

近十五年国内英语听力测试研究现状与展望

陈琳琳

曲阜师范大学外国语学院, 山东 济宁

收稿日期: 2025年3月7日; 录用日期: 2025年4月15日; 发布日期: 2025年4月28日

摘要

本研究采用文献计量分析法, 系统考察了2009~2023年间发表在国内核心期刊的英语听力测试研究文献。作为语言能力评估体系的重要组成部分, 英语听力测试不仅能够全面反映学习者的语言综合能力, 还对教学实践具有重要的反拨作用。研究发现: 1) 研究对象分布不均, 以大学生群体为主, 初高中阶段研究明显不足; 2) 研究范畴集中于全国性统一考试, 对区域性及校本测试关注度较低; 3) 研究主题呈现多元化特征, 主要聚焦于测试题型设计、信效度验证、反拨效应分析、认知诊断理论应用、真实性探讨及难度控制等方面; 4) 新兴研究领域如技术辅助测试及听力理解的心理认知机制等尚未得到充分重视。通过系统梳理近十五年的研究进展, 揭示了该领域的研究现状与不足, 为未来的研究方向提供了理论参考和实践指导。

关键词

英语听力测试, 测试效度, 测试的反拨效应, 研究综述

Research Status and Prospect of Domestic English Listening Tests in the Past Fifteen Years

Linlin Chen

College of Foreign Languages, Qufu Nominal University, Jining Shandong

Received: Mar. 7th, 2025; accepted: Apr. 15th, 2025; published: Apr. 28th, 2025

Abstract

This study employs bibliometric analysis to systematically investigate research literature on English listening assessment published in Chinese core journals between 2009 and 2023. As a crucial component of language proficiency evaluation systems, English listening tests not only comprehensively reflect learners' integrated language abilities, but also exert significant backwash effects on teaching practices. The research reveals the following characteristics: 1) Research subjects are unevenly

distributed, with predominant focus on college students and insufficient attention to secondary education levels; 2) Research scope primarily concentrates on national standardized tests, while regional and school-based assessments receive limited attention; 3) Research themes are diversified, mainly focusing on test item design, reliability and validity verification, backwash effect analysis, application of cognitive diagnostic theory, authenticity exploration, and difficulty control; 4) Emerging research areas such as technology-assisted testing and psychological-cognitive mechanisms of listening comprehension remain understudied. Through systematically reviewing the research progress over the past fifteen years, this study reveals the current status and limitations in this field, providing theoretical references and practical guidance for future research directions.

Keywords

English Listening Test, Test Validity, Test's Backwash Effect, Research Review

Copyright © 2025 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

在英语作为第二语言的学习过程中，听力理解能力的评估需要通过标准化的测试来实现，这是衡量该语言能力的主要手段。而有效可信的英语听力测试离不开与之相关的科学研究。鉴于其重要性，近年来国内学者对于英语听力测试的研究有所加强，由此，研究领域得到了拓展，研究的广度和深度得到了提高。本研究采用文献计量分析方法，对 2009~2023 年期间国内英语听力测试研究进行系统性综述，重点考察研究热点、方法演进，以揭示该领域的发展轨迹与未来研究方向。

2. 英语听力测试的研究趋势

基于中国知网学术趋势分析工具，本研究对“英语听力测试”、“听力理解测试”以及“听力测试效率”三个关键词进行检索，得出以下发现：首先，就“英语听力测试”研究而言，2009~2011 年间呈现显著上升态势，其中 2011 年达到峰值，年度发文量达 34 篇(见图 1)。然而，2012~2023 年期间，该领域研究热度整体呈下降趋势。其次，“听力理解测试”研究趋势同样呈现先升后降的特征。数据分析显示，2010 年为该领域研究高峰期，当年中国知网收录相关文献 40 篇(见图 2)。最后，关于“听力测试效率”的研究呈现出波动性发展态势。值得注意的是，2015 年该领域研究达到顶峰，年度发文量为 7 篇(见图 3)，此后虽有起伏，但整体维持在相对稳定的水平。

3. 英语听力测试的研究方法

本研究采用高一虹等(1999)提出的研究方法分类框架，应用语言学研究方法划分为实证性研究与非材料性研究两大范畴[1]。以此为依据，对 24 篇英语听力测试研究的文章就其研究方法进行了分类，其中实证性研究为 16 篇，占 66.7%，非材料性研究为 8 篇，占 33.3% (表 1)。国内英语听力测试研究中实证性研究总体比例高于非材料性研究，这与语言测试研究自身的特征密不可分，因为离开实验研究和数据的语言测试无异于纸上谈兵。

