

数智驱动背景下的语音识别技术在英语口语教学中的应用研究

宁晓婷

扬州大学外国语学院, 江苏 扬州

收稿日期: 2025年10月22日; 录用日期: 2025年11月19日; 发布日期: 2025年11月26日

摘要

全球化与智慧教育背景下, 大学英语口语教学受大班制反馈不足、学生自主练习低效等问题制约。本文定位为基于系统性文献回顾的教学模式构建与应用提案, 基于纠正性反馈理论, 系统梳理国内外相关研究与典型案例, 分析数智驱动下语音识别技术在英语口语“识别-教学-反馈”环节的应用。研究表明, 语音识别技术可实现段音规模化检测、客观评估与即时反馈, 缓解教师反馈压力, 支撑“课前-课中-课后”学习; 但在超段音检测、错误原因解释及情感支持上存在局限, 需与教师协同形成混合反馈模式。本文通过综述明确语音识别技术的应用价值与局限, 为大学英语口语教学智能化转型提供参考, 也为纠正性反馈理论在新技术场景的应用拓展思路。

关键词

数智驱动, 语音识别技术, 英语口语教学, 纠正性反馈理论, 混合教学模式

A Study on the Application of Speech Recognition Technology in English Speaking Instruction under the Data-Intelligence-Driven Context

Xiaoting Ning

College of International Studies, Yangzhou University, Yangzhou Jiangsu

Received: October 22, 2025; accepted: November 19, 2025; published: November 26, 2025

Abstract

In the context of globalization and smart education, college English speaking instruction faces challenges such as limited feedback in large-class settings and low efficiency in students' autonomous practice. This study positions itself as a teaching model construction and application proposal based on a systematic literature review. Based on corrective feedback theory, this paper systematically reviews relevant domestic and international research and representative cases, analyzing the application of Automatic Speech Recognition (ASR) technology in the "recognition-instruction-feedback" stages of English speaking under intelligent data-driven conditions. The study finds that ASR can achieve large-scale phoneme-level detection, objective evaluation, and instant feedback, thereby alleviating teachers' feedback burden and supporting learning before, during, and after class. However, limitations remain in suprasegmental feature detection, error cause interpretation, and emotional support. Thus, a hybrid feedback model integrating teacher participation is necessary. Through this review, the paper clarifies the value and limitations of ASR in college English speaking instruction, providing references for the intelligent transformation of English teaching and expanding the application scope of corrective feedback theory in emerging technological contexts.

Keywords

Data-Intelligence-Driven, Speech Recognition Technology, English Speaking Instruction, Corrective Feedback Theory, Blended Teaching Model

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

在全球化与信息化不断加深的背景下，英语口语能力在学术交流与社会交往中的地位日益凸显。然而在我国英语教学实践中，口语教学长期处于薄弱环节。我国高校大学英语口语教学长期受限于“大班制授课、教师反馈覆盖不足、学生自主练习缺乏指导”的现实困境。在 40~50 人的常规班级中，教师每节课仅能对 5~8 名学生进行发音纠错，其余学生的“隐性发音偏误”难以得到及时修正，进而形成固化习惯[1]。与此同时，学生课后自主练习时，因缺乏即时反馈机制，常陷入“机械重复错误发音”的困境，导致练习效率低下。这一现实困境使得如何有效提高英语口语教学效率成为外语教育急需解决的问题。

语音识别技术的发展为改善这一局面提供了新的可能。与传统人工反馈相比，ASR 具有客观性、一致性和可扩展性的优势，可以大规模应用于课堂与课外学习场景，为教师减轻重复性的检测任务，也为学习者提供更多自主练习机会。在“智慧教育”和“数智驱动”的大背景下，探索 ASR 在英语口语教学中的应用已成为学界和教育实践的重要议题[2]。Inceoglu 对 19 名韩国 EFL 大学生的研究显示，移动端 ASR 练习可显著改善英语 /i/ 元音发音，68% 学习者认可其实用价值[3]。但 ASR 也存在明显局限，Wang & Young 于 2015 年指出仅 30% 青少年能通过 ASR 反馈理解错误根源，其余需教师补充元语言解释[4]。

