

超声技术在语音生理研究中的应用综述

刘帅东¹, 吕士良^{2*}

¹西北民族大学语言与文化计算教育部重点实验室, 甘肃 兰州

²西北民族大学甘肃省民族语言文化智能信息处理重点实验室, 甘肃 兰州

收稿日期: 2026年5月29日; 录用日期: 2026年7月17日; 发布日期: 2026年7月29日

摘要

超声技术凭借自身优势, 从传统医学延伸至语音生理研究各领域, 成为该领域重要技术支撑。自20世纪70年代以来, 其应用不断拓展, 研究从发音器官、发音现象的基础研究, 逐步覆盖语言健康、语言教学、语言资源建设等方向, 形成多学科交叉的研究体系。本文运用文献计量、主题归纳等方法, 系统梳理了2000年以来约24年的国内外相关研究进展。研究发现, 国际研究已在言语治疗、二语习得、发音机制量化等方面形成了成熟研究范式, 国内研究在言语障碍诊疗、跨学科技术融合等层面仍有不足。伴随人工智能发展, 语音超声技术将向智能化、精准化发展, 以期更好地服务于未来语言事业的兴旺发展。

关键词

超声技术, 语音生理, 研究进展, 学科融合, 言语研究

A Review of the Application of Ultrasound Technology in the Study of Speech Physiology

Shuaidong Liu¹, Shiliang Lyu^{2*}

¹Key Laboratory of Linguistic and Cultural Computing, Ministry of Education, Northwest Minzu University, Lanzhou Gansu

²Key Laboratory of Minzu Languages and Cultures Intelligent Information Processing, Northwest Minzu University, Lanzhou Gansu

Received: May 29, 2026; accepted: July 17, 2026; published: July 29, 2026

Abstract

Thanks to its inherent advantages, ultrasound technology has expanded from traditional medicine

*通讯作者。

文章引用: 刘帅东, 吕士良. 超声技术在语音生理研究中的应用综述[J]. 现代语言学, 2026, 14(7): 1050-1060.
DOI: 10.12677/ml.2026.147738

into various fields of speech physiology research, becoming a key technical pillar in this area. Since the 1970s, its applications have continued to expand, with research progressing from basic studies of speech organs and phonetic phenomena to gradually encompassing areas such as speech health, language teaching and language resource development, thereby forming a multidisciplinary research framework. This paper employs methods such as bibliometrics and thematic analysis to systematically review the progress of relevant domestic and international research over approximately 24 years since 2000. The findings reveal that international research has established mature research paradigms in areas such as speech therapy, second language acquisition and the quantification of articulation mechanisms, whilst domestic research still has shortcomings in areas such as the diagnosis and treatment of speech disorders and the integration of interdisciplinary technologies. With the development of artificial intelligence, the voice ultrasound technology will evolve towards intelligence and precision, with the aim of better serving the thriving development of future language undertakings.

Keywords

Ultrasound Technology, Speech Physiology, Research Developments, Interdisciplinary Integration, Speech Research

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

超声技术针对语言本体的研究集中于发音器官和发音现象, 随后又扩展到了语言健康、语言战略等方面。语言健康是人类基本的健康素养, 超声技术在言语治疗方面的应用效果已经得到了证实且仍在持续发挥着作用。同时, 新时代语言事业需要努力建设语言资源以提供全面支撑, 也需要努力开发语言产品以提供高效服务。时至今日, 超声技术已在语言教学、方言和多语种研究中积累了一定数量的语言资源, 为部分语言文字工作的高效开展提供了有力保障。当下, 随着智能新技术的日益兴起、经济与社会的持续发展, 面对语言全面发展的新要求、超声技术在发音生理领域的相关研究将继续推动语言事业的蓬勃发展。

2. 超声技术在发音生理方面的研究现状

通过对 2000 年以来约 24 年的国内外文献整理, 采用 CNKI 中文数据库进行“超声”“生理”“语音”等关键词检索, 共获得 26 篇相关文献; 利用 SCI 核心数据库将“超声”“语言学”等关键词进行检索, 共获得 301 篇相关文献, 分别对两个数据库的搜索结果进行发文量分析, 得出了关于超声发音生理研究分别在 CNKI 与 SCI 的文献趋势图, 主题分布图和关键词网络化可视图。

通过文献整理分析可知(见图 1), 国际上的超声生理语音研究的成果则更为丰硕。从年发文量来看, 国外的研究起步早、数量多, 在 2016 年之前相对呈增长趋势, 单位年度的文献量最高达到了 30 篇左右, 该数据有力表明了利用超声技术研究发音生理等相关领域是当时的热点和重心; 在 2016 年之后, 研究的热度有所降温, 年发文量基本保持在 20 篇左右, 由此可知, 超声技术在发音生理有关方面的研究出现了瓶颈, 时代的变化要求研究者们发展新思路、发展新研究范式。在 VOSviewer 中进行关键词可视化分析时, 将关键词最小出现的次数设置为 17, 则在 1112 个关键词中有 21 个词达到了阈值。