研究数据显示，英语听力测试研究方法呈现出显著的范式转变。蒋显菊(2007)的研究表明，1998~2007 年间，非材料性研究方法占据主导地位(80.1%)，而实证性研究仅占 19.9% [2]。彭康洲和袁巧玲(2015)的研究进一步揭示，2000~2015 年期间，非材料性研究比例下降至 55%，实证性研究则上升至 45% [3]。基

于本研究的数据统计,近15年来实证性研究方法的使用率已超越非材料性研究方法33.4个百分点(见表1),这一显著变化反映了该领域研究方法的实质性转变。



Figure 1. Trends in listening tests research (2009~2023)

图 1. 听力测试研究趋势(2009~2023 年)

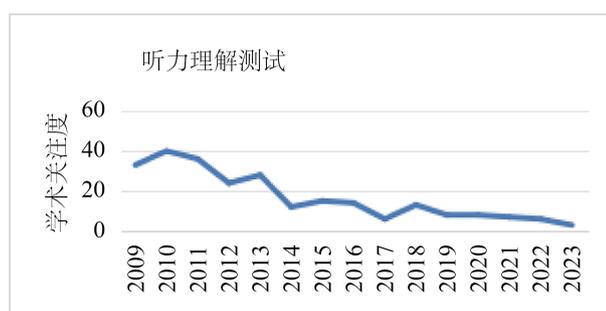


Figure 2. Trends in listening comprehension research (2009~2023)

图 2. 听力理解研究趋势(2009~2023 年)



Figure 3. Trends in research on the validity of listening tests (2009~2023)

图 3. 听力测试效度研究趋势(2009~2023 年)

Table 1. Status of research methods for English listening test (2009~2023)

表 1. 英语听力测试研究方法情况(2009~2023 年)

分类	实证性研究	非材料性研究	总计
篇数	16	8	24
比例	66.7%	33.3%	100%

究其原因,我们认为这一转变主要源于以下两个方面:首先,语言测试理论的纵深发展推动了理论建构与实证研究的有机结合;其次,研究者对定量研究方法和分析工具的掌握程度显著提升,为开展实证研究提供了方法学支撑。

4. 英语听力测试的研究对象

基于研究对象的教育阶段属性,本研究对听力测试研究进行了系统分类。数据分析表明(见表2),大学生群体最受研究者关注,相关文献达13篇,其中非英语专业学生研究12篇,英语专业学生研究1篇。其次为高中生群体,相关研究6篇。相比之下,初中生听力测试研究明显不足,仅2篇文献涉及。

Table 2. Current research status on examinees as the subject

表 2. 以考生主体为对象的论文情况

考生	大学		中学		其他
	英语专业	非英语专业	高中	初中	无特别说明研究对象的年龄段
篇数	1	12	6	2	3
比例	4.2%	50%	25%	8.3%	12.5%

研究数据显示(见表3),2009~2023年间,听力测试研究主要集中于全国性统一考试,而针对区域性考试及校本测试的研究明显不足。这里所研究的大学统一考试包括:大学英语四、六级(CET4/6),专业英语四、八级(TEM4/8)、雅思、托福等考试。

Table 3. Current research status on test types as the subject

表 3. 以测试类型为研究对象的论文情况

试题	全国性统一听力考试 (大学统一考试、全国高考)	非统一性听力考试 (地区考试、校本考试)	其他无特别说明的考试
篇数	11	3	10
比例	45.8%	12.5%	41.7%

5. 听力测试的研究内容

基于对24篇文献的系统分析(见表4),本研究采用主题分类法将其归纳为六大研究类别。在分类过程中,通过文本细读和内容分析,重点考察了各文献的核心研究主题,并妥善处理了主题交叉问题。从研究主题分布来看,听力测试的反拨效应及信效度研究占据主导地位,而命题研究、认知诊断理论探讨及测试方法创新等主题的文献数量相对均衡。

Table 4. Classification of literature

表 4. 文献分类情况

类别	篇数	所占比例
反拨效应	7	29.2%
题型、命题方法	4	16.7%
信效度	6	25%
认知诊断评估	4	16.7%
真实性、难度	3	12.5%

5.1. 听力测试的反拨效应

反拨效应指语言测试对教学实践和学习过程产生的影响。自 20 世纪 90 年代以来, 该领域研究受到学界持续关注, 形成了多维度的研究范式。代表性研究包括: 张宁娇和杜苗(2010)从测试内容、形式维度, 结合教材、教师、学生三重视角, 系统考察了新四级听力测试对大学英语教学的双向反拨效应[4]。杨维秀和陈洪宇(2012)构建了四级考试反拨效应的三阶段模型: 早期阶段受限于测试信息不足, 正向效应未能充分显现; 中期阶段随着测试经验的积累, 教与学逐步趋于理性; 后期阶段则出现教学重心向应试技巧偏移的现象, 导致正向效应减弱[5]。董连忠(2017)聚焦高考英语听力测试, 在肯定其提升学生听力能力的同时, 提出了优化考纲、提高效率、改进教法等对策建议[6]。

