

# 人机协同视域下英语口语译课程学习评价生态的重构

杜雨欣<sup>1</sup>, 肖展鹏<sup>2,3\*</sup>

<sup>1</sup>广西民族大学相思湖学院外国语言文学学院, 广西 南宁

<sup>2</sup>广西国际商务职业技术学院人工智能学院, 广西 南宁

<sup>3</sup>马来西亚理科大学计算机与科学学院, 马来西亚 槟城

收稿日期: 2025年12月1日; 录用日期: 2025年12月28日; 发布日期: 2026年1月5日

## 摘要

在以大语言模型(Large Language Model, LLM)和自动语音识别(Automatic Speech Recognition, ASR)为代表的新一代人工智能技术迅速发展的背景下, 翻译行业生态与口译人才能力结构正发生深刻变化, 高校英语口语译课程的学习评价体系面临前所未有的重构压力。现有研究多聚焦于智能技术赋能教学模式与课堂活动的设计, 对于人工智能条件下学习评价范式如何整体转型、评价目标-主体-证据关系如何在“人机协同”视域下重构关注不足。本文在系统梳理教育评价范式演进与学习导向评价理论的基础上, 结合人工智能介入教育的最新发展, 揭示当前高校英语口语译课程学习评价在目标与标准偏狭、主体结构单一、证据生成碎片化等方面的结构性困境。在此基础上, 引入教育生态学与人机协同学习理论, 提出以“学习证据链”为关键中介的英语口语译课程学习评价生态重构框架, 细化学习证据链的构成要素、数据来源、分析方法与呈现路径, 并构建基于LLM与ASR的技术实现流程。进一步, 以某高校英语专业《英汉交替传译》课程为例, 构建基于学习证据链的人机协同评价教学情境, 展示师-生-机多主体协同的互动方式, 讨论其在课堂实践中的可能成效与潜在挑战。最后, 从技术使用偏离教育初衷的风险、数据隐私与学术诚信、制度规范与素养培育等维度, 探讨人机协同学习评价的伦理边界与治理路径。研究旨在为人工智能时代高校英语口语译课程学习评价的理论探索与实践改革提供一个兼具规范性与操作性的分析框架。

## 关键词

人工智能, 英语口语译课程, 学习评价, 人机协同, 学习证据链

## Reconstructing the Learning Assessment Ecology of English Interpreting Courses from a Human-AI Collaborative Perspective

Yuxin Du<sup>1</sup>, Zhanpeng Xiao<sup>2,3\*</sup>

\*通讯作者。

文章引用: 杜雨欣, 肖展鹏. 人机协同视域下英语口语译课程学习评价生态的重构[J]. 教育进展, 2026, 16(1): 373-378.  
DOI: 10.12677/ae.2026.161051

<sup>1</sup>School of Foreign Languages and Literature, Xiangsihu College of Guangxi Minzu University, Nanning Guangxi

<sup>2</sup>School of Artificial Intelligence, Guangxi International Business Vocational College, Nanning Guangxi

<sup>3</sup>School of Computer Sciences, Universiti Sains Malaysia, Penang, Malaysia

Received: December 1, 2025; accepted: December 28, 2025; published: January 5, 2026

## Abstract

With the rapid advancement of a new generation of artificial intelligence technologies represented by Large Language Models (LLMs) and Automatic Speech Recognition (ASR), the ecology of the translation industry and the competence structure of interpreters are undergoing profound changes, placing unprecedented pressure on the reconstruction of learning assessment systems in university English interpreting courses. Existing studies primarily focus on how intelligent technologies empower teaching models and classroom activities, while paying insufficient attention to how assessment paradigms should be transformed as a whole under AI conditions, and how the relationships among assessment goals, agents and evidence can be reconfigured from a human-AI collaborative perspective. Drawing on a systematic review of the evolution of educational evaluation paradigms and theories of learning-oriented assessment, and incorporating recent developments of AI in education, this paper identifies several structural dilemmas in current learning assessment for English interpreting courses, including narrow assessment goals and standards, a single and teacher-dominated agent structure, and fragmented generation of assessment evidence. Building on educational ecology and human-AI collaborative learning theories, it then proposes a reconstruction framework for the learning-assessment ecology of English interpreting courses with the “learning evidence chain” as a key mediator, and further specifies its components, data sources, analytical methods and modes of presentation, together with a technical pipeline based on LLMs and ASR. Taking a university English interpreting course (Chinese-English consecutive interpreting) as an illustrative case, the paper designs a human-AI collaborative assessment scenario grounded in the learning evidence chain, and discusses the potential benefits and challenges of multi-agent collaboration among teachers, students and machines in classroom practice. Finally, from the perspectives of risks associated with the misuse and overreliance on technology, data privacy and academic integrity, as well as institutional regulation and literacy cultivation, it explores the ethical boundaries and governance pathways of human-AI collaborative learning assessment. The study seeks to provide a normative yet operational analytical framework for theoretical inquiry and practical reform of learning assessment in English interpreting courses in the age of artificial intelligence.

