

一项基于 AI 的多语同步口语平台在新加坡小学的应用探索

“Learn”: An AI-Based Synchronous Multilingual Oral Learning Platform for Primary Schools in Singapore

胡月宝, 林庆荣

南洋理工大学国立教育学院亚洲语文与文化学部

guatpoh.aw@nie.edu.sg

khengyong.lim@nie.edu.sg

【摘要】 随着人工智能,尤其是机器学习在教育领域的快速发展,智能教材的设计逐渐转向以学习者为中心的个性化方向。“乐言”是由新加坡理工大学与南洋理工大学联合开发的图画驱动 AI 口语学习平台,专为小学低年级学生设计,支持华语、马来语和淡米尔语的多语同步学习。平台融合自然语言处理、语音识别与个性化反馈机制,提供涵盖发音、词汇与口语对话的图像引导式任务练习。本文围绕“多语同步开发”理念,介绍平台的构建路径与三语协同机制,分析其在语言对等设计、智能反馈与跨语适配方面的关键策略,并结合试点应用,探讨系统面临的挑战与优化方向。

【关键词】 教育人工智能;机器学习;自然语言处理;多语学习;语音交互

Abstract: *LEARN is a picture-based AI oral learning platform co-developed by the Singapore Institute of Technology and Nanyang Technological University. Designed for lower primary students, the platform supports synchronous multilingual learning in Chinese, Malay, and Tamil. It integrates natural language processing, speech recognition, and personalized feedback to provide contextualized oral practice through image-guided tasks. Centered on the concept of multilingual synchronous development, this study outlines the platform's design process and collaborative framework across languages. Key features such as content alignment, intelligent feedback, and cross-linguistic adaptation are examined. Preliminary pilot use highlights both the platform's effectiveness and areas for improvement, offering insights into AI-driven approaches to oral language learning in multilingual primary education settings.*

Keywords: AI in education, machine learning, natural language processing, multilingual learning, speech interaction

1. 研究背景与问题陈述

近年来,人工智能技术,特别是机器学习与自然语言处理(NLP),在教育领域的应用迅速扩展,推动了智能教材与个性化学习工具的发展(Chen et al., 2020; Holmes et al., 2019)。深度学习模型在语音识别、语言生成和语义理解方面的突破,使得构建互动性强、响应灵敏的语言学习平台成为可能(Zawacki-Richter et al., 2019)。语言模型(如Transformer架构)和语音识别系统的不断进化,为提升儿童语言输入质量和学习反馈的即时性提供了技术基础(VanLehn, 2006)。在低龄儿童语言教育场景中,机器学习的自适应能力尤其适合处理初学者多变的语言输入,提供可调整的学习内容与反馈路径(Chen et al.,

2020)。同时，语音转文本（ASR）、语义理解（NLU）与对话生成（NLG）等关键技术的结合，使 AI 系统能够理解学生表达、识别语言偏误，并提供即时响应，从而构建具备互动、诊断与支持功能的智能会话环境（VanLehn, 2006）。机器学习技术，尤其是具备自适应与迁移能力的语言模型，能够针对本地学生的语言背景设计定制化的输入输出路径，从而弥补语言环境的不足（Zawacki-Richter et al., 2019）。通过对大量新加坡本地化语料（如校本语言表达、口语练习答案、生活化用语）的训练，AI 系统可以模仿真实语境中的语言使用方式，增强语言输入的文化关联性与情境真实感。与此同时，智能反馈机制也能根据学生具体表达方式给予语言风格与语法层面的微调建议，有效提升他们在非母语语境下的表达信心与准确度（Holmes et al., 2019）。

基于此，本研究提出核心问题：如何借助融合语音识别与自然语言理解的人工智能技术，构建一个支持多语同步开发的口语学习平台，服务于新加坡小学一至二年级学生的母语学习需求，推动跨语口语教材的智能化转型与本地化适配。

2. 项目目标

本研究历经两年，在新加坡理工大学智能机器学习团队与南洋理工大学国立教育学院语言教育团队（NIE）的通力合作下，开发了一个图画驱动、任务导向、互动性强的 AI 语言学习平台，旨在提供新加坡小学一至二年级学生在日常生活中使用母语（华语、马来语、淡米尔语）进行简短对话的自主习得机会。该平台整合多模态互动机制（听、说、看、反馈），并将其应用于仿说练习、词汇学习与智能对话任务中。同时，系统设计了即时反馈模块，帮助学生提升自主学习能力与自我监控意识，为儿童语言习得提供个性化、沉浸式的技术支持。