5.2. 听力测试的题目类型、命题

此类研究聚焦听力测试的题型设计及其影响。蒋京丽(2019)基于语篇、情境和语用三个维度, 构建了系统的命题框架, 并据此开发了适用于不同年级和主题的听力测试题型, 为听力测试命题提供了系统的理论框架和实践指导[7]。雷雪梅和席仲恩(2016)通过对比实验, 探讨了试题语言对听力测试的影响, 发现母语命制的试题虽能提供辅助信息, 但在主旨理解类题目中可能产生干扰效应。上述研究在理论构建和实践应用方面均作出了重要贡献[8]。然而, 现有研究多局限于特定测试类型, 缺乏对不同文化背景下的普适性探讨, 这为未来研究指明了方向。

5.3. 听力测试的信度和效度

在听力测试效度研究领域, 学者们从多维度展开深入探讨。莫玉秀和覃元姣(2011)采用多元效度验证框架, 指出四级网考听力测试的共时效度可通过规范命题和试测程序得以保证, 但其预测效度和内容效度仍存在改进空间[9]。钟伟轩(2012)强调在听力测试设计中平衡信度与效度的重要性[10]。李廉和杨舒(2018)基于整体效度观, 运用因子分析法提取听力构念, 验证了仿真测试的效度建构可行性, 为提高测试效度提供了新思路[11]。上述研究系统探讨了听力测试效度的关键问题, 从理论构建到实践应用均有所突破, 为提升测试质量提供了重要理论支撑和实践指导。

5.4. 认知诊断评估

认知诊断评估作为传统评价方式的重要补充, 近年来受到学界广泛关注。肖云南和罗娟(2019)运用 G-DINA 模型, 从语言学与认知心理学视角识别出 8 项关键认知属性, 并系统分析了学生群体、分层群体及个体三个层面的听力认知特征与缺陷, 揭示了不同水平学习者的显著差异[12]。何莲珍和张娟(2021)则聚焦《中国英语能力等级量表》的应用研究, 通过开发个性化诊断报告, 将量表描述语言与补偿式教学有机结合, 有效提升了学习者的语言能力[13]。认知诊断研究在模型构建和应用实践方面取得显著进展, 为传统评价方式的革新提供了新的研究方向和实践路径。

5.5. 听力测试的真实性、难度

此类研究涉及测试的各种特征、各种测试方法。刘宁(2012)探讨了雅思听力测试的真实性。她指出雅思听力测试的真实性体现在真实生活用语的使用及其交互性方面, 并指出真实性是指听力材料应当是真实的、现实生活中所用的自然语言[14]。何莲珍等(2018)采用 Coh-Metrix 文本分析工具与 SPSS 统计分析软件, 通过多维度的实证研究, 系统考察了测试方法对英语听力任务难度的作用机制及其特征间的相互关系。该研究不仅为听力测试设计提供了理论框架, 还通过实证数据支持了相关假设, 具有重要的方法论意义[15]。大部分都是对于大学的听力测试研究, 对中小学听力测试几乎无研究, 未来可参照教材探究其不同特征。

6. 结语

近年来,国内在英语听力测试领域的研究取得了显著进展。具体体现在以下几个方面:首先,研究方法经历了从单纯的理论探讨和问题分析向更多实证性研究的转变,这种趋势的增长对于我国听力教学与测试任务的科学化设计及开发提供了重要支持;其次,研究视角从单纯关注听力测试本身转向更加重视应试者群体,这一变化体现了“以评促学”的理念,强调了考生在学习过程中的主体地位;再者,跨学科研究逐渐受到重视,通过借鉴其他学科的理论和方法,丰富了听力测试的研究手段;最后,随着多种自然语言处理工具的应用,听力测试研究的精确度得到了有效提升。

然而,当前国内听力测试研究仍存在一些不足之处:一方面,多数研究集中于高等教育阶段的听力测试,而对基础教育阶段的相关测试关注不足;另一方面,研究多侧重于大规模、高风险测试,而对于小规模或校本课堂测试的研究则相对稀缺;此外,尽管现代科技手段为改进听力测试方法、实现测试智能化与个性化提供了可能,但相关的探索仍然有限;在题型研究方面,更关注客观题,而对主观题关注不够。更重要的是,从2015年开始,国内英语听力测试的研究数量呈下降趋势,研究者的兴趣可能转向了其他领域,如口语、写作或综合语言能力评估。在未来的听力测试研究中要更全面地考虑不同层次教育的需求以及题型多样性;进行多模态研究,研究听力与其他语言技能(如口语、阅读)的相互作用。