图 1. SCI 文献趋势图

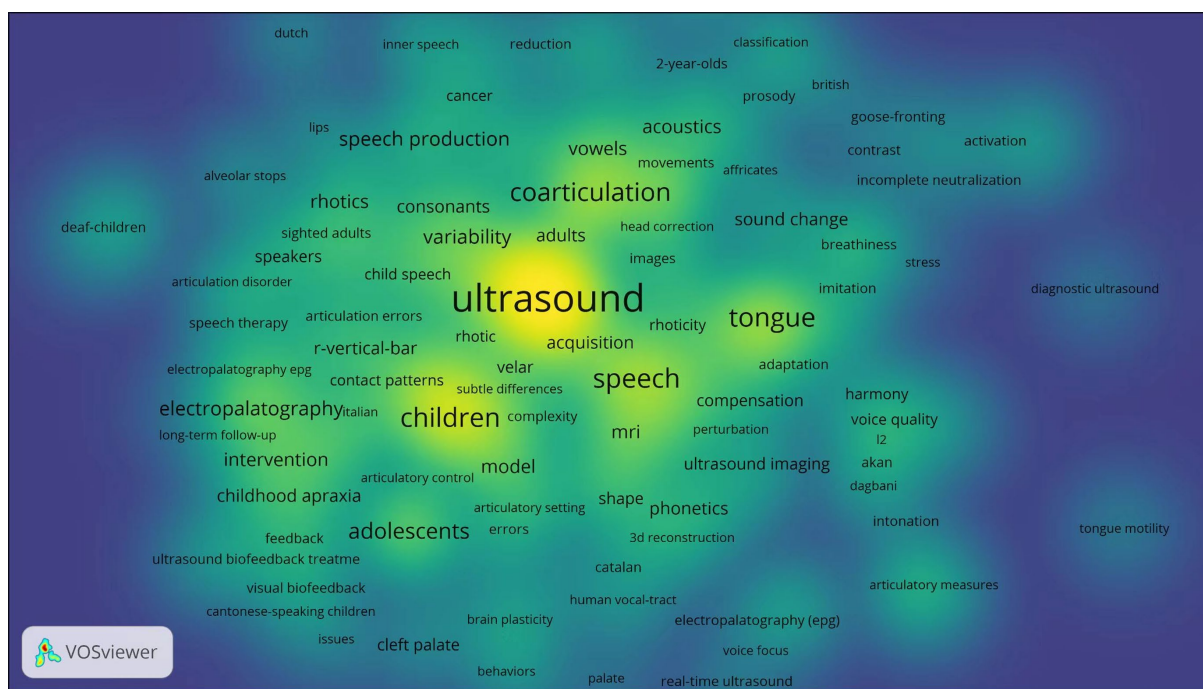


图 2. SCI 文献关键词网络化可视图

与国外研究相比,国内超声发音生理研究主要围绕多模态分析、语言数据库、发音器官探究、语言教学四大方向展开(见图3)。从文献分布来看,聚焦多模态技术与语言数据库的研究成果数量最多,可见学界普遍赞同超声语音研究在语言资源建设中的价值;发音器官、汉语教学、藏语方言等关键词则体现

出研究对象的多元化特征, 整体研究重心集中在发音生理、语言教学与语言数据库三大板块。然而, 国内现有的相关成果体量有限, 各类研究均有待进一步深化完善。相较于国际研究成果, 国内超声技术结合言语治疗的应用研究发展滞后, 相关方向尚未得到足够重视。进一步利用新兴智能技术补充完善语言应用领域的不足无疑是当下研究的重点任务与热点趋势, 语言应用领域中的言语治疗也将成为未来极具探索价值的潜在方向。

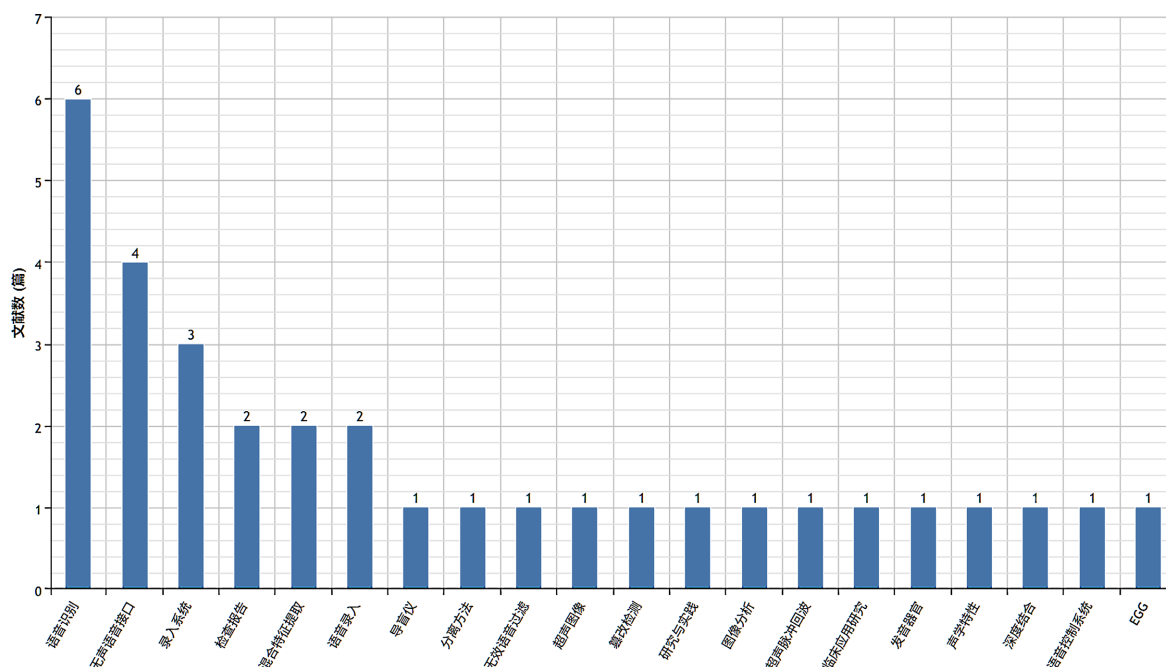


Figure 3. CNKI distribution chart of major themes in ultrasound physiology

图 3. CNKI 超声生理语音主要主题分布图

3. 超声技术在发音生理研究中的应用

通过对国内外超声技术在发音生理领域应用现状的梳理, 超声技术首要应用于发音器官和发音现象的解释研究, 为语言健康和语言战略提供了基础支撑。认识语音的本质是一个复杂的过程, 语言影响着人们日常的工作生活, 也关乎国家未来的发展, 故而涉及到语音的语言健康和语言战略(包括语言资源和语言文字工作)同样值得被关注, 二者需要超声技术的介入支持。简言之, 超声技术在发音生理方面的应用在今后依然具有广阔的发展空间。

