质医疗资源的规模化应用。通过手机 app 部署，家长可通过智能终端对语言障碍儿童完成评估并得到个性化支持。大语言模型能突破地域限制，缓解专业治疗师短缺问题，从而降低相关成本。

4 大语言模型在语言康复服务中面临的挑战

4.1 干预准确性

大语言模型可能输出不准确或非事实信息，这种情况被称为模型幻觉[7]。模型在推理过程中会产生逻辑上与事实信息相矛盾的输出，或者输出无法被证实或者证伪内容。在医疗场景中，医生的诊疗建议通常具有相当程度的信任属性，因为患者往往缺乏足够的医疗知识，只能选择相信医生的专业判断来继续采取诊疗措施。然而，当大语言模型错误地使用了知识或者使用了错误的知识，却依然顺畅地输出语言内容时，在缺乏信息检验和修正的情况下，患者可能会难以辨别信息的准确性或者潜在的危险性。这种情况下，可能会导致一系列严重的后果，包括误诊、不适当的治疗甚至有害的治疗结果。因此通过引入更多高质量的医疗数据进行训练，以增强模型对医疗知识的掌握，不断提高大语言模型能力为患者提供更加安全、可靠的医疗服务。

4.2 隐私与数据安全性

医疗健康大模型的训练离不开医学相关数据，而这种数据的获取和使用往往涉及复杂的伦理和法律问题。用于训练医疗健康大模型的数据中可能包含健康状况、家庭信息、临床信息等涉及患者隐私的敏感数据。这些数据的泄漏不仅会对患者的个人生活造成严重干扰，还可能引发医疗纠纷和法律问题，产生不可挽回的损失。为了应对这些挑战，研究者们正在积极探索各种隐私保护技术，通过差分隐私技术，可以在保证模型训练效果的同时，对数据进行有效的隐私保护。利用对抗性训练技术，使模型在面对各种复杂情况时更具鲁棒性，以及开发专门的验证机制，对模型生成的输出进行实时监测和校正。由于系统联网部署，网络数据安全管理显得尤为重要，制定并实施一套全面的数据安全策略，包括建立严格的数据访问权限控制和先进的加密机制，不断提升数据安全防护水平，从而确保医疗数据都在安全的网络环境中得到妥善处理与存储。

4.3 数据治理公平性

医疗领域大模型训练需要大量数据，这些数据可能包含偏见性样本，比如未纳入不同地域或者经济地位相关样本。不公正相关训练会导致模型诊断过程中出现误诊进而造成个人健康受损与医疗资源浪费。因此数据集应当尽可能地包容不同种族、性别、年龄和社会经济地位，从而避免模型在训练过程中吸收并放大某些特定群体的偏见。同时结合人工标注和算法识别，在模型训练前有效地减少数据集中的偏见，确保模型在后续的应用中更加公正和客观。在模型后入使用后，定期进行公平性评估，通过一系列的措施避免因数据偏见而导致的歧视和不公，使得大语言模型技术能够更好地服务于广泛的社会群体。

5 总结

随着跨学科的高度融合发展，大语言模型可以通过精准诊断、个性化治疗方案推荐等方式，提升儿童语言障碍诊断的准确性和效率。通过语言障碍儿童数据的深入分析和挖掘，大语言模型还能为政府儿童健康政策的制定提供科学依据。在保障患者福祉方面，医疗健康大模型能够通过提供更加个性化、精准的医疗服务，提升患者的治疗效果和生活质量，从而真正实现以患者为中心的服务理念。总之，跨学科合作不仅是推动大语言模型发展的关键手段，也是确保其能够持续、健康发展的保障。通过这种合作，大语言模型将在未来为儿童的健康事业带来更多的突破和福音。

参考文献

[1] 石定栩,杨洋.中国语言障碍与言语治疗——现状分析和发展思路[J].语言战略研究,2020(2):17-24.

[2] 盛欐,俞稼钰,郝颖.汉语发展性语言障碍的现状及建议[J].早期儿童发展,2022(4):73-83.

[3] 徐林荔,何晓炜,徐津津,孙蓝.汉语特殊型语言障碍儿童量词习得研究[J].外语教学,2020(4):63-68.

[4] 周彦每.发展性语言障碍儿童临床诊断:研究进路及损伤特征[J].河南师范大学学报(哲学社会科学版), 2023(1): 107-113.

[5] 陈炜,张莉,饶睿,赵韵韵,周林路.145 例学龄前言语语言障碍儿童病因及特征分析[J].听力学及言语疾病杂志, 2022 (2):202-205.

[6] 罗薇,史泱,张庆苏.国内儿童言语语言障碍治疗现状调查[J].中国听力语言康复科学杂志,2024(4): 397-401.

[7] 刘雪曼.语言发育迟缓与语言障碍的临床干预[J].中国儿童保健杂志,2022(8): 813-817+821.

[8] 刘通,黄成凤,张雯. 医疗健康大语言模型的发展、挑战与对齐[J]. 医学信息学杂志,2024 (09): 1-9+18.

[9] Gu Y, Tinn R, Cheng H, et al. Domain-specific Language Model Pretraining for Biomedical Natural Language Processing[J]. ACM Transactions on Computing for Healthcare,2021 (1):1-23.

[10] Xu X, Yao B, Dong Y, et al. Mental-LLM: Leveraging Large Language Models for Mental Health Prediction via Online Text Data[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2024 (1):1-32.

[11] 闫温馨,胡健,曾华堂.人工智能大语言模型在基层医疗卫生服务中的应用与挑战[J]. 中国全科医学, 2025(1): 1-6.

版权声明：©2025 作者与开放获取期刊研究中心（OAJRC）所有。本文章按照知识共享署名许可条款发表。
http://creativecommons.org/licenses/by/4.0/
OPEN ACCESS