## Keywords

Artificial Intelligence, English Interpreting Courses, Learning Assessment, Human-AI Collaboration, Learning Evidence Chain

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

新一轮以生成式人工智能为核心的技术革命,正在重塑全球语言服务行业的生产流程与价值链条。智能同声传译、机器翻译后编辑、人机协同翻译等新业态不断涌现,使译员的工作任务由线性语言转换转向复杂的人机协作与多模态信息整合。在此背景下,高校英语口语课程的人才培养目标已难以停留在

“具备扎实语言基础与口译技能”的传统表述,而需要面向语言能力、跨文化胜任力、信息素养与技术素养等维度的综合发展[1]。与此同时,以LLM和ASR为代表的人工智能技术逐步进入课堂,学习过程和学习结果日益呈现出数据化、可视化和可追踪的特征,为学习评价体系的重构提供了前所未有的技术条件。

然而,与教学模式和课堂活动创新的热度相比,高校英语口译课程学习评价体系的改革明显滞后。当前不少课程仍以期末考试和教师主观评分为主,评价目标高度聚焦于译文准确性和语音流利度,评价主体集中于教师一方,评价证据主要来自少数几次高风险任务和教师的印象性判断[2]。这种结果本位、甄别导向的评价范式不仅弱化了评价的学习支持功能,也难以充分吸纳人工智能技术带来的证据资源与分析能力。已有研究在一定程度上注意到了智能技术在学习分析、自动评分和形成性评价中的潜力,但对于人工智能条件下学习评价范式整体转型、特别是人机协同视域下“教-学-评”关系如何重构的问题,相关讨论仍相对有限。

基于此,本文试图将研究视角从“如何利用LLM与ASR构建新型教学模式”上移至“人工智能时代英语口译课程学习评价如何实现系统转型与重构”,重点回答三个相互关联的问题:其一,在学习评价范式演进的宏观脉络中,当前高校英语口译课程学习评价存在哪些具有结构性和时代性的困境?其二,以人机协同和教育生态学为分析视角,如何在评价目标、主体、证据与功能等维度上重构英语口译课程学习评价生态?其三,在重构过程中,应当如何处理技术使用偏离教育初衷、数据隐私与学术诚信等伦理问题,构建相应的治理路径?本文采用文献分析与规范性理论研究的方法,对学习评价理论、人工智能教育应用与翻译教育研究进行综合梳理,并在此基础上提出“人机协同学习评价生态”的分析框架,旨在为人工智能时代高校英语口译课程学习评价的理论探索与实践改革提供学理支撑与结构性思路。

## 2. 学习评价范式的根基与演进

学习评价不仅是一套技术程序,更是对“何为学习、何为发展”的价值表达。在考试测量传统中,评价主要被视为通过标准化工具测量学生成绩、区分水平高低的手段,评价对象以学习结果为主,强调信度、效度与客观性。随着课程改革与教学研究的推进,评价的视角逐渐扩展至课堂教学过程本身,教学评价开始关注教师教法与教材实施效果,并通过评价结果改进教学设计[3]。建构主义学习观与“以学生为中心”理念的兴起,使学习评价范式进一步从“围绕教学”转向“围绕学习”:学习导向评价将学习者的理解建构、策略调整与自主反思视作评价的核心目标,评价开始被理解为“学习的一部分”,而非仅仅是学习之后的“检验”。