3. 平台设计与功能

3.1. 文化协同与图像内容的统一设计

平台的学习内容由贴近日常生活的主题图画所驱动，围绕个人、家庭与学校等常见情境精心绘制，风格活泼、生动，富有趣味。图画内容体现新加坡多元文化的社会生活，并根据华族、马来族与印度族三大族群的文化背景进行适当调整，旨在增强学生的跨文化认同与情境理解。开发过程中，图像内容须兼顾各族文化的共通性与差异性。例如，原设定的生日宴会场景中出现的“烧猪”图因不适用于马来族而被提出修改；最终，为确保三语版本内容统一，平台将相关图像统一调整为冷盘食物。

3.2. 仿说练习（基本口语词库 Wordbank）

每幅图配套句子，学生反复聆听录音并模仿发音。系统提供三次录音机会，并通过语音识别显示错误词语。每轮结束后，自动总结学生表现，展示“好词”与“多练习”词语。请参考下列图像 1 至 3。

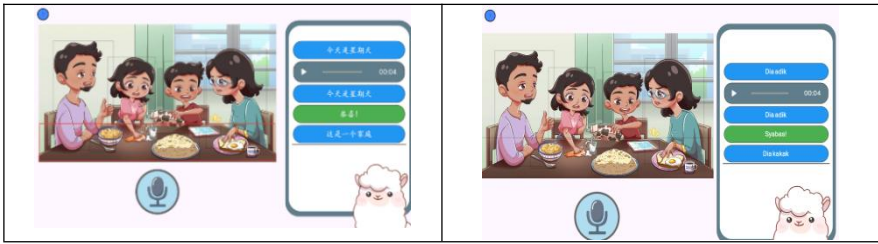


图 1 成功答题系统会以赞语回应，并自动切换到下一题。以上为华语和马来语平台的示例。

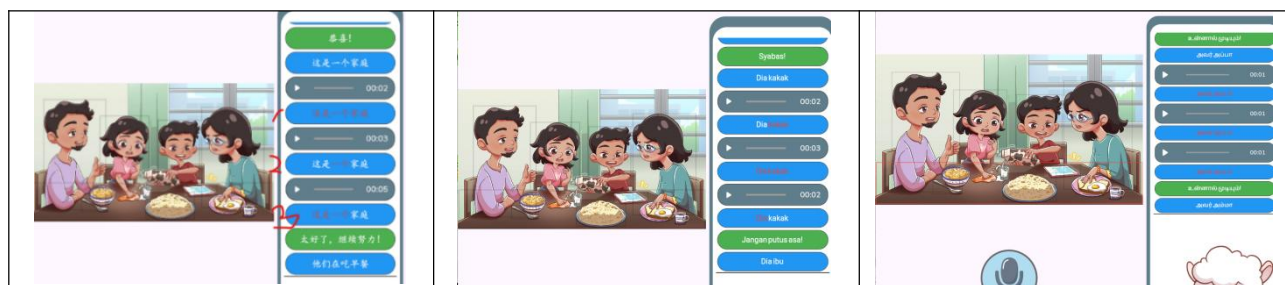


图 2 读错的字词，系统会亮出红字，显示答错。以上为华语、马来语和淡米尔平台的示例。

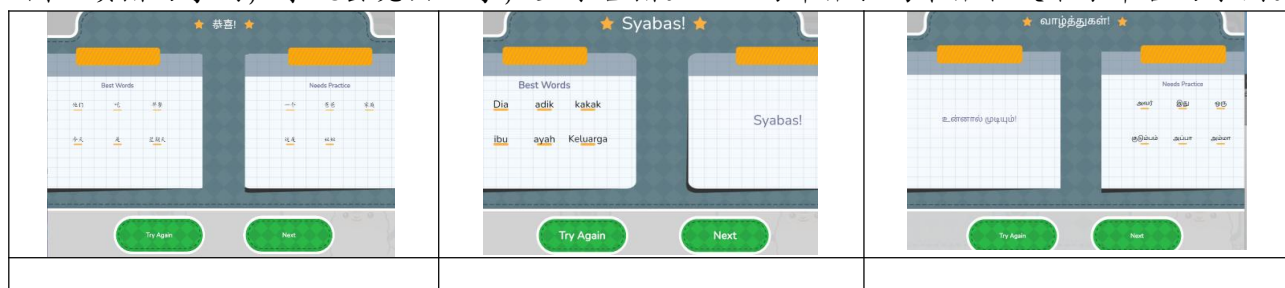


图 3 在“好词”与“多练习”词汇表里，每种语言都有独立的二表，每个类别的总结上限为六个。若所有词语都能正确读取，“多练习”词汇表会亮出赞语‘恭喜！’，类似下列马来语图示；若全部读错，“好词”词汇表会亮出类似下列淡米尔语图示的鼓励语。

3.3. 多选择词语练习 (MCQ)

系统自动播放题干与选项音频，学生听后选择正确词语。每题提供三次作答机会：答对即亮出赞语；前两次答错将给予鼓励语引导重试；第三次答错则提供正确答案及语音反馈。答题结束后，系统将根据整体表现给予相应的鼓励反馈。请参考下列图像 4 至 7。



图 4 每图设有四题，每题共有四个选项。