3.1. 发音器官及发音现象

早在 20 世纪 60 年代的超声语音研究领域, Kelsey [1] 等人率先使用超声技术来观察生理发音情况, 并证实了超声技术具有能量易集中、方向性优和穿透力强等优点, 可贵之处在于对人体无害。在发音器官的研究方面, 舌体是主要观察对象, 因而正确地认识并利用超声舌像则是首要前提。Preston JL [2] 等人对超声图像的基本知识进行了介绍和说明, 指出超声在多个学科领域被高频地使用, 正确分析超声图像是开展研究的首要工作。文章主要从超声图像的解释、评估舌形的方法、对错误舌形的划分和引出适合的功能性的舌形这四个方面来规范超声技术的应用范式, 为专业研究人员提供了必要指导。Wommack BA [3] 指出: “超声的使用早已从医用扩展到了语音测试。” Stone, Sonies, Shawker, Weiss 和 Nadel (1982) 测试了五位正常成年人持续产生元音 /i/, /a/, /u/ 和 “rest” 位置。他们通过该方法以期能够实现舌体表面

的可视化。随后, Stone M [4]就此做了专门介绍。在开展舌体的超声波检测时, 研究者需将探头放置在发音人的下颌部位并根据研究需要调整探头方向, 当探头的方向与头部的前后方向平行时, 所得到的超声波图像为舌体的矢状轮廓切片; 当探头的方向与头部的前后方向垂直时, 则会得到舌体的冠状轮廓切片, 将舌体矢状轮廓和冠状轮廓切片结合起来就是舌体的三维形态。一般而言, 语言学领域的相关研究以观察舌体的矢状轮线在发音过程中的运动变化为主。

3.1.1. 发音器官——舌

在有关舌体的研究中, Fusheini Hudu [5]等人利用超声技术探讨了舌超声成像在记录非洲语言中的作用。其中着重关注了舌根和谐、元音收缩和软腭前移三个现象, 该研究帮助了非洲语言中舌根现象的直接映射理论的进一步完善。Vazquez-Alvarez Y [6]等人对于“低谷效应”, 即英语双唇塞音在对称 VCV 序列中通常出现舌头下垂的现象, 进行了有关其成因的探究。研究结果表明, 利用超声分析多位受试者大量舌面位移运动的数据是可行且实用的方法。Palo P [7]等人利用超声图像作为舌头发音测量的概念证明, 探究了不同阶段(从言语准备到言语发声)的发音过程数据为实现自动化识别舌体变化速率与其所反映出的像素差异轮廓等相关内容。Rivera Campos A [8]等人利用超声成像技术对舌后部在颤音产生过程中的构型进行了细致描述, 其研究成果对于分析舌体构型变化以获取目标有效音的发音策略是有价值的。在国内, 索昂代占[9]等人利用超声探究了藏语单辅音的语音现象, 根据舌位图像的声学特征, 发现了不同辅音间的区别特征和具体的音位信息。龙代吉草[10]等人利用超声技术分析了藏语安多方言的舌位运动过程, 与古藏语相比, 研究发现安多方言的元音系统已经产生了历时性变化, 此举对于研究藏语不同方言之间的差异与共性均有重要的理论意义和参考价值。

3.1.2. 发音器官——声喉

在有关发音器官中的声带和咽喉方面, Hertz CH [11]等人于 1970 年率先将超声用于观察声带运动, 记录了声带运动的清晰曲线。国内的刘劲荣[12]等人利用超声技术对汉语普通话基础元音的声带运动进行了细致的分析, 加强了发音音系学理论的解释能力。前文已知, Kelsey 等人描述了在言语产生过程中使用诊断性超声技术监测头部和颈部各种生理参数的技术。讨论了这些技术的基本原理以及直接应用于言语生理学研究的基本仪器系统, 包括咽壁深度的 A-扫描测量、气管的 B-扫描、移动侧咽壁的时间 - 运动显示和声带速度的多普勒监测。文章的结论是诊断超声可以成为语音研究的有用工具。Gugatschka, M [13]等人利用超声在大量受试人群中对喉部嗓音进行了评估分析, 进而表明了喉部发音的特殊性。

3.1.3. 发音现象——鼻音

对于鼻音研究, Carignan C [14]等人为解决鼻音化对元音声学独立的影响, 使用超声波和鼻音预测舌状物构型对鼻化元音的共振峰频率进行了分析探究, 得出了鼻化元音有逆时针链移现象的结论。Bouavichith DA [15]等人利用超声成像来确定马拉雅拉姆语具有七个对比鼻音发音部位(唇音、齿音、齿龈音、卷舌音、腭音、软腭音和腭化软腭鼻音/ɲ/)是否在实际中是动态稳定的, 通过一位马拉雅拉姆母语人士来分析舌与腭接触的位置和时间, 集中讨论了腭化软腭鼻音与腭化辅音的不同, 以及软腭鼻音和其他腭化辅音的区别之处。国内的马良[16]等人使用超声技术比较了中派和新派发音人鼻化元音与纯口元音的发音差异, 明确了不同年龄段发音人之间的变化以及普通话对上海话鼻化元音的影响, 研究结果显示, 中派发音人能够区分前鼻化元音和后鼻化元音; 而新派发音人则不能区分二者的不同。

3.1.4. 发音现象——协同发音

在有关超声发音现象的研究中, 协同发音的研究成果最为丰富, Krebs VL [17]等人利用超声技术来观察三位受试者舌根收缩的位置变化, 研究结果表明, 舌根参与了舌背前元音的协同发音。Saito [18]等