基于此，本研究以系统性文献回顾为基础，梳理理论与既有研究，结合典型案例分析揭示 ASR 的功能与价值，构建可落地的混合反馈教学模式并配套评估方案，最终形成“理论梳理 - 模式构建 - 教学提案 - 评估设计”的完整逻辑，为英语口语教学的数智化发展提供兼具理论性与实践性的参考。

2. 文献综述

2.1. 理论框架

语音识别技术源于 20 世纪 50 年代，90 年代隐马尔可夫模型(HMM)提升其识别准确率并推动实用化。最初，ASR 主要用于语言测试的自动评分，随后逐渐扩展至发音训练、口语练习和交互式学习。其核心功能不仅是将语音转写为文字，更在于通过算法检测学习者的发音偏误并提供即时反馈，从而形成“检测 - 反馈 - 修正”的学习闭环[5]。

在教育学和二语习得理论视角下，ASR 的功能与反馈理论高度契合。Swain 的输出假说指出，学习者在语言产出过程中通过觉察与修正偏误实现习得，而这一过程依赖有效反馈[6]。纠正性反馈(Corrective Feedback, CF)理论进一步揭示了反馈在语言学习中的关键作用：其一，反馈帮助学习者注意到输入与输出之间的差距，形成“注意假说”所强调的显性觉察[7]；其二，反馈通过即时或延迟的形式引导学习者修正，促进“形式聚焦”与语言内化[8]。

Lyster 和 Ranta 在其经典研究中将课堂纠正性反馈分为六类，不同反馈类型对学习者的注意与 uptake 的影响存在差异[9]。为明确 ASR 与教师的功能边界及协同机制，表 1 将六种反馈类型与 ASR 功能、教师/同伴角色、学习者行为进行精准对应。

Table 1. Types of corrective feedback and corresponding ASR functions

表 1. 纠正性反馈类型与 ASR 功能对应关系

反馈类型	ASR 功能实现情况	教师/同伴角色	学习者行为
明示性纠正	可实现：标红错误音素/单词、播放标准语音、生成正确率数据	补充复杂错误解释、验证 ASR 检测结果	对比 ASR 提供的标准音与自身发音、修正错误并重复练习
重述	可实现：将错误发音转写为正确文本、播放重述后的完整句子音频	强化重述场景性、纠正 ASR 未识别的语境化偏误	模仿 ASR 或者教师的重述内容、在真实交际中运用正确表达
澄清请求	不可实现：无法理解学习者的语境意图，无法主动发起“请再说一遍”等请求	发起澄清提问、引导同伴间相互澄清	反思发音模糊的原因、重新组织语言并清晰表达
元语言解释	有限实现：仅能提示错误类型，无法提供规则解释	提供语法/语音规则解释、举例拓展	记录元语言规则、结合规则修正发音
引导提示	有限实现：可提供错误位置提示，无法引导自主发现错误原因	设计提示问题、引导同伴相互提示	基于提示自主探索错误原因、尝试多种发音方式直至正确
重复	可实现：重复播放学习者的错误发音与标准发音，形成对比	重复学习者的错误发音以引起注意、组织同伴重复练习	跟随 ASR 或教师重复正确发音、通过多次重复固化正确发音习惯

2.2. 国内外研究现状

2.2.1. 国外研究现状

国外对 ASR 与二语口语教学的研究已形成多维度实证体系，围绕 ASR 在语音教学中的技术适配性与教学有效性展开系统探索，核心成果集中于三个领域。在音段方面，研究者通过元分析与对照实验明确了 ASR 在音素、单词层面的教学价值。Ngo 等人整合 15 项实证研究，发现 ASR 对段音学习的平均效应量达 $g = 0.82$ (大效应)，其中对学习者的母语中缺失的英语音素纠错效果最为显著[10]。