图 5 答对时，系统会以赞语弹窗回应，并让用户点选下一题。



图 6 前两次若选择错误，系统会亮示鼓励语让学生继续尝试。

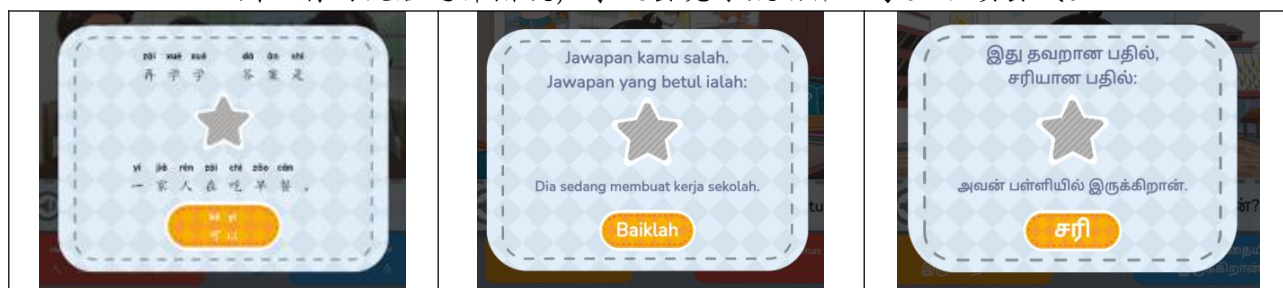


图 7 第三次答错，系统就会亮示鼓励语，并自动展示正确答案。

为帮助用户更专注内容理解，系统会随机排列每道题的选项，避免用户通过记住选项的位置或颜色机械作答，影响练习效果。

完成四道题目后，系统会以星形标记正确答题数，并按各语提供表 1 的赞语或鼓励语。学生点击“可以/OK/சரி (OK)”键后，系统会返回题目选择主页。请参考下列图像 8。

表 1 各语的赞语或鼓励语。

	华语	马来语	淡米尔语
四题全对，四星：	恭喜！	Syabas! 好样的！	வாழ்த்துகள் 恭喜
答对三题，三星：	真棒！	Bagus! 很好！	நல்ல முயற்சி 不错的尝试
答对一、两题，一、二星：	太好了，继续努力！	Jangan putus asa! 不要放弃！	உன்னால்/ உங்களால் முடியும் 你能做到
四题全错，无星：	再努力，你可以的！	Ingin mencuba lagi? 想再试一次吗？	மீண்டும் முயற்சி செய்யலாமா 我们可以再试一次吗？



图 8 四题全对时系统的赞语弹窗。

3.4. 智能对话 (Chatbot)

基于图画与任务情境，学生与 AI 展开简短会话（约 7-13 轮）。采用生成式 AI 技术，依据学生回应动态调整对话路径。

“乐言”系统的智能对话在展示新对话时，首先播放相关句子。学生听完句子后，如果不清楚，可以点击对应按键重复播放。确认听清楚后，学生点击麦克风按钮进行答复。完成后，学生可点击录音回放检查是否满意。如同仿说练习（Word Bank），如果不满意，学生可以继续录制，直到达到理想效果再点击发送。如图所示，该学生录制了四次，并最终选择最满意的一次发送。收到答复信息后，系统自动生成回复，开展下一个对话。目前，不论答对与否，系统都统一回应，“正确的答案是……”，如表 2 和图 9 所展示。

