

# Practice Research of AI Speech Tool in Junior High School English Oral Homework Based on “Integration of Teaching and Evaluation”

Yingkui Gao

Tanggu No.5 Middle School, Binhai New Area, Tianjin, 300450, China

## Abstract

This study explores the effective integration of AI-powered speech tools into junior high school English speaking assignments to establish a closed-loop practice system integrating teaching, learning, and assessment. Addressing challenges in traditional oral exercises such as delayed feedback, monotonous formats, and lack of personalization, the research conducted a semester-long teacher action study in a second-year junior high class. By systematically designing a cyclical model—“goal-oriented task design (teaching) → AI-assisted practice with instant feedback (learning/assessment) → data-driven instructional intervention (teaching)” —the study analyzed data from AI platforms, student pre-and post-test recordings, and interviews. Results demonstrated that this model significantly increased speaking practice frequency and engagement, with marked improvements in pronunciation accuracy and intonation. The AI’s real-time objective feedback enhanced students’ metacognitive awareness, fostering self-directed learning. Teachers transitioned from mere grader to learning analytics specialists and personalized instructors. This study provides actionable pathways and reflective insights for implementing the “integration of teaching, learning, and assessment” concept in technology-enhanced oral education.

## Keywords

Integration of teaching and assessment; AI-powered speech tools; Junior high school English; Oral assignments; Action research; Formative assessment

## 基于“教学评一体化”的AI语音工具在初中英语口语作业中的实践研究

高英奎

天津市滨海新区塘沽第五中学，中国·天津 300450

## 摘要

本研究旨在探索如何将人工智能（AI）语音工具有效整合于初中英语口语作业中，以构建一个“教学评一体化”的实践闭环。针对传统口语作业存在的反馈滞后、形式单一、难以个性化等困境，研究以某初二年级一个班级为对象，开展了为期一学期的教师行动研究。通过系统设计“目标导向的作业布置（教）—AI辅助练习与即时反馈（学/评）—数据驱动的教学干预（教）”的循环模式，并收集AI平台数据、学生前后测录音及访谈等资料进行分析。研究发现，该模式能显著提升学生口语练习的频次与积极性，在语音、语调的准确性上进步明显；同时，AI提供的即时、客观反馈增强了学生的元认知意识，促使其成为更自主的学习者。教师角色则从单纯的批改者转变为学情分析师与个性化指导者。本研究为在技术环境下落实“教学评一体化”理念、革新口语教学提供了具体的实践路径与反思。

## 关键词

教学评一体化；AI语音工具；初中英语；口语作业；行动研究；形成性评价

## 1 引言

随着《义务教育英语课程标准（2022年版）》对学生口语交际能力要求的日益提高，以及“教—学—评”一体化设计理念的强调，初中英语口语教学面临着新的挑战与机遇。然而，在“大班额”的教学现实中，口语教学常陷入困

境：课堂时间有限，难以给予每位学生充分练习与指导；课后口语作业形同虚设，因其“难布置、难监督、难反馈”，往往流于简单的朗读背诵，教师无法获知完成过程与质量，评价严重滞后，导致“教”“学”“评”三者脱节。

近年来，基于人工智能语音识别与评测技术的教育应用（如“口语100”、“E听说”、“流利说”等）为破解这一难题提供了技术可能。这些工具能对学生的跟读、模仿、自由表达进行即时评分和智能诊断，提供发音、流利度、完整度等多维度反馈。然而，当前实践中，这些工具常被师生

【作者简介】高英奎（1971—），男，中国天津人，本科，高级教师，从事初中英语研究。

作为零散的“练习器”或“评分机”使用，未能深度融入系统化的教学设计与评价体系，其产生的海量学习数据也未有效用于教学决策。

因此，本研究尝试超越对技术工具效能的简单验证，聚焦于如何以“教学评一体化”理念为统领，将AI语音工具系统化、结构化地嵌入初中英语口语作业的全流程，构建一个可持续、可操作的教学模式。本研究拟解决以下核心问题：

(1) 基于“教学评一体化”理论，如何设计并实施一个整合AI语音工具的口语作业闭环流程？(2) 该实践对学生的口语产出能力（语音、语调、流利度）有何具体影响？(3) 该实践如何影响学生的口语学习动机、策略及自主学习能力？

## 2 文献综述与理论框架

### 2.1 “教学评一体化”的内涵与发展

“教学评一体化”源于形成性评价（Formative Assessment）与“为了学习的评价”（Assessment for Learning）理念，其核心主张是评价并非教学的终点，而是贯穿于教学全过程，旨在促进学习、改进教学（Black & Wiliam, 1998）。它强调教学目标、学习活动与评价任务的高度一致性，评价结果被即时用于调整教学策略与满足学生需求（王蕾等, 2019）。在本研究中，“一体化”体现为：口语作业的目标（教）是评价设计的依据，AI工具实施的评价（评）驱动和优化学生的练习过程（学），而评价产生的数据又反过来指导教师下一步的教学决策（教）。