在超音段检测的探索中，国外研究者尝试突破 ASR 在语调、重音等超音段特征教学中的技术瓶颈。Liakin 针对法语二语学习者设计“ASR + 语调可视化”整合工具，通过波形图实时展示学习者语调的基

频变化,实验结果显示该工具使法语语调纠错效果显著高于英语场景下 ASR 的超段音纠错效果(5%) [11]; Derwing 与 Munro 通过跨语言对比研究进一步指出, ASR 超段音检测效果的差异与目标语言的语音特征直接相关,法语语调依赖固定节奏模式,其声学特征易被算法捕捉,而英语语调随交际语境动态变化,导致 ASR 检测难度显著增加,这一发现为后续超段音领域的技术优化提供了方向指引[12]。

在混合反馈模式的实践尝试中,国外研究重点探索 ASR 与教师、同伴反馈的协同路径,以弥补单一反馈的不足。在 ESL 课堂中, McCrocklin 设计了“ASR 课前诊断-教师课中讲解”的混合模式,学习者课前通过 ASR 完成段音自主检测并生成个人错误报告,教师课中基于数据聚焦班级高频错误音素开展针对性讲解,实验结果显示该模式下学习者的段音正确率达 83%,显著高于纯 ASR 自主练习组[13]; 2024 年后, Zhong、Xie 与 Yao 构建了“ASR 错误检测-LLM 元语言解释”协同反馈机制, ASR 模块负责精准定位超音段错误, LLM 模块则为错误生成规则解释与纠正建议。实验数据显示,该机制使学习者的超段音正确率平均提升 17.8%,显著高于单一 ASR 模式(8.9%),初步实现了“技术检测 + 智能解释”的融合应用[14],这一研究为未来基于数智驱动的个性化语音反馈系统提供了启示。

综上,国外研究侧重从算法性能、反馈形式及学习机制构建系统化框架,形成以实证研究为核心的技术与教学融合体系,探讨 ASR 教学功效与学习者认知过程,趋势体现为从段音识别拓展至超段音检测、从结果性反馈转向解释性反馈、引入大语言模型与多模态交互探索一体化反馈模式。

2.2.2. 国内研究现状

国内研究起步相对较晚,重点解决 ASR 技术在国内英语教学场景中的适配性问题。在本土化 ASR 工具的研发与应用方面,研究者针对中国学习者的典型发音偏误优化技术功能。陈桦团队开发的“智能口语评测系统”,通过调整声学模型的特征权重,将汉语学习者 /θ/、/v/ 音素的错误识别准确率提升至 85%,该系统在国内 10 余所高校公共英语课堂应用中,实现了 40 人以上大班的全员段音实时检测[15]; 2022 年,涂惠燕等人基于手机移动平台开发的 ASR 口语练习工具,其“错误音素即时标红 + 标准音对比播放”功能,适配中学高考口语备考场景,在广东省 30 余所中学的应用中,工具使用率达 90%,有效缓解了教师课后口语作业批改的压力[16]。

在高校与中学场景的教学实证方面,国内研究主要聚焦于 ASR 在标准化英语教学中的有效性验证。中学场景中,于娇娜以某省重点中学高二 200 名学生为对象,设计人工智能辅助英语口语教学实验,结果显示一个学期内,学生口语听力平均分提高 3.91 分,其中低分段学生提升最为显著,说明 ASR 与语音评测技术结合可有效提高学生发音准确率与流利度[17]。霍雨佳进一步指出,借助智能化语音识别工具,教师可实现学生发音检测与自动纠错,其纠错速度与精度均优于人工听评[18]。在技术与考试评价的结合方面,国内研究探索 ASR 在口语考试备考与教学质量评估中的应用价值。ASR 辅助英译汉任务显示 ASR 模式下译文产出速度提高 4%~5%,表达质量显著优化,学生接受度评分 6.3 [19]。类似研究也尝试将 ASR 反馈数据与教学质量评估挂钩,如基于 ASR 自动生成的“班级错误热力图”,可直观呈现不同音素错误频率,为教师优化教学重点与资源分配提供数据依据。