英语口译教育较晚嵌入这一范式演进轨迹,其实践长期深受“高风险情境”“高标准输出”的行业逻辑影响。在传统专业培养模式下,口译课程评价高度集中于期末考试、模拟会议等高难度任务,强调译文准确性、完整性和语言得体,形成以结果为中心、以甄别与选拔为主要功能的评价结构。即便在课程评价逐步引入口译实验、课堂演练和实践汇报之后,评价关注的焦点仍多停留在最终表现质量本身,学生在任务准备、资料整合、策略运用与情绪调节等过程性努力难以获得系统呈现与价值确认[4]。人工智能技术的加入,一方面为捕捉学习过程数据提供了技术可能,另一方面也提出了新的问题:如果仍沿用以往“终结性表现-教师打分”的逻辑,只是在评分环节增加部分智能工具,那么评价范式的深层结构并未改变,技术也就难以真正实现“赋能学习”的目标。

进入人工智能时代,尤其是生成式人工智能和语音识别技术逐渐成熟并进入课堂实践之后,学习评价在证据形态、实施方式与权力结构方面出现了显著的新趋势。一方面,学习过程日益被各类平台与工具“编码”下来,语音信号、文本轨迹、交互记录、任务完成时间等多源数据构成前所未有的“学习痕迹”,为基于证据链的能力推断提供了条件;另一方面,自动评分、智能诊断和学习分析仪表盘等工具

在提高评价效率与精度的同时, 也在悄然改变“谁在评价”“谁有权解释评价结果”的传统格局[5]。如果在理论与制度上缺乏充分准备, 容易滑向“技术决定论”或对数据的单一化迷信, 使技术工具在教育实践中“喧宾夺主”, 并加剧评价功能的失衡。因此, 在人工智能情境下讨论英语口语译课程学习评价, 必须同时看到范式演进的机遇与风险, 从结构上重新审视评价目标、主体与证据之间的关系。

### 3. 英语口语译课程学习评价的结构性困境

在学习导向评价成为国际教育改革重要方向的背景下, 高校英语口语译课程学习评价表层上呈现出一定的多元化趋势: 人才培养方案中频繁出现“综合素养”“跨文化能力”“信息素养”等术语, 课程大纲中写入了“过程性评价”“形成性评价”等条目, 一些课程还开始尝试引入口译实验、课堂展示与实践汇报等多元任务。但在更深的层面上, 评价范式仍深受传统“结果本位-甄别导向”逻辑支配, 具体体现在以下几个方面。

① 评价目标与标准偏狭, 与人才培养要求错位。尽管文件层面普遍强调口译人才应具备语言转换、信息整合、跨文化沟通、应急决策和技术应用等多维能力, 但在具体评价实践中, 标准往往集中于译文是否准确、表达是否流利、术语是否规范等易于观测和打分的结果指标。学生在选材准备、提纲构建、策略选择、情绪管理和与技术工具协同等环节的努力与成长, 较少进入正式评价视野。评价于是固化为对少数关键表现性任务的终局判定, 难以引导学生形成对自身整体学习状态的深度理解, 也难以支撑能力结构的全面发展。

② 评价主体结构单一, 权力高度集中于教师。多数口译课程仍采取教师独立设计任务、独立评分、独立解释结果的模式, 学生主要在成绩公布后被动接受评价, 很少有机会理解评价标准的生成逻辑, 更谈不上参与评价规则制定或对评价结果进行合理辨析。人工智能尚未介入时, 这一结构极易造成评价权力失衡与学生评价素养 deficit; 在 LLM 与 ASR 等技术广泛进入课堂之后, 如果仍坚持“教师独评”的旧范式, 就会产生一种更隐蔽的不平衡: 技术通过自动计分和智能诊断事实上参与了评价, 却缺乏透明的角色定位与使用规范; 学生虽能从系统获得即时反馈, 却仍被排除在评价话语权之外, 在教师与算法的双重裁决中被动应对。这不仅削弱了学习评价的教育性, 也埋下评价公正与责任归属方面的隐患。

③ 评价证据生成碎片化, 学习过程长期“缺席”。传统口译课程主要依靠期末考试、阶段模拟、课堂展示等少量高风险任务获取评价证据, 教师日常观察形成的大量印象性判断难以系统化, 学生在课前准备、课后反思和自主练习中产生的语音、文本与思维轨迹更少被纳入正式证据体系。结果是, 评价证据既低频又静态, 只能给出若干时间切片上的“能力快照”, 难以呈现学习过程的动态变化与能力发展的趋势。即便在技术条件已经可以支持语音转写、过程追踪和数据可视化的今天, 如果评价观念和制度仍固守“孤立评分事件”的组织方式, 大量过程性证据依旧在实践中缺位, 学习评价难以从“看结果”走向“看过程”“看变化”。