人利用超声技术研究了词形频率效应和遗留或预期共音的相互作用, 对德语动词中的元音[a]的协同发音进行了超声成像分析, 在实验中通过控制动词的尾缀和前面的代词来诱发共音效应。研究结果显示, 低频词的舌根更靠前, 舌体更低, 舌尖更高; 而高频词的舌形更圆, 舌根更缩回, 由此可知, 高频词具备更加优化的舌部运动。国内的宋清逸[19]等人利用超声波技术对拉萨藏语元音进行了发音舌位和元辅音间的协同发音现象的观察, 研究结果验证了元辅音协同发音的普遍规律, 总结了影响协同发音的因素。

4. 超声技术在语言资源研究中的应用

语言资源作为语言事业发展的基础工程, 对当下的语言研究提出了系统化、时代化的新要求。超声技术在发音生理方面的研究需要与时俱进, 通过对技术手段的适当调整和改进能够于今后对语音研究产生积极作用。Shawker T.H.与 Stone M [20]等人借助超声技术更多地了解了发音的生理活动, 以确定言语产生的正常主体间的变化, 并试图将这些变化与年龄、性别和语言环境联系起来, 此举有望在语音合成方面积累大量相关的数据库信息。Lawson E [21]等团队使用超声舌像语音语料库对苏格兰英语中后置词/r/的变化进行了社会语言学研究, 得出的结论是, 发音水平的不同变化与当地的社会分层息息相关。辨析发音中变异现象有时需要借助如超声的手段方能有助于探究其中的真实原委。Porras-Plata D [22]等人利用超声系统获取、分析和评估了舌部运动的发音信息, 提出了两组特征来表示发音信息, 并与仅使用声学信息训练的 SV 系统进行比较。研究结果表明, 所提出的特征包含高度区分性的信息, 并且与说话者身份相关; 此外, 这些特征可以通过在特征级别将此类信息与倒谱系数相结合来补充和改进现有系统。

武丹[23]设计了一个用于稳定超声探头和 EGG 配合使用的设备。在确保头部和超声探头可以自由灵活的移动的前提下, 解决了传统稳定设备重量大、挤压下颌等问题。从评价结果和多模态发音器官运动的分析来看, 该装置的性能良好, 获得的数据质量较高, 可以为语音研究提供优质数据。寇贇[24]利用超声波技术对舌位线进行了提取, 搭建了一个基于 MATLAB 平台开发的舌位参数提取及分析系统, 该系统界面的功能模块划分清晰, 实现了对舌位数据的统计、分析及绘图的代码功能, 为今后处理分析更多的舌位发音运动数据提供了有效办法。

此外, 部分研究者还阐述了超声技术存在的问题和完善措施, 该工作有助于语言资源的精细化和科学化发展, 如 Wommack BA [3]指出: “Stone, Sonies, Shawke 等人发现超声语音实验存在探头定位的变异性及无可见的硬结构在图像中作为参考舌位这两大问题”。Van Lieshout PH [25]等人发现声波的一个缺点是缺乏整体触发图像, 对于舌尖的反映成像可能变得不可见; 且超声波对于所拍摄图像的分析只能逐个处理。Stone M [26]等研究者试着用计算机中的网络-数字转换器系统为超声图像所提取的舌部测量数据提供相对恒定的参考, 该方法进一步弥补了超声语音分析的不足之处, 能够对可视化的超声舌位图像进行更加科学细致的描述。总体而言, 经过超声技术分析后的发音数据使语音中个别事实不清, 界限不明的现象或问题得到了科学地解决。

5. 超声技术在语言教学研究中的应用

语言教学通常被视为语言文字工作的主阵地。发音生理研究与语言文字工作密不可分, 让学习者(包括儿童和成年人的语言习得)熟练地掌握目标语言的发音原理一直是语言教学研究的热点。因此, 可视化的超声图像为提升学生学习发音的效率做出了贡献, 在语言教学过程中所积累的实际数据充实了语言资源, 助力语言事业的发展。

在国外的超声语言教学领域中, Gick B [27]等人讨论了第二语言习得和教学问题, 概括介绍了超声

技术进行发音训练等一系列教学活动的使用说明, 文中所列举的实际案例起到了很好的示范效果。Davidson L [28]等人利用超声成像的数据表明, 第二语言的学习者使用如补语的传统修复方法会对目标语言的习得产生影响, 尤其体现在学习者所必须获得的知识类型上。Tsui HM [29]等人设置了将英语作为第二语言的成年日本人关于 /l/ 和 /ɹ/ 的发音为实验样本, 利用超声波的视觉反馈进行对应的发音训练, 实验结果表明, 发音人训练前后的变化明显, 舌部超声波训练法有益于教学实际。Suzuki K [30]等人开发了一款面向第二语言学习群体以及语音学专业学生的英语发音软件。使用者可以在学习过程中看到超声波舌头的有关视频数据, 帮助提升学习者练习发音效率, 以便达到理想的听说水准。Mennnen I [31]等从事第二语言研究的学者们在原有语音规范的基础上使用超声技术进一步捕捉第二语言学习者在习得过程中的发音倾向, 如唇形圆展程度、嘴唇和舌头的张力、发声类型等语音配置, 为相关的语言教学工作提供了有力支持。Wilson I [32]等人进一步指出, 由于发音研究可以直接观察发音器官(如舌头, 下颚等), 因此在实际发音教学中, 使用超声波这样的直接视觉反馈方法, 对于语言学习者和对有兴趣于观察语言本体内部、跨语言交流及对比发音差异的研究人员而言无疑会起到巨大的帮助作用。Joanne C [33]等人将超声技术应用用于舌读训练, 为日后实现视觉生物反馈治疗的成功做出了一些尝试与探索。