综合来看, ASR 在二语口语教学中的研究与应用呈现“从识别准确到学习价值、从技术优化到教学生态构建”的演进趋势。

3. 研究设计

本章节基于文献综述与理论分析的进行教学方案设计,在系统梳理国内外关于语音识别技术(ASR)与纠正性反馈理论的研究成果基础上,构建“识别-教学-反馈”教学模式框架。目的在于以设计形式呈现文献证据的整合结果,为高校英语口语教学提供可操作的模式参考与理论落地路径。

3.1. 研究目的与问题

研究旨在探讨数智驱动下语音识别技术在英语口语教学中的应用模式。研究聚焦“识别-教学-反馈”三个环节，分析其在课堂中的作用与局限，从而为英语教师如何在实际教学中有效使用 ASR 提供理论支撑和实践思路。研究问题包括：

- (1) ASR 在发音识别环节能为教师和学生提供怎样的支持？
- (2) 教师如何在课堂中将 ASR 识别结果嵌入教学过程？
- (3) ASR 反馈如何与教师反馈互补，共同提升学习效果？

3.2. 研究范围

研究聚焦大学英语口语课程中的基础发音教学模块，重点覆盖“音素对比训练”“单词重音感知”“基础对话语音准确性”三大内容维度，探索 ASR 技术在“课前诊断-课中训练-课后巩固”全流程中的应用路径。

研究对象确定为大学英语基础口语课程学生，大学生作为数字原生代，对移动端 ASR 工具的操作熟练度较高，且高校普遍具备公共英语课程体系与多媒体教室，ASR 工具的部署与推广无需额外投入大量硬件资源，研究成果易复制、易落地。

3.3. 研究方法

- (1) 文献分析法：梳理国内外研究，提炼 ASR 在语言学习中的应用模式与理论依据；
- (2) 案例研究法：分析典型案例，从“识别-教学-反馈”总结成效与不足；
- (3) 教学设计的法：结合文献与案例，提出 ASR 课堂教学设计方案。

3.4. 研究价值

理论价值：将纠正性反馈理论与数智技术深度结合，明确 ASR 在“明示性纠正”“重述反馈”中的功能定位，补充“技术辅助反馈”在二语语音习得理论中的研究空白；

实践价值：为高校一线英语教师提供“含具体操作细节+评估标准”的教学设计方案，解决当前 ASR 教学应用中“重流程、轻落地”的问题，帮助教师快速将技术融入课堂。

4. 教学设计课例

基于前文所形成的教学应用方案，该章节为理论驱动的教学设计示例，其作用在于将文献研究中的结论转化为具体教学流程与评估指标，展示语音识别技术在“识别-教学-反馈”链条中的可行应用方式，为后续实证研究提供设计蓝本。

4.1. 教学目标

- (1) 知识目标：掌握 3 组核心最小对比对发音规则：/θ/与/s/；/i:/与/ɪ/；/eɪ/与/e/。
- (2) 技能目标：发音准确性：课后用 ASR 工具读 20 个目标词，三组对比对音素正确率 $\geq 80\%$ ；“听音辨词”测试(10 题)正确率 $\geq 90\%$ 。
口语可懂度：双人对话用 ASR 录制，转写准确率 $\geq 95\%$ ；同伴用“简化二语可懂度量表”评分，平均 ≥ 4 分；教师随机点评评分平均 ≥ 4.2 分。
- (3) 情感目标：对 ASR 辅助练习“接受度评分 ≥ 4 分”(5 分制)；课后自主用 ASR 额外练习学生占比 $\geq 70\%$ 。

4.2. 教学对象与时间

对象：大一非英语专业学生(词汇量 2500~3000，存典型发音偏误)，40 人班；
时间：45 分钟。

4.3. 教学流程

(1)导入(5 min)