在人工智能已深度介入语言学习的情境下, 上述目标偏狭、主体单一和证据碎片化三方面的问题, 不再只是效率或技术问题, 而是范式层面的掣肘。如果不能在理论上对其进行有意识的更新与调整, 人工智能的引入就极易沦为旧有评价框架上的“技术贴皮”, 既放大既有范式的局限, 又难以真正释放智能技术在支持学习发展方面的潜力。

### 4. 人机协同视域下英语口语译课程学习评价生态的重构路径

要在人工智能时代构建新的学习评价范式, 关键在于增添若干技术工具, 而在于在“人机协同”视域下整体重构评价生态, 使评价在课程系统中发挥连接教学设计、学习实践与能力发展的枢纽功能。

表 1 呈现了评价范式的主要转换方向。



Table 1. Evaluation paradigm conversion table  
表 1. 评价范式转换表

维度	传统口译课程评价	人机协同学习评价
评价理念	结果本位、甄别导向	学习本位、发展导向
评价目标与内容	聚焦语言准确性与流利	兼顾语言、策略、跨文化与技术素养
评价主体	教师垄断话语权	师 - 生 - 机多主体协同
评价证据与工具	低频、静态、零散	多源、连续、可追踪的证据链
评价过程与反馈方式	终结性、一锤定音	过程性、多轮反馈与调节
评价功能与伦理关注	选拔、认证为主	诊断、促进与甄别并重

① 以学习证据链为核心重构评价观。与以往围绕单次成绩的“终点评价”不同，人机协同学习评价强调围绕能力发展目标，将任务设计、学习表现、过程记录、模式识别与反馈调节等环节有机串联，构成“学习证据链”。在英语口语译课程中，ASR 可以将学生口译输出转化为可检索的文本，记录语速、停顿和重音等语音特征；LLM 可以在预设维度上对语言准确性、信息完整性、语篇连贯性等提供初步诊断；学习管理平台可以汇聚多次任务的表现数据，形成时间序列的过程性证据。教师与学生在此基础上围绕证据展开对话，实现从“事后判定”向“过程介入”、从“单点量化”向“结构理解”的转变。

② 构建师 - 生 - 机多主体协同结构。人机协同学习评价并不意味着技术取代人，而是在评价链条中为不同主体赋予合理角色。教师是评价价值取向与专业判断的最终守护者，在能力目标设定、指标体系建构和高风险决策中承担核心职责；学生通过自评、互评与对技术反馈的再解读，逐步形成理解标准、运用标准与反思自我的能力，成为评价过程的积极参与者；智能系统则被明确定位为数据采集、模式识别与诊断性分析的工具性主体，为师生判断提供透明、可追溯的证据信息，而不是在关键情境中取代人的裁决。通过制度化安排三方在不同环节的参与方式与权重，可以缓解技术工具“喧宾夺主”的风险，提升评价的教育性。

③ 拓展评价指标维度并重塑评价功能定位。在内容层面，评价指标应从单一的语言输出质量拓展为涵盖语言准确与得体、信息完整与逻辑条理、即时策略与问题解决、跨文化敏感性与适应力、信息素养与 AI 使用素养等多维结构。智能技术可以在部分维度提供精细化数据支持，但何者为核心能力、何者为关键发展潜力，仍需在专业共同体中形成共识。在功能层面，评价需要从“主要用于甄别与选拔”的单一取向，转为“诊断 - 调节 - 促进 - 甄别”多功能的有机统一：形成性与发展性评价成为常态，终结性评分则置于这一大框架之中。

由此可见，人机协同学习评价并非对传统模式的简单“技术增强”，而是在理念、结构与功能上进行系统性重构，使英语口语译课程学习评价更好地服务于人工智能时代复合型口译人才的培养。

5. 人机协同学习评价的伦理边界与治理路径

在肯定人机协同学习评价潜力的同时，必须正视其可能带来的伦理与治理问题。首先，技术使用偏离教育初衷与评价功能失衡的风险不容忽视。若评价目标被“可量化性”所绑架，技术工具便可能从服务教育的手段转变为在无形中左右教育决策的主导力量，评价活动会从面向学习者整体发展的教育实践滑向对有限指标的机械追逐和排名。其次，数据隐私与学术诚信问题日益凸显。人机协同学习评价高度依赖对学生语音、文本与行为轨迹的持续采集和分析，学习数据一旦在收集、存储和使用过程中缺乏明确规范与边界，极易引发隐私泄露与滥用风险。同时，生成式人工智能的参与模糊了“学习成果”的归