在国内, 谭亚鑫[34]等人利用超声技术观察拼音方案中元音字母的实际舌位, 提出了针对性发音训练策略, 有助于进行普通话元音舌位的教学和训练。张文平[35]主要论述了汉语语音生理语料库在对外汉语教学中的应用情况, 该语料库的数据是由天津大学计算机科学学院自主研发的超声音频系统采集获得的, 研究结果表明: 基于超声音频系统构建的语料库在对外汉语教学实际中具有积极作用。魏远[36]指出: 基于超声和 MRI 图像融合的可视化发音过程可视化系统相比于传统的发音教学, 其显著优势在于能够为学习者提供实时的视觉反馈, 快速发现问题、提升学习发音的效率, 并在一定程度上帮助听障人士矫正发音。从长远来看, 国内超声技术在教学中的应用仍具有无限的发展潜力, 未来对于国家通用语的推广普及、区域方言和多语种的规范保护等工作产生重要影响, 以最终服务于语言发展的大局。

6. 超声技术在语言健康研究中的应用

国际上, 语言健康的超声研究主要聚焦于言语病理治疗方面, 发音患者主要分为青少年和儿童两大群体。对于不同年龄段和不同类型的发音患者而言, 超声技术具有良好的帮助与治疗效果, 在客观上也为相关的医学实践提供了参考价值。Barberena LD [37]等人证实了超声技术在言语病理治疗和听力科学中的积极作用, 其旨在提供更准确的医疗诊断以实现相关治疗方案的可行性与科学性。Saigusa H [38]团队证明了彩色多普勒超声成像在观察正常受试者和构音障碍者等其他肌病患者在执行音节重复任务时舌体运动的有效性。Zharkova N [39]等人利用超声舌像作为定量分析腭裂患者舌头功能的潜在方式, 该措施可用于对比患者治疗前后的舌头功能, 对于更高效地选择或改进相应的治疗方案极具参照价值。Bernhardt B [40]等人讨论了使用超声技术进行言语康复的效用, 表明超声波治疗技术已经可以治疗英语元音、流音等发音障碍, 治疗的反馈效力非常实用。Cavin M [41]等人使用二维动态式的超声波对患者发北美英语 /r/ 音进行矫正治疗。其结果有效提高了发音人发目标音的准确性。

6.1. 青少年语言健康

对于青少年的发音治疗, Bernhardt B [42]等人进行了一项对四名患有不同程度的神经性听力损失的青少年的治疗研究。治疗目标包括对比 /s/ 与 /z/、流音 /l/ 与 /r/ 等, 结果均显示出使用电腭分析和超声波辅助治疗后, 四名青少年的症状得到了有效改善。Bryfonski L [43]等人使用超声比较了视觉和口头纠正反馈对第二语言发音发展的影响, 为英语青少年提供了视觉反馈, 帮助他们掌握了 /r/ 的延迟。

6.2. 儿童语言健康

与青少年相比, 儿童言语的健康情况则更受关注。Ribeiro MS [44]等人讨论了超声舌像在自动检测语音发音错误方面的贡献。其研究价值体现为帮助诊断和治疗患有言语声音障碍的儿童尽早脱离困境, 并将由此产生的影响降到最小。Zharkova N [45]等人立足于处理幼儿的舌形信息, 分析对比了儿童的超声数据, 表明了语音舌形差异能够量化, 提升了对舌运动微小变化的敏感程度, 超声波在儿童言语治疗中容易用作诊断工具。Wiethan F [46]等人回顾了语音技术的发展脉络, 文中提到的一项研究表明超声图像比较了儿童和成人舌头的共振峰特性, 该研究有力地支持了儿童的协同发音程度明显高于成人的观点。文中另一项关于超声检查在修复北美英语因素/r/方面的潜在效用的调查结果显示出超声辅助治疗舌头的功效有望进一步提升。最后, 该篇论文指出在语音领域进行其他研究是必要性的工作。McKeever L [47]等人使用超声图像来评估儿童的言语运动, 实验选取了八名患有自闭症谱系障碍的儿童和九名正常发育儿童, 并进行了传统的速率、准确性和一致性测量以及慢速和快速发音的超声舌形分析。测量结果表明自闭症儿童的舌位运动变化较大, 超声舌像是一种获得言语运动控制的有效方法。Spencer C [48]等人为了探究患有残留言语障碍(RSSD)的儿童在言语生成神经功能方面是否存在差异, 通过选取 16 名患有 RSSD 的儿童在语言治疗计划之前(时间 1)和之后(时间 2)完成了 fMRI(功能性磁共振成像)实验, 该计划使用超声视觉反馈来纠正/r/发音错误。实验结果表明, 残留言语障碍(RSSD)和超声视觉反馈治疗可能与言语产生过程中的言语运动和视觉联想过程的神经差异有关。

相较于国外, 我国依托超声技术的言语治疗相关研究成果较为匮乏, 学科发展受限主要源于三方面核心问题。其一, 学科认知与领域划分固化。言语治疗长期归属于医疗范畴, 与语言本体研究脱节, 新兴的临床语言学在我国尚处于起步阶段, 整体研究水平落后于国际前沿[49]。同时, 语言康复作为交叉领域, 未形成成熟的学科体系与人才培养机制, 多学科协同不足, 致使语言理论与临床治疗实践割裂, 形成“两张皮”的现象[50], 国内语言学研究极少参与言语障碍诊疗工作, 进一步加剧领域壁垒。其二, 跨学科研究理念落地受阻, 单一研究范式制约发展。交叉研究是语言学与医学发展的必然趋势, 培养复合型语言康复人才也契合“健康中国”战略, 但国内医学交叉领域普遍存在学科壁垒难以打破、融合理念难以落地的问题[51]。国内相关研究多局限单一思维模式, 缺乏系统化、多学科的研究思路, 严重阻碍了超声言语治疗领域的创新发展。其三, 研究受试群体选取失衡, 临床研究供需错位。国内语言学研究样本多以正常语言使用者为主[52], 针对构音障碍人群的专项临床研究体量有限, 虽已有部分患儿干预研究[53], 但整体成果稀缺。同时, 国内言语治疗专业人才缺口大、儿童言语障碍需求攀升, 临床服务需求与学术研究供给差距显著, 间接制约领域成果产出。此外, 超声语音技术结合深度学习技术也是未来跨学科融合研究的重要发展方向之一。如 Al Ani S [54]等人强调了深度学习技术在超声舌像评估和诊断语言障碍方面的独特优势。其研究结果表明, 超声语音技术能够与深度学习技术配合使用, 二者将共同服务于言语障碍的相关治疗。