### 2.2 AI技术在语言口语教学中的应用

大量研究证实了自动语音识别（ASR）与评测系统在提高发音准确性、提供即时反馈方面的有效性（Li et al., 2021）。然而，既有研究多聚焦于工具本身的技术精准度或短期使用效果，较少从课程与教学论的宏观视角，探讨如何将其有机整合进完整的教学周期，并审视其对师生角色、学习文化产生的深层影响（Chen et al., 2020）。本研究试图填补这一空白，将AI工具定位为“教学评一体化”闭环中的关键赋能者，而非孤立的技术插件。

### 2.3 本研究的理论框架

基于以上综述，本研究构建了如下理论框架（图1）：以清晰的口语能力目标为起点，AI语音工具作为核心的“评价中枢”，在“学”的环节提供即时、个性化的反馈，支持学生的刻意练习；同时，该工具作为“数据中枢”，将学习过程数字化，为教师的“教”提供精准的学情诊断依据。教师通过解读数据，进行共性问题的课堂补救或个性化辅导，从而形成一个动态、持续改进的“目标—练习/反馈—教学调整”闭环。这个闭环使得评价真正服务于学与教，实现一体化设计。

## 3 研究设计与方法

### 3.1 研究情境与参与者

本研究在天津市滨海新区塘沽第五中学开展。研究对象为初二年级一个由39名学生组成的自然班，学生英语水平呈正态分布。研究周期为一个学期（约16周）。选用“E

听说”平台作为AI工具，因其功能（跟读模仿、角色对话、情景问答）与初中教材同步，且后台数据分析报告较为详细。

### 3.2 研究方法：教师行动研究

采用教师行动研究范式，遵循“计划—行动—观察—反思”的螺旋式循环（Kemmis & McTaggart, 1988）。研究者本人即授课教师，在教学实践中进行研究，通过两轮循环不断完善教学模式。

### 3.3 “教学评一体化”口语作业模式设计

第一轮（第1-8周）：初步构建与调试。

目标层（教）：每周根据单元主题（如“健康生活”）与语法重点（如一般过去时），设定1-2个具体的口语微技能目标（如“能使用过去时态描述上周活动，动词过去式发音基本正确”）。

实施层（学/评）：每周布置2次AI作业，包含：①基础跟读：针对目标词汇、句型的语音校准；②情景模仿：与AI虚拟角色进行目标句型对话；③自由表达：围绕主题进行1-2分钟的录音陈述。学生完成即获AI评分（百分制）及逐句发音诊断（绿/黄/红色标识）。

反馈与调整层（教/评）：教师每周查看平台“班级报告”，关注“共性错误排行榜”（如/th/发音、动词过去式-ed结尾），在下一节课前5分钟进行集中纠音与讲解。

第二轮（第9-16周）：模式深化与优化。

基于第一轮反思，进行优化：①引入“挑战任务”，设置稍高于平均水平的评分门槛，激励学生反复练习追求“三星”；②要求学生每周提交一次“自我反思笔记”（用中文），记录自己最大的进步与仍困惑的音；③教师根据平台“个体轨迹图”，识别出“持续低迷”或“突然下滑”的学生，进行课下面谈或简短语音指导。

### 3.4 数据收集与分析

定量数据：AI平台导出的全班前后测（教材配套标准化口语测试）平均分、每周作业平均完成率与平均分趋势、共性错误数据统计。

定性数据：

学生录音样本：选取高、中、低水平学生各3名，对比其学期初与学期末同一类型任务的录音。

半结构化访谈：学期末，从三个水平层中各随机选取2名学生进行访谈，了解其感受与策略变化。

教师反思日志：每周记录实施情况、观察到的学生反应及教学调整思考。

数据分析：定量数据进行描述性统计与趋势分析；定性数据采用主题分析法，对访谈和反思日志进行编码，提炼主题。

## 4 研究过程与发现

### 4.1 第一轮循环：发现问题与初步适应

初期，学生表现出浓厚兴趣，作业完成率达98%。但教师反思日志显示问题：(1) 部分学生追求“刷分”，同一任务重复数十次直至满分，但并未真正关注错误原因；

“班级报告”仅显示错误单词，教师难以理解错误根

源（如是因为清浊辅音不分还是重音不对）；

（3）部分后进生因得分持续偏低产生挫败感。

调整措施：在课堂中示范如何解读 AI 反馈（如查看波形图对比），强调“听懂反馈比高分更重要”；对后进生降低单项任务要求，并安排小组长进行同伴鼓励。

## 4.2 第二轮循环：深化应用与模式固化

引入“挑战任务”和“反思笔记”后，学习行为发生质变。学生 S1 在笔记中写道：“这周我发现我的 /æ/ 和 /e/ 老是分不清，AI 总是标黄。我看了平台里的‘发音秘籍’视频，知道了嘴巴张开的大小不同，现在好多了。”教师根据个体轨迹，发现学生 S2 在“自由表达”任务上分数骤降，面谈后发现是因话题“描述一次竞赛经历”对其有难度，随即提供了关键词汇支架和句型模板。