播放 3 组对比对标准音频(think/sink 等)，学生判断发音是否相同；
结合 PPT 口型图，引导感知差异，引出核心内容。

(2) 课前诊断回顾(5 min)

工具：提前 1 周让学生下载“科大讯飞口语评测 APP”，完成 20 个目标词课前诊断；
课堂：展示班级共性偏误(如/θ/发成/s/错误率 65%)，邀请 2~3 名学生分享个人错误。

(3) 讲解与示范(12 min)

结合外教口型视频、发音规则卡，拆解 3 组对比对发音要点(如/θ/舌尖位置、/i:/音长)；学生跟练，教师巡视纠错，用 ASR 实时检测抽查学生发音。

(4) 操练与即时检测(10 min)

4 人一组用操练任务单(5 个短句，词汇 ≤ 3000)练习；学生用 APP 朗读并获取错误标红与标准音对比，组内互评；网络问题备用同伴互听流程，恢复后补测。

(5) 交际任务(10 min)

两人一组“餐厅点餐”角色扮演(对话含 10 个目标音句子，如“Can I have three cakes?”)；用 APP 录制对话生成发音报告，同伴填互评表，教师随机点评 3 组。

(6) 总结与作业(3 min)

回顾“ASR 诊断→教师讲解→ASR 检测→互评”链条；
作业：完成 5 次 ASR 课后自测，录制 1 分钟含目标音的自我介绍上传班级群。

4.4. 教学评估

基于文献分析与教学设计框架，本文提出表 2 教学评估体系，以展示语音识别技术(ASR)在混合反馈教学模式中的应用逻辑。

Table 2. Teaching evaluation framework
表 2. 教学评估体系

评估维度	工具	指标	达标标准	评估时间
发音准确性(音素)	科大讯飞口语评测 APP	3 组对比对音素正确率	≥80%	课后
发音准确性(单词)	教师听音辨词测试	10 题正确率	≥90%	课堂
口语可懂度(ASR)	APP 对话录制	转写准确率	≥95%	交际任务环节
口语可懂度(同伴)	简化可懂度量表	4 项指标平均得分	≥4 分	交际任务环节
口语可懂度(教师)	简化可懂度量表	4 项指标平均得分	≥4.2 分	交际任务环节
学习情感与自主性	APP 后台数据 + 接受度问卷	自主练习比例、接受度	≥70% ≥4 分	课后 1 周

4.5. 教学效果综合分析

为弥补当前研究中实证数据的缺失，本部分明确教学设计的预期价值、可能面临的问题，并构建一套完整的效度评估方案，供未来实证研究直接使用。

4.5.1. 预期效果

基于文献研究与案例分析，本教学设计的预期效果从知识、技能、情感三个维度展开：

(1) 知识维度：课后 1 周内，85%以上学生能准确复述 3 组对比对的发音规则；

(2) 技能维度：

发音准确性：课后 ASR 自测中，目标音素正确率 $\geq 80\%$ 的学生占比从课前的 30% 提升至 75% 以上；

口语可懂度：交际任务中，ASR 对话转写准确率 $\geq 95\%$ 的小组占比达 80%，教师评分 ≥ 4.2 分的学生占比达 70%；

情感维度：课后接受度问卷中，对 ASR 辅助练习的评分 ≥ 4 分(5 分制)的学生占比 $\geq 80\%$ ，自主额外练习的学生占比 $\geq 75\%$ 。

4.5.2. 潜在挑战

(1) 超段音检测偏弱：以“ASR 基础诊断 + 教师示范/同伴模仿 + 可视化节律练习的三联策略补强。

解释性反馈不足：引入“ASR 定位 + 教师/LLM 规则解释”流程卡，统一话术与范例。

(2) 情感负荷与挫败感：设置“二次尝试 - 分层达标”机制与成长曲线可视化，辅以同伴正向回馈脚本。

(3) 设备与网络：准备离线操练材料与同伴互评备用流程，网络恢复后补测并合并成绩。

4.5.3. 效度评估

本研究的效度评估采用“量化测评与质性访谈相结合”的混合方法：除前后测与问卷外，还设计了配套的半结构式访谈提纲(见附录)，用于收集学习者对 ASR 反馈、教师协同及情感体验的深度信息。