7. 总结与讨论

本文详细梳理了 2000~2024 年国内外超声语音相关的文献, 较为全面地梳理了超声技术在语音生理领域的发展脉络与主要应用场景。结果显示, 超声技术现已成为语音生理研究不可或缺的工具之一, 被广泛运用于发音机制解析、语言教学、言语障碍诊疗、语言资源建设等多个方向。目前该领域正逐步构建起语言学、医学、计算机科学等多学科交叉的研究框架, 在阐释不同人群、不同语种的发音生理规律、优化可视化发音教学模式、辅助言语障碍临床干预、构建多模态语言数据库等方面, 均已取得了较为突出的研究进展。

相较而言, 国内外超声语音生理研究的发展水平仍存在较为显著的差距。国际领域研究起步更早, 理论范式与研究方法更为成熟, 研究重心偏向言语治疗、二语习得等应用型方向, 普遍采用大样本长期追踪的实证研究方案。部分研究成果已逐步走向落地试用, 陆续出现超声言语评估软件(Articulate Assistant Advanced)、辅助诊断软件(Ultra-CATS 与 Sonospeech Ultra-CATS)等原型产品与试点应用, 但放眼全球, 尚未形成全覆盖、标准化的成果转化体系; 规模化、普及化的商用产品明显不足。国内超声语音相关研究虽起步相对滞后, 但近年来发展速度较快, 在多模态数据融合、汉语方言与多语种发音生理探究、语言资源数据库搭建等方向逐步形成自身特色与优势。但整体而言, 国内研究仍以基础理论探索为主, 应用类研究发展薄弱, 仍存在科研成果向实际应用转化周期长、落地难度大, 部分研究设计规范性有待提升, 跨领域合作广度与深度不足等现实问题。

现如今, 超声语音领域仍面临诸多共性难题: 全球范围内尚未建立统一的超声语音数据采集规范, 超声影像智能识别与分析的精度仍有提升空间, 多模态数据同步融合技术存在瓶颈, 研究样本覆盖面与代表性不足, 技术产业化整体发展速度缓慢。针对以上问题, 本文建议未来可依托新兴人工智能与深度学习技术推动超声语音研究向智能化转型; 研发并拓展便携、低成本的专用超声设备, 降低技术使用门槛、推动行业普及; 进一步拓宽研究范畴, 针对特殊言语群体、方言及多语种开展多元化探索, 拓展远程言语诊疗、智能语音教学、濒危语言保护等新兴应用场景。同时, 加快制定统一的数据标准, 搭建数据共享平台, 推动超声语音生理研究走向规范化、体系化发展, 逐步探索产业化落地路径, 使其更好地服务于语言建设与保护及全民言语健康保障工作, 期待后续能够涌现出更多高质量、高水平的研究成果。

基金项目

国家社科基金一般项目“裕固语语音声学 and 发音生理特征研究”(23BYY086)。

参考文献

- [1] Kelsey, C.A., Minifie, F.D. and Hixon, T.J. (1969) Applications of Ultrasound in Speech Research. *Journal of Speech and Hearing Research*, **12**, 564-575. <https://doi.org/10.1044/jshr.1203.564>
- [2] Preston, J.L., McAllister Byun, T., Boyce, S.E., Hamilton, S., Tiede, M., Phillips, E., et al. (2017) Ultrasound Images of the Tongue: A Tutorial for Assessment and Remediation of Speech Sound Errors. *Journal of Visualized Experiments*, No. 119, Article ID: 55123. <https://doi.org/10.3791/55123-v>
- [3] Wommack, B. (1988) A Selective Review of Current Technology Applicable to Orofacial Myology Research. *International Journal of Orofacial Myology*, **14**, 17-26. <https://doi.org/10.52010/ijom.1988.14.3.4>
- [4] Stone, M. (2005) A Guide to Analysing Tongue Motion from Ultrasound Images. *Clinical Linguistics & Phonetics*, **19**, 455-501. <https://doi.org/10.1080/02699200500113558>
- [5] Hudu, F., Miller, A. and Pulleyblank, D. (2009) Ultrasound Imaging and Theories of Tongue Root Phenomena in African Languages. *Proceedings of Conference on Language Documentation and Linguistic Theory*, London, 13-14 November 2009, 153-163.
- [6] Vazquez-Alvarez, Y. and Hewlett, N. (2007) The ‘Trough Effect’: An Ultrasound Study. *Phonetica*, **64**, 105-121. <https://doi.org/10.1159/000107912>
- [7] Palo, P., Schaeffler, S. and Scobbie, J.M. (2015) Effect of Phonetic Onset on Acoustic and Articulatory Speech Reaction Times Studied with Tongue Ultrasound. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*, Glasgow, 10-14 August 2015, 300-303.
- [8] Rivera Campos, A., Boyce, S.E., Hwu, F. and DeMott, B.N. (2022) Tongue Root Configuration of the Apicoalveolar Trill /r/: An Ultrasound Imaging Study. *Communication Disorders Quarterly*, **44**, 143-151. <https://doi.org/10.1177/1525740122111531>
- [9] 索昂代占, 石海强, 更太加, 等. 基于舌位变化和声学信号的藏语方言单辅音生理语音研究[J]. 信息化研究, 2023, 49(1): 25-31.
- [10] 龙代吉草, 更太加, 石海强, 等. 基于超声成像的藏语安多方言元音研究[J]. 信息化研究, 2024, 50(4): 42-47.
- [11] Hertz, C.H., Lindström, K. and Sonesson, B. (1970) Ultrasonic Recording of the Vibrating Vocal Folds: A Preliminary