表 2 各语开展对话的范例节选。

	华语	马来语	淡米尔语
系统提问：	“欢迎！让我们看一下图片里一个拿着书的男孩。这个男孩在做什么？”	Selamat datang! Mari kita lihat budak lelaki yang memegang buku dalam gambar. Apakah budak lelaki itu sedang lakukan? (译文：欢迎！让我们看一下图片里一个拿着书的男孩。这个男孩在做什么？)	வரவேற்கிறேன்! படத்தில் ஒரு புத்தகம் பிடித்திக் கொண்டிருக்கும் ஒரு சிறுவனை நோக்குக. சிறுவன் என்ன செய்கிறான் (译文：欢迎！让我们看一下图片里一个拿着书的男孩。这个男孩在做什么？)
用户选择最佳答复，点送录音内容：	看书。	Baca buku. (译文：读书。)	புத்தகம் படிக்கிறார் (译文：他在读书)
系统回应：	正确的答案是“看书。”。男孩在书上做什么？”	Jawapan yang betul adalah: 'Baca buku'. Apa yang budak itu lakukan dengan buku itu? (译文：正确的答案是“读书。”。男孩在书上做什么？)	சரியான பதில்: 'புத்தகம் படிக்கிறார்'.. அந்த சிறுவன் புத்தகத்துடன் என்ன செய்கிறான்? (译文：正确的答案是“他在读书”。男孩在书上做什么？)