前测工具(教学前 1 周实施，时长 20 分钟)

(1) 音素识别测试(10 题)：播放 3 组对比对的单词音频，让学生判断是否相同，考查基础辨别能力；

(2) 单词发音测试(20 题)：让学生朗读目标词，由 2 名英语专业教师独立评分(Kappa 系数 ≥ 0.85)，记录正确率；

(3) 简短对话测试(1 段)：让学生朗读“餐厅点餐”对话，ASR 转写并记录准确率。

后测工具(教学后 1 周实施，与前测结构一致，题目替换为同难度新题)：仅新增“听音辨词”测试(10 题)，考查发音规则的迁移应用能力。

5. 结论与讨论

5.1. 结论

本研究以纠正性反馈理论为核心，结合系统性文献回顾、典型案例分析及“课前 - 课中 - 课后”全流程教学设计，聚焦数智驱动下语音识别技术在大学英语口语教学中的应用。

研究明确了纠正性反馈理论与 ASR 的适配关系，即 ASR 可有效实现纠正性反馈，契合二语习得中“偏误觉察 - 修正”逻辑，却在“澄清请求”“元语言解释”“引导提示”上存在功能缺口，需教师补充，这一适配性填补了该理论在数智教学场景的应用空白。

在“识别 - 教学 - 反馈”链条中，ASR 呈现鲜明功能特征：其优势集中于段音规模化检测与教学流程优化，可解决传统教学反馈不足问题；局限则体现在超段音检测薄弱、无规则解释能力及缺乏情感支持，需通过技术协同或教师介入弥补。

“ASR + 教师 + 学习者”三位一体混合反馈模式为可行路径, ASR 承担批量检测与即时反馈, 教师负责深度解释与情感引导, 学习者通过“对比 - 互评 - 修正”形成闭环, 既适配大班教学, 又助力口语教学向智能协同转型。

此外, 研究在理论上完善了二语习得与技术融合的论证, 在实践中提供了可落地的教学设计; 未来需强化 ASR 的语境适应性与情感功能, 纳入情感变量研究, 并探索其与 LLM 等技术的融合, 以完善数智化口语教学生态。

5.2. 讨论

本研究将 ASR 应用与二语习得理论深度融合, 进一步明确了技术在理论实践中的定位。ASR 的明示性纠正与重述反馈契合课堂纠正性反馈类型, 多模态反馈触发输出假说的偏误觉察, 帮助学习者感知发音差距。呼应形式聚焦理论中兼顾效率与深度的主张, 为技术辅助反馈的理论落地提供路径。然而本文也存在以下局限: 第一, 研究对象局限, 仅围绕大学英语口语基础发音教学展开; 第二, 情感变量缺失, 研究未纳入学习者情感因素分析, ASR 的实时错误标红、正确率量化显示可能引发低水平或内向学习者的焦虑情绪, 而这一影响未被纳入模式设计, 难以全面覆盖学习者的学习体验。

在智能化学习环境中, 大学英语口语教学应从基础发音逐步过渡到高阶口语能力, 构建“语音 - 语流 - 语篇”的递进体系。阶段 A 注重音段准确性(如最小对立音、词重音与节律操练); 阶段 B 聚焦语流迁移, 通过跟读与复述发展语速与节奏; 阶段 C 提升语篇整合与交际可懂度。ASR 在 A 阶段提供高密度纠错, 在 B 阶段量化节律与停顿, 在 C 阶段与教师、同伴协同实现综合评估。LLM 可弥补 ASR 在解释性反馈和情感支持上的不足, 实现“检测 - 解释 - 练习”闭环, 并通过差错档案与成长型反馈增强学习动机。教师作为中枢, 利用学习分析可视化数据, 促进“以教促评、以评促学”的循环。未来的智能口语教学可构建多层协同框架: ASR 负责检测, LLM 提供解释与个性化支持, 学习分析连接课堂与课后, 教师与同伴承担情感与语篇指导, 共同实现“技术检测 - 智能解释 - 人本协同 - 数据驱动改进”的一体化教学路径。