- Report. *Acta Oto-Laryngologica*, **69**, 223-230. <https://doi.org/10.3109/00016487009123357>
- [12] 刘劲荣, 何根源, 张琪. 基于超声检测的汉语普通话基础元音声带运动研究[J]. 汉语史与汉藏语研究, 2019(1): 56-74.
- [13] Gugatschka, M., Kiesler, K., Chibidziura-Priesching, J., Schöckler, B. and Friedrich, G. (2009) Pitch Synchronous Changes of the Anterior Cricothyroid Gap by Using Ultrasonography. *Journal of Voice*, **23**, 610-613. <https://doi.org/10.1016/j.jvoice.2008.01.011>
- [14] Carignan, C. (2018) Using Ultrasound and Nasalance to Separate Oral and Nasal Contributions to Formant Frequencies of Nasalized Vowels. *The Journal of the Acoustical Society of America*, **143**, 2588-2601. <https://doi.org/10.1121/1.5034760>
- [15] Bouavichith, D.A., Namboodiripad, S. and Garellek, M. (2018) A Contrastive Place of Articulation between Palatals and Velars: An Ultrasound Study of Malayalam Palatal-Velar Nasals. *The Journal of the Acoustical Society of America*, **143**, 1755-1755. <https://doi.org/10.1121/1.5035742>
- [16] 马良, 岳暘, 陈忠敏. 基于超声影像的上海话鼻化元音发音研究[J]. 语言研究集刊, 2023(1): 361-371, 395.
- [17] Krebs, V.L., Sedarous, Y. and Miller, A.L. (2013) Consonant-Vowel Coarticulation in Velar Plosives. *Proceedings of Meetings on Acoustics*, **19**, Article ID: 060284. <https://doi.org/10.1121/1.4800991>
- [18] Saito, M., Tomaschek, F., Ching-Chu Sun, R. and Baayen, H. (2021) An Ultrasound Study of Frequency and Co-Articulation. PsyArXiv. <https://doi.org/10.31234/osf.io/6yndk>
- [19] 宋清逸. 基于超声波检测的拉萨藏语元音发音舌位运动研究[D]. 上海师范大学, 2020.
- [20] Shawker, T.H., Stone, M. and Sonies, B.C. (1985) Tongue Pellet Tracking by Ultrasound: Development of a Reverberation Pellet. *Journal of Phonetics*, **13**, 135-146. [https://doi.org/10.1016/s0095-4470\(19\)30741-7](https://doi.org/10.1016/s0095-4470(19)30741-7)
- [21] Lawson, E., Scobbie, J.M. and Stuart-Smith, J. (2011) The Social Stratification of Tongue Shape for Postvocalic /r/ in Scottish English. *Journal of Sociolinguistics*, **15**, 256-268. <https://doi.org/10.1111/j.1467-9841.2011.00464.x>
- [22] Sepulveda Sepulveda, F.A., Porras-Plata, D. and Sarria-Paja, M. (2020) Speaker Verification System Based on Articulatory Information from Ultrasound Recordings. *DYNA*, **87**, 9-16. <https://doi.org/10.15446/dyna.v87n213.81772>
- [23] 武丹. 基于超声的多模态发音数据采集辅助设备的研究[D]: [硕士学位论文]. 天津: 天津大学, 2021.
- [24] 寇贇. 基于超声影像的舌位参数提取及分析系统[D]: [硕士学位论文]. 兰州: 西北民族大学, 2019.
- [25] Van Lieshout, P.H., Peters, H.F. and Hulstijn, W. (1994) Articulo-Motorisch Onderzoek: Een overzicht van technieken met de nadruk op Elektro-Magnetische Medio-sagittale Articulografie (EMMA). *Stem-, Spraak-en Taalpathologie*, **3**, 241-261.
- [26] Stone, M., Sonies, B.C., Shawker, T.H., Weiss, G. and Nadel, L. (1983) Analysis of Real-Time Ultrasound Images of Tongue Configuration Using a Grid-Digitizing System. *Journal of Phonetics*, **11**, 207-218. [https://doi.org/10.1016/s0095-4470\(19\)30822-8](https://doi.org/10.1016/s0095-4470(19)30822-8)
- [27] Gick, B., Bernhardt, B.M., Bacsfalvi, P. and Wilson, I. (2008) 11. Ultrasound Imaging Applications in Second Language Acquisition. *Phonology and Second Language Acquisition*, **36**, 309-322. <https://doi.org/10.1075/sibil.36.15gic>
- [28] Davidson, L. (2005) Addressing Phonological Questions with Ultrasound. *Clinical Linguistics & Phonetics*, **19**, 619-633. <https://doi.org/10.1080/02699200500114077>
- [29] Tsui, H.M. (2012) Ultrasound Speech Training for Japanese Adults Learning English as a Second Language. Ph.D. Thesis, University of British Columbia.
- [30] Suzuki, K., Wilson, I. and Watanabe, H. (2017) Visual Learning 2: Pronunciation App Using Ultrasound, Video, and MRI. *IN-TERSPEECH 2017*, Stockholm, 20-24 August 2017, 831-832.
- [31] Mennen, I., Scobbie, J.M., de Leeuw, E., Schaeffler, S. and Schaeffler, F. (2010) Measuring Language-Specific Phonetic Settings. *Second Language Research*, **26**, 13-41. <https://doi.org/10.1177/0267658309337617>
- [32] Wilson, I., Gick, B., O'Brien, M.G., Shea, C. and Archibald, J. (2006) Ultrasound Technology and Second Language Acquisition Research. *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)*, 148-152.
- [33] Cleland, J., McCron, C. and Scobbie, J.M. (2013) Tongue Reading: Comparing the Interpretation of Visual Information from inside the Mouth, from Electropalatographic and Ultrasound Displays of Speech Sounds. *Clinical Linguistics & Phonetics*, **27**, 299-311. <https://doi.org/10.3109/02699206.2012.759626>
- [34] 谭亚鑫, 沈向荣. 基于超声动态舌形的普通话元音发音特点研究[J]. 听力学及言语疾病杂志, 2024, 32(4): 317-321.
- [35] 张文平. 汉语语音生理语料库在对外汉语教学中的应用研究[D]: [硕士学位论文]. 天津: 天津大学, 2014.
- [36] 魏远. 基于超声和 MRI 图像融合的发音过程可视化系统[D]: [硕士学位论文]. 天津: 天津大学, 2018.