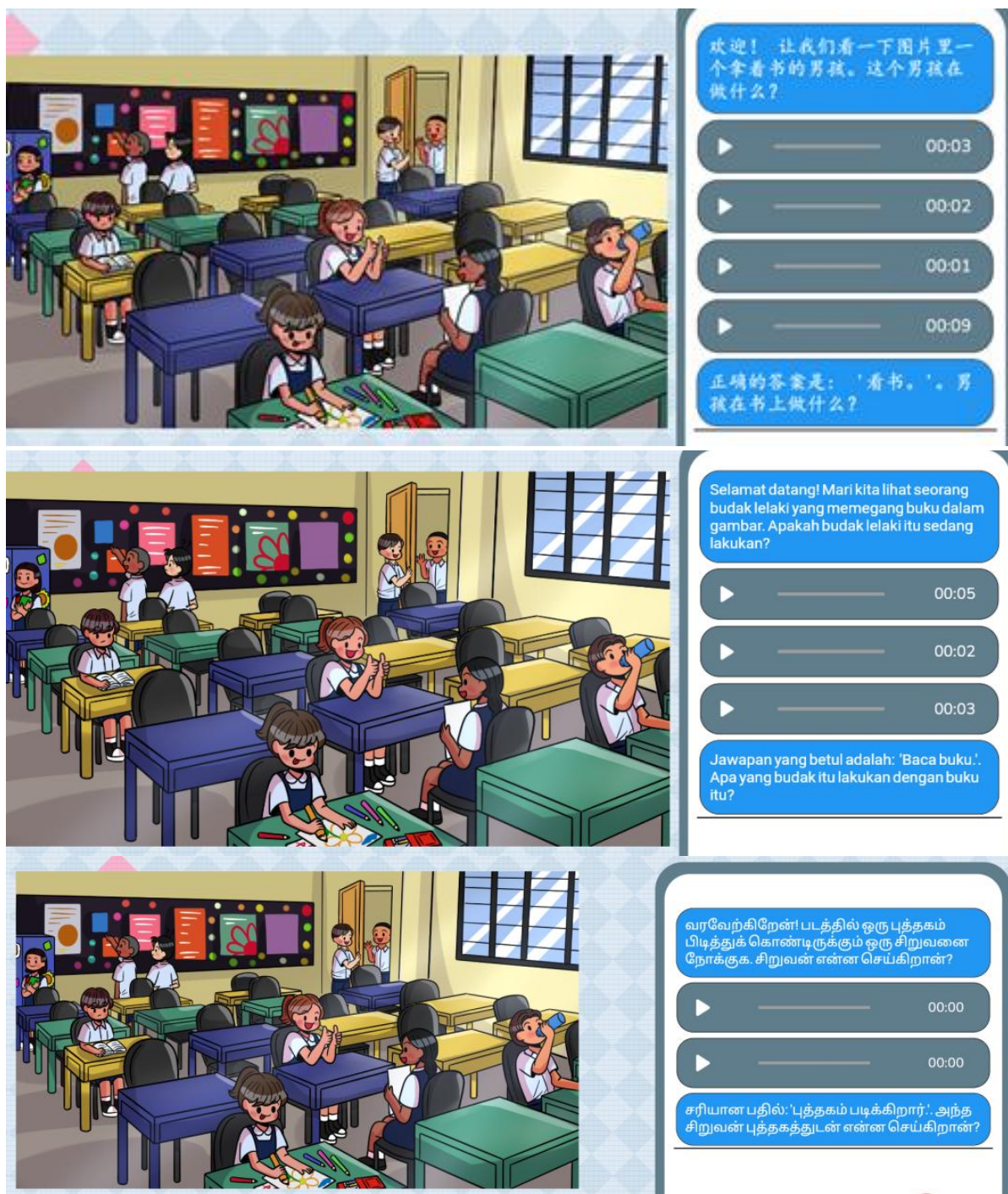


图9 学生可重复录音，并择优提交。

三个语言平台在界面布局与使用功能上保持高度一致，体现统一的设计框架，其核心语言学习辅助功能均以语音播放为主。在此基础上，平台根据各语言的语音特点、教学进度安排，以及日常用语和文化表达习惯的差异，进行内容层面的本地化适配，实现“结构统一、内容因语而异”的三语同步设计理念。表3列举了三种语言在内容呈现上的部分差异示例：

表3 三语内容呈现差异概览

	华文	马来文	淡米尔文
平台语言辅助差异	通过拼音注音作为支持，帮助学生绕过汉字识读障碍，专注于口语	无	无

	华文	马来文	淡米尔文
	练习本身，提升语言输出的流畅度。		
发音	属于声调语言，有四个声调，声调是第二语言习得的难点。	音节较简单，没有字调问题。	包含丰富的辅音-元音组合，音节结构相对复杂；语音系统的学习难度明显高于华语和马来语。
语系	属于汉藏语系，广泛用于中华地区及全球华人社区。	属于南岛语系，广泛用于马来西亚、新加坡、印度尼西亚等地。	属于达罗毗荼语系，主要用于印度南部和斯里兰卡。

4. “乐言”——智能机器学习的特色与创新点

“乐言”平台融合了机器学习、自然语言处理与教育科技设计的优势，体现出智能语言学习系统的关键特色与创新路径，主要包括以下四点：

4.2 智能对话系统

基于自然语言处理技术，平台能够模拟日常情境中的语言交流，实现与学生进行多轮互动式对话，提升语言输出的真实性、频率与表达质量。

4.3 个性化反馈机制

系统结合语音识别与学习数据分析，对学生的发音、语调、词汇准确度等维度进行即时诊断，提供针对性的改进建议，有效支持差异化教学需求。

4.4 可视化学习体验

以图像引导语言输入与输出，通过图画描述任务降低认知负担，同时激发学生对场景语言的理解力与表达兴趣。

4.5 自主学习支持

平台具备高度可调节性，学习者可依据自身节奏自由选择学习内容与练习频率，强化学生的学习主控感与自我监控能力。

5. 实践反思与系统改进

在新加坡“英语主导、母语传承”的双语教育政策框架下，学生除修读英语，并以英语为各个课程的教学媒介语外，亦须选择一门族群母语（华语、马来语或淡米尔语）作为第二语言。由于母语课程紧密承载各族群的语言与文化传统，不同语言在语用、表达方式及内容组织上存在因此显著差异。这使得在统一课程目标与教学任务下，如何实现多语教学内容的等效设计，成为课程开发与资源建设中亟待解决的关键难题。

本项目在技术团队与三大母语教育专家（华文、马来文、淡米尔文）的多轮协商、测试与迭代基础上，遵循“异中存同、同中尊异”的设计原则，尊重各语言的语法逻辑与学习规律，推动平台内容的等效性与本土化适配。在协作过程中，以英语作为跨语种的沟通媒介，协助多语言专家与技术专家团队达成共识，实现内容的同步开发。