未来研究可以从三个方向进一步深入: 一是开展实证实验, 结合前测后测与纵向追踪, 验证 ASR 在不同水平学习者中的长期效果; 二是探索 ASR 在超段音、交际策略和跨文化表达等高阶能力训练中的潜力; 三是关注学习者情感因素, 考察 ASR 使用过程中对学习动机、焦虑水平与自我效能感的影响。通过这些方向的扩展, ASR 在外语教学中的角色将被进一步明确与优化。

参考文献

- [1] Dai, Y. and Wu, Z. (2021) Mobile-Assisted Pronunciation Learning with Feedback from Peers and/or Automatic Speech Recognition: A Mixed-Methods Study. *Computer Assisted Language Learning*, **36**, 861-884. <https://doi.org/10.1080/09588221.2021.1952272>
- [2] Eskenazi, M. (2009) An Overview of Spoken Language Technology for Education. *Speech Communication*, **51**, 832-844. <https://doi.org/10.1016/j.specom.2009.04.005>
- [3] Inceoglu, S., Lim, H. and Chen, W. (2020) ASR for EFL Pronunciation Practice: Segmental Development and Learners' Beliefs. *The Journal of AsiaTEFL*, **17**, 824-840. <https://doi.org/10.18823/asiatefl.2020.17.3.5.824>
- [4] Wang, Y. and Young, S.S. (2014) Effectiveness of Feedback for Enhancing English Pronunciation in an ASR-Based CALL System. *Journal of Computer Assisted Learning*, **31**, 493-504. <https://doi.org/10.1111/jcal.12079>
- [5] Cucchiaroni, C. and Strik, H. (2017) Automatic Speech Recognition for Second Language Pronunciation Training. In: Kang, O. and Thomson, R.I., Eds., *The Routledge Handbook of Contemporary English Pronunciation*, Routledge, 556-569. <https://doi.org/10.4324/9781315145006-35>
- [6] Swain, M. (1985) Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in Its Development. In: Gass, S. and Madden, C., Eds., *Input in Second Language Acquisition*, Newbury House, 165-179.
- [7] Schmidt, R.W. (1990) The Role of Consciousness in Second Language Learning. *Applied Linguistics*, **11**, 129-158.