- [37] Barberena, L.d.S., Brasil, B.d.C., Melo, R.M., Mezzomo, C.L., Mota, H.B. and Keske-Soares, M. (2014) Ultrasound Applicability in Speech Language Pathology and Audiology. *CoDAS*, **26**, 520-530. <https://doi.org/10.1590/2317-1782/20142013086>
- [38] Saigusa, H., Saigusa, M., Aino, I., Iwasaki, C., Li, L. and Niimi, S. (2006) M-Mode Color Doppler Ultrasonic Imaging of Vertical Tongue Movement during Articulatory Movement. *Journal of Voice*, **20**, 38-45. <https://doi.org/10.1016/j.jvoice.2005.01.003>
- [39] Zharkova, N. (2013) Using Ultrasound to Quantify Tongue Shape and Movement Characteristics. *The Cleft Palate Craniofacial Journal*, **50**, 76-81. <https://doi.org/10.1597/11-196>
- [40] Bernhardt, B., Gick, B., Bacsfalvi, P. and Adler-Bock, M. (2005) Ultrasound in Speech Therapy with Adolescents and Adults. *Clinical Linguistics & Phonetics*, **19**, 605-617. <https://doi.org/10.1080/02699200500114028>
- [41] Cavin, M. (2015) The Use of Ultrasound Biofeedback for Improving English /r/. *Working Papers of the Linguistics Circle*, **25**, 32-41.
- [42] Bernhardt, B., Gick, B., Bacsfalvi, P. and Ashdown, J. (2003) Speech Habilitation of Hard of Hearing Adolescents Using Electropalatography and Ultrasound as Evaluated by Trained Listeners. *Clinical Linguistics & Phonetics*, **17**, 199-216. <https://doi.org/10.1080/0269920031000071451>
- [43] Bryfonski, L. (2023) Is Seeing Believing? The Role of Ultrasound Tongue Imaging and Oral Corrective Feedback in L2 Pronunciation Development. *Journal of Second Language Pronunciation*, **9**, 103-129. <https://doi.org/10.1075/jslp.22051.bry>
- [44] Ribeiro, M.S., Cleland, J., Eshky, A., Richmond, K. and Renals, S. (2021) Exploiting Ultrasound Tongue Imaging for the Automatic Detection of Speech Articulation Errors. *Speech Communication*, **128**, 24-34. <https://doi.org/10.1016/j.specom.2021.02.001>
- [45] Zharkova, N., Gibbon, F.E. and Lee, A. (2016) Using Ultrasound Tongue Imaging to Identify Covert Contrasts in Children's Speech. *Clinical Linguistics & Phonetics*, **31**, 21-34. <https://doi.org/10.1080/02699206.2016.1180713>
- [46] Wiethan, F., Ceron, M.I., Marchetti, P., Giacchini, V. and Mota, H.B. (2015) O uso da eletroglotografia, eletromiografia, espectrografia e ultrassom nos estudos de fala—Revisão teórica. *Revista CEFAC*, **17**, 115-125. <https://doi.org/10.1590/s1516-18462013005000049>
- [47] McKeever, L., Cleland, J. and Delafield-Butt, J. (2021) Using Ultrasound Tongue Imaging to Analyse Maximum Performance Tasks in Children with Autism: A Pilot Study. *Clinical Linguistics & Phonetics*, **36**, 127-145. <https://doi.org/10.1080/02699206.2021.1933186>
- [48] Spencer, C., Vannest, J., Preston, J.L., Maas, E., Sizemore, E.R., McAllister, T., *et al.* (2023) Neural Changes in Children with Residual Speech Sound Disorder after Ultrasound Biofeedback Speech Therapy. *Journal of Speech, Language, and Hearing Research*, **66**, 3223-3241. https://doi.org/10.1044/2023_jslhr-22-00430
- [49] 吕立松, 李清华. 跨界与创新的语言与医学交叉研究综述[J]. 医学与哲学, 2024, 45(24): 72-75.
- [50] 马文, 苏永刚. 新医科背景下语言康复国际化人才培养模式探索与实践[J]. 中国外语, 2023, 20(5): 18-23.
- [51] 施浩, 杨君洁, 刘虹, 等. 身体哲学视角下医学学科交叉发展的问题与对策研究[J]. 医学与哲学, 2023, 44(14): 20-24.
- [52] 周频. 论科学轨道上的语言学学科建设[J]. 当代外语研究, 2012(10): 2-8, 76.
- [53] 王慧, 鲍毓, 徐琳, 等. 儿童功能性构音障碍的临床分析及言语治疗[J]. 中华物理医学与康复杂志, 2011, 33(11): 835-838.
- [54] Al Ani, S. (2023) Systematic Review of Deep Learning Models in Ultrasound Tongue Imaging for the Detection of Speech Disorders. Authorea Preprints. <https://doi.org/10.36227/techrxiv.22699291>