稳定的操作流程最终形成（见图 2，此处应有一流程图，描述为：教师设定目标与任务 → 学生 AI 练习获得即时反馈 → 平台生成个体与班级报告 → 教师分析数据 → 实施课堂集中补救或个别指导 → 调整下一轮目标与任务）。

## 4.3 研究发现

针对研究问题一（模式构建）：本研究成功构建了一个以数据为驱动的口语作业“教学评一体化”闭环模式。其核心环节在于：目标的具体化、评价的自动化与过程化、以及数据解读的教学化。

针对研究问题二（对能力的影响）：

定量证据：班级口语后测平均分较前测提升 15.6%。共性错误数据中，诸如“think”读作“sink”的 /th/ 发音错误率从期初的 41% 下降至期末的 18%。学生每周在 AI 作业上的平均有效练习时间从最初的 8 分钟增加到后期的 15 分钟。

定性证据：对比学生录音样本发现，高水平学生 L1 的语调更为丰富；中等学生 M1 的单词尾音（如复数 s、过去式 ed）清晰度显著提高；后进生 L1 从几乎无法完成完整句子到能够产出基本达意的短语音频。

针对研究问题三（对情感与策略的影响）：

访谈分析揭示出三个主题：

从“怕开口”到“敢开口”：“反正 AI 不会笑话我，说错了就重来。”（学生 A）

元认知意识的觉醒：“以前不知道自己哪里读得不对，现在看那个红色的词就知道了，会专门去练它。”（学生 B）

学习自主性的提升：“我会自己去找‘提升训练’里的专项练习了，感觉像打游戏过关一样。”（学生 C）

教师反思日志也指出，学生开始从被动完成作业转向主动寻求诊断和提升。

## 5 讨论与反思

### 5.1 综合讨论

本研究证实，当 AI 工具被系统地置于“教学评一体化”框架下时，其价值远超“电子评分员”。它实现了评价的“高频化”与“隐形化”，使形成性评价在日常作业中得以常态化实施。更重要的是，它推动了师生角色的深刻转变：学生

成为拥有“数字画像”的自主学习管理者，而教师则进化为基于证据进行决策的“学情分析师”与“个性化学习设计师”。这与新时代教育评价改革的方向高度契合。

### 5.2 实践启示

对一线教师的启示在于：

始于目标，终于发展：技术应用的首要前提是明晰的教学目标，所有活动与评价都应对准目标，避免为用技术而用技术。

读懂数据，智慧干预：教师需提升数据素养，学会从平台报告中识别共性问题与个体异常，并将此转化为精准的教学行动。

融合人机，情感关怀：AI 反馈虽客观但冰冷，必须与教师的鼓励、同伴的分享以及面向困难学生的情感支持相结合，构建有温度的技术增强型学习环境。

### 5.3 研究局限性

本研究为单组行动研究，缺乏严格的对照组，结论的普适性受到一定限制。研究周期仅为一个学期，长期效果有待观察。此外，AI 工具对口语交际中更复杂的互动能力、思辨能力的培养作用尚不明确。

### 5.4 未来展望

未来研究可探索：如何将 AI 对“准确性”的评价与教师、同伴对“内容深度”、“交际得体性”的质性评价相结合，构建更全面的口语评价体系；如何利用 AI 数据支持更大尺度的差异化教学；以及如何设计基于 AI 的创造性口语任务，推动学生从“模仿输出”走向“意义创造”。

## 6 结语

本研究通过一学期的行动研究，初步探索并验证了一个基于“教学评一体化”理念、整合 AI 语音工具的初中英语口语作业模式。该模式通过技术赋能，实现了口语练习的即时反馈、学习过程的精细化管理以及教学决策的数据驱动，有效促进了学生口语能力与学习自主性的协同发展。实践证明，技术的成功应用，关键在于以先进的教育理念为引领，进行精心的教学设计与持续的专业反思。让技术服务于“人的发展”，才是教育数字化转型的根本要义。

### 参考文献

- [1] Black P, Wiliam D. Assessment and classroom learning [J]. Assessment in Education: Principles, Policy & Practice, 1998,5 (1):7-74.
- [2] Chen H, Wang Y, Zhao S. A review of research on AI-powered speaking assessment [J]. Language Learning & Technology, 2020,24 (3):1-19.
- [3] Kemmis S, McTaggart R. The action research planner [M]. Geelong: Deakin University Press, 1988.
- [4] Li K, Qian X, Wang S. The impact of ASR-based pronunciation training on EFL learners' speech intelligibility [J]. Computer Assisted Language Learning, 2021:1-25.
- [5] 王蕾, 李亮. 推动核心素养背景下英语课堂教—学—评一体化: 意义、理论与方法 [J]. 课程·教材·教法, 2019,39 (5):114-120.