- <https://doi.org/10.1093/applin/11.2.129>
- [8] Long, M.H. (1996) The Role of the Linguistic Environment in Second Language Acquisition. In: Doughty, C.J. and Long, M.H., Eds., *Handbook of Second Language Acquisition*, Elsevier, 413-468. <https://doi.org/10.1016/b978-012589042-7/50015-3>
 - [9] Lyster, R. and Ranta, L. (1997) Corrective Feedback and Learner Uptake: Negotiation of Form in Communicative Classrooms. *Studies in Second Language Acquisition*, **19**, 37-66. <https://doi.org/10.1017/s0272263197001034>
 - [10] Thi-Nhu Ngo, T., Hao-Jan Chen, H. and Kuo-Wei Lai, K. (2023) The Effectiveness of Automatic Speech Recognition in ESL/EFL Pronunciation: A Meta-Analysis. *ReCALL*, **36**, 4-21. <https://doi.org/10.1017/s0958344023000113>
 - [11] Liakin, D., Cardoso, W. and Liakina, N. (2015) Learning L2 Pronunciation with a Mobile Speech Recognizer: French /y/. *CALICO Journal*, **32**, 1-25. <https://doi.org/10.1558/cj.v32i1.25962>
 - [12] Derwing, T.M. and Munro, M.J. (2005) Second Language Accent and Pronunciation Teaching: A Research-Based Approach. *TESOL Quarterly*, **39**, 379-397. <https://doi.org/10.2307/3588486>
 - [13] McCrocklin, S. (2019) Learners' Feedback Regarding ASR-Based Dictation Practice for Pronunciation Learning. *CALICO Journal*, **36**, 119-137. <https://doi.org/10.1558/cj.34738>
 - [14] Zhong, H., Xie, Y. and Yao, Z. (2024) Leveraging Large Language Models to Refine Automatic Feedback Generation at Articulatory Level in Computer Aided Pronunciation Training. *Interspeech 2024*, Kos Island, 1-5 September 2024, 2600-2604. <https://doi.org/10.21437/interspeech.2024-1005>
 - [15] 陈桦, 吴奎, 李景娜. 英语口语自动评测新方法——中国学生英语朗读自动评测系统[J]. 外语电化教学, 2019(1): 72-77
 - [16] 涂惠燕, 陈一宁. 基于语音识别和手机平台的英语口语发音学习系统[J]. 计算机应用与软件, 2011, 28(9): 64-66, 147.
 - [17] 于娇娜. 人工智能在英语口语教学中的应用[C]//人民教育出版社人教数字教育研究院. 数字教材·数字化教学——第四届中小学数字化教学研讨会论文案例集. 北京: 北京市通州区永乐店中学, 2019: 414-416.
 - [18] 霍雨佳. 智能化工具在英语教学中的辅助作用[J]. 基础教育课程, 2019(14): 75-77.
 - [19] 苏雯超, 招晓杏, 李德凤. 计算机辅助口译技术在翻译实践中的可用性研究[J]. 外语教学, 2022, 43(4): 69-75.

附录

本访谈旨在了解语音识别技术(ASR)在大学英语口语教学中的应用效果与反馈机制,探讨其在纠正性反馈理论框架下的人机协同作用。研究以学习者和教师为对象,通过 20~30 分钟的半结构式个别访谈,收集其在 ASR 辅助教学中的使用体验、反馈感知及教学策略,以分析 ASR 在口语教学中的优势与局限,并为构建混合反馈模式提供参考。

(一) ASR 使用总体体验

1. 你多长时间使用一次 ASR (如讯飞听见、SpeechAce 等)进行英语口语练习?
2. 在哪些类型的任务中(朗读、对话、自由表达)你觉得 ASR 最有帮助?
3. 使用 ASR 时,你的感受如何?(如兴趣、信心、焦虑、挫败感等)

(二) 对反馈类型的感知与反应

4. ASR 系统通常会给出哪些反馈形式?(如错误标记、标准读音、分数、文本对比等)
5. 收到 ASR 反馈后,你会如何处理?(立即改正、忽略、记录下来等)
6. 与教师反馈相比,你认为 ASR 反馈在准确性、实用性和清晰度方面如何?
7. 在六种纠正性反馈类型中(明示性纠正、重述、澄清请求、元语言反馈、引导提示、重复),你认为 ASR 在哪些方面表现较好,哪些方面仍有不足?

(三) 教师—技术协同

8. 你的教师在课堂上是否使用或讲解 ASR 的反馈结果?如何使用?
9. 当 ASR 的结果与教师的反馈不一致时,你更倾向于相信哪一方?为什么?
10. 你认为教师应如何帮助学生更好地理解和利用 ASR 反馈?

(四) 学习效果与反思

11. 使用 ASR 之后,你在发音、流利度或自信心方面有哪些变化?
12. 你是否觉得 ASR 促进了自主学习或自我监控能力的提升?
13. 使用过程中,你遇到了哪些困难或挑战?(如技术问题、反馈不清、情绪影响等)
14. 你希望 ASR 系统今后增加哪些功能或支持?(如解释原因、情感激励、同伴反馈等)

(五) 未来展望与建议

15. 你认为未来 ASR 或人工智能(如大语言模型)能如何更好地支持口语学习?
16. 如果可以选择,你更希望独立使用 ASR、与同伴协作使用,还是在教师指导下使用?为什么?
17. 你会给其他即将使用 ASR 进行口语学习的学生或教师什么建议?