展望未来,笔者认为国内的听力测试研究可以进一步从以下几个方面进行拓展。一,拓展研究对象,多关注小规模测试和基础教育测试。比如在基础教育测试方面,设计适合不同学段的听力测试框架,结合课程标准,开发符合学生认知发展水平的测试任务。在一些小规模测试中鼓励校本测试研究,探索基于课堂的听力评估模式,开发适用于小规模测试的题库和评分标准。

二,拓展对影响听力能力测试表现因素的全面性探究,深入分析影响听力理解能力的各种因素,如语言输入环境、语音识别能力、心理因素(焦虑、自信心)、认知策略运用以及跨文化交际意识等,并探讨如何通过教学干预改善这些影响因素。比如探讨听力理解中的认知策略运用,如预测、推断、总结等,并开发相应的教学干预方案。例如,可以通过实验研究,比较不同认知策略对学生听力理解能力的影响,并设计相应的教学策略,如预测训练、推断训练等。

三,开发更具挑战性和适应时代变化的新颖题型,比如利用视频、音频等多媒体资源,模拟真实场景对话,测试学生在复杂语境下的听力能力。例如,可以设计基于视频的听力测试题目,如观看一段英语电影片段后回答问题,测试学生对真实语境下听力材料的理解能力。增加主观题比例,如听后复述、听说测试(Listen to Speak)、听写测试(Listen to Write)等,以更全面地评估学生的听力理解能力。同时,制定更为细致、科学且能够体现多元技能水平的评分标准。最后,改进研究方法。运用多维度数据分析方法,如结构方程模型(SEM),探究变量间的复杂关系,分析语言输入环境、心理因素、认知策略和跨文化交际意识对听力能力的综合影响。

利用先进的技术手段,如眼动实验研究学生在听力理解过程中的注意力分配情况,分析不同听力材料对学生注意力的影响。为听力理解过程提供更加客观、直观的数据支持,从而深化对听力理解机制的认识。

本文的研究结果对未来听力测试研究具有一定的参考意义,但仍然存在不足,即研究样本量较少,可能限制了研究结果的普适性。未来研究可扩大样本范围,验证结果的普适性。

参考文献

- [1] 高一虹,李莉春,吕玉君. 中、西应用语言学研究方法发展趋势[J]. 外语教学与研究, 1999, 31(2): 8-16.
- [2] 蒋显菊. 国内英语测试研究:十年回顾与展望[J]. 外语界, 2007(2): 89-96.
- [3] 彭康洲,袁巧玲. 国内近十五年英语听力测试研究述评[J]. 外语测试与教学, 2015(2): 21-27.

- [4] 张宁娇, 杜苗. 新四级听力测试对大学英语听力教学的反拨作用[J]. 首都师范大学学报(社会科学版), 2010(S3): 22-24.
- [5] 杨维秀, 陈洪宇. 英语听能测试对学习者的反拨作用研究[J]. 东北师大学报(哲学社会科学版), 2012(3): 216-218.
- [6] 董连忠. 高考英语听力考试的反拨效应研究[J]. 外语测试与教学, 2017(1): 42-49.
- [7] 蒋京丽. 基于学科能力的初中英语听力测试命题研究[J]. 中国考试, 2019(3): 20-27.
- [8] 雷雪梅, 席仲恩. 听力考试试题命制中的母语和外语效应[J]. 外语测试与教学, 2016(2): 23-31.
- [9] 莫玉秀, 覃元姣. 大学英语四级网考听力测试的效度探析[J]. 广西师范大学学报(哲学社会科学版), 2011, 47(2): 84-88.
- [10] 钟伟轩. 英语听力测试设计中存在的问题与对策[J]. 教学与管理, 2012(11): 138-139.
- [11] 李廉, 杨舒. 英语听力测试构念效度仿真建构研究[J]. 外语测试与教学, 2018(2): 15-26.
- [12] 肖云南, 罗娟. 大学英语分级测试听力理解认知诊断研究[J]. 湖南大学学报(社会科学版), 2019, 33(1): 113-118.
- [13] 何莲珍, 张娟. 《中国英语能力等级量表》在补偿式教学与学习中的应用[J]. 外语测试与教学, 2021(3): 1-11.
- [14] 刘宁. 雅思听力测试的真实性及其教学改革的思考[J]. 中国劳动关系学院学报, 2012, 26(6): 106-109.
- [15] 何莲珍, 陈大建, 闵尚超. 英语听力测试中测试方法对任务难度的影响研究[J]. 现代外语, 2018, 41(1): 43-54+146.