属边界, 如果评价机制缺乏相应识别与引导, 可能损害评价结果的真实性。

因此, 有必要在制度层面建立清晰的数据治理规则, 明确学习数据的所有权归属、存储期限、访问权限与用途范围, 落实知情同意、最小必要与去标识化等原则; 在课程层面, 将“合理使用 AI 工具”“维护学术诚信”纳入明确的学习目标, 通过案例分析和情境演练帮助学生理解不同使用方式的边界; 在教师专业发展层面, 通过培训与学术共同体交流, 提升教师对人工智能原理、优势与局限的理解能力, 避免陷入技术崇拜或简单拒斥。只有在制度规范与素养培育协同推进的前提下, 人机协同学习评价的生态重构才能在保障教育公正与学生权利的基础上稳步推进。

## 6. 结语

人工智能技术的迅猛发展, 使学习过程可以以前所未有的方式被记录、分析与呈现, 为高校英语口译课程学习评价的范式转换提供了现实契机。本文在梳理教育评价范式演进与学习导向评价理论的基础上, 指出当前高校英语口译课程学习评价在目标与标准僵化、主体结构单一和证据生成碎片化等方面存在结构性困境, 强调在人工智能广泛介入的情境下, 若不对既有评价观念与制度安排进行有意识的更新与调整, 技术加持可能加剧评价权力失衡与功能偏移。围绕评价生态重构的问题, 本文提出以“学习证据链”为关键中介的人机协同学习评价生态重构框架, 从评价观、主体结构以及指标与功能三个维度阐释了人工智能时代英语口译课程学习评价的重构路径, 并进一步从技术使用偏离教育初衷的风险、数据隐私与学术诚信、制度规范与素养培育等方面探讨了人机协同学习评价的伦理边界与治理要求。

从理论层面看, 本文尝试将教育评价理论、教育生态学、人机协同学习理论与翻译教育研究相整合, 为理解人工智能时代英语口译课程学习评价的生态重构提供了一个具有解释力的分析框架; 从实践层面看, 本文为高校在口译课程中设计学习导向、技术赋能的人机协同评价体系提供了若干可资借鉴的原则与路径。当然, 本文主要采用规范性分析方法, 尚缺乏大样本、长周期实证数据对所提框架进行系统检验。未来研究可在不同类型高校和课程情境中对人机协同学习评价生态的运行机制与成效展开纵深考察, 并进一步探索该框架在其他外语课程乃至跨学科教学中的迁移与适配性, 以不断丰富和修正相关理论。

## 基金项目

外教社 2024 年度全国高校外语教学科研项目“人工智能时代英语口译课程教学模式的破与立”(全国高校 202412636GX); 广西民族大学相思湖学院 2024 年度校级教育教学改革工程立项一般项目“人工智能技术驱动下的英语专业英语口译课程教学改革探索”(2024JGYB59); 广西国际商务职业技术学院 2024 年度校级一般科研项目“RCEP 背景下广西高职学生数智化思维与素养提升策略”(GSB2024018)。

## 参考文献

- [1] 张莉. 数字赋能英语读写教学评价: 概念、路径和理据[J]. 创新教育研究, 2025, 13(7): 77-83. <https://doi.org/10.12677/ces.2025.137499>
- [2] 袁磊, 徐济远, 刘沃奇. 数智教育生态下人机协同教学范式转型[J]. 开放教育研究, 2025, 31(2): 108-117.
- [3] Yang, S.J.H., Ogata, H., Matsui, T. and Chen, N. (2021) Human-centered Artificial Intelligence in Education: Seeing the Invisible through the Visible. *Computers and Education: Artificial Intelligence*, **2**, 100008. <https://doi.org/10.1016/j.caeai.2021.100008>
- [4] Brusilovsky, P. (2024) AI in Education, Learner Control, and Human-Ai Collaboration. *International Journal of Artificial Intelligence in Education*, **34**, 122-135. <https://doi.org/10.1007/s40593-023-00356-z>
- [5] 红鸽. 人工智能在中小学英语教学中的影响: 现状、挑战与建议[J]. 国际教育学, 2025, 7(3): 68-72.