新加坡这一独特的多语教育生态为本研究提出关键议题：如何借助融合语音识别与自然语言理解的人工智能技术，开发一套支持多语同步推进的口语学习平台，从而推动语言教材

的智能化转型与跨语种应用。作为新加坡首次以三语同步开发的 AI 口语学习系统，“乐言”不仅回应本地教育政策的多语需求，也为国际社会提供了跨语言智能教材开发的实证范式。

作为第一代智能口语学习平台，“乐言”在实践过程中积累了宝贵经验，也暴露出一些需优化的系统问题。本项目已完成初步开发，并在真实教育场景中进行了应用测试。

首先，第一代系统在反馈机制方面仍有待完善。目前的设计在部分任务中未能提供清晰的“正误”判断，容易造成学生对答案准确性产生疑惑，影响学习信心。特别是在开放性问题中，现有系统无法提供任何回应，而是直接跳转到下一个话题。这种 AI 回应的模糊性与单一反馈方式，使学生难以理解其表达是否得当。新工大目前正尝试引入新的语言模型来改善回应机制，具体进展尚不明确。此外，淡米尔语的问题也最为明显，原因有待查验。

其次，在对话交互系统中，存在跳题、回应逻辑不连贯的问题。这影响了语言任务的连贯性与真实感，削弱了学生的沉浸体验。未来版本应进一步优化对话策略的上下文保持能力，提高 AI 在连续对话中的表现一致性。

此外，系统对于不同表达方式的识别度还需提升，例如学生用词与参考答案不同但语义相近时，系统反馈仍显僵化，未能体现智能教学的包容性与灵活性。

随着自然语言处理技术与生成式 AI 的持续进化，这些问题在未来版本中有望得到解决。目前团队已针对上述问题制定功能更新路线图，逐步完善反馈精度、对话连贯性与语义识别能力，以进一步提升学习成效与用户体验。

6. 结论

“乐言”平台展现了人工智能在跨语言教学领域的创新潜能，尤其在多语学习环境中构建“智能一机多语”的应用框架，回应了新加坡作为双语教育国家对语言平等与教育公平的现实需求。平台在设计之初即采用跨语同步开发策略，确保教学任务、图像场景与交互机制在华语、马来语和淡米尔语三种母语中实现等效表达，既尊重语言差异，又促进内容协同。

平台融合图画驱动、任务导向与个性化反馈机制，为学生构建出一个具有高度互动性、即时应答与语言迁移支持能力的学习空间。通过人机对话模拟日常交流情境，即便学生处于母语输入相对匮乏的家庭语境中，亦能借助 AI 系统获得持续、结构化的语言练习机会，从而提升其口语流利度、语用准确性与表达信心。实际应用结果显示，系统有效拓展了学生接触母语的时间与频率，也为跨语智能教材的开发提供了实践路径与可推广模型。

展望未来，随着自然语言处理与语音技术的持续优化，以及虚拟现实（VR）与增强现实（AR）在教育场景中的逐步普及，“乐言”平台有望发展为一个集智能问答、情境互动与虚拟角色扮演为一体的沉浸式语言学习系统。它不仅能进一步提升学习兴趣与表现，还将为人工智能在语言教育领域的深层融合提供实践范式与研究借鉴。“乐言”平台展示了人工智能在小学母语教育中的创新应用。通过图画驱动、任务导向、反馈增强的语言学习系统，学生在家庭中获得更多互动机会，提升口语能力。未来结合 VR/AR 等沉浸式技术，“乐言”将为语言学习者提供更丰富、真实、有趣的学习体验，为 AI 教育实践提供重要参考路径。

鸣谢：

本文研究成果源自新加坡人工智能署（AI Singapore）资助项目 **LEARN: Language automated Evaluation by generating Answers/questions from caRtooNs**（项目编号：AISG-SIT 12/22 SA）。特此致谢。

参考文献

- Chen, X., Xie, H., Zou, D., & Hwang, G. J. (2020). *Application and theory gaps during the rise of artificial intelligence in education*. *Computers and Education: Artificial Intelligence*, 1, 100002. <https://doi.org/10.1016/j.caeai.2020.100002>
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Boston: Center for Curriculum Redesign.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). *Systematic review of research on artificial intelligence applications in higher education – where are the educators?* *International Journal of Educational Technology in Higher Education*, 16(1), 1–27. <https://doi.org/10.1186/s41239-019-0171-0>
- VanLehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education*, 16(3), 227–265.