

基于“汉音语声”小程序的汉语语音学习效能驱动机制研究

石柳

赵文君

王譞翀*

华北水利水电大学, 中国

摘要

研究围绕感知与产出教学模式, 以语音学习模型 (Speech Learning Model, SLM) 与自我决定理论 (Self-Determination Theory, SDT) 为框架, 探讨了“汉音语声”小程序在提升第二语言 (L2) 学习者汉语语音习得中的成效。研究邀请 70 名汉语学习者作为研究对象, 随机分为实验组 (n=35)、对照组 (n=35)。通过 16 周的准实验设计, 比较“汉音语声”融入课堂教学与传统面授教学的差异。结果显示, 实验组在短期表现和长期保持方面均显著优于对照组, 且遗忘率更低。机制分析表明, 自主性显著调节了工具使用与即时进步的关系 ($\beta=1.409, p=0.006$), 关联性促进了持续进步 ($\beta=1.623, p<0.001$), 而能力感与短期收益显著相关 ($r=0.306, p<0.01$)。整体而言, 研究验证了 SLM 在数字化语境下的感知-产出机制, 并揭示了 SDT 核心需求在语音学习成效中的中介作用, 为数字技术赋能的汉语语音教学提供了理论与实践启示。

关键词

“汉音语声”小程序, 感知-产出教学模式, 语音学习效能, 驱动机制

1 引言

在教育数字化和全球化的深度融合背景下, 第二语言教育正经历着显著的变革。近年来, 数字技术辅助语言学习逐渐发展为一个独立的研究领域, 其“移动性、通用性、交互性、易用性、多媒体集成和个性化”等特点 (Cook, 2010; 姜丽萍、王立, 2021; 李宝敏等, 2018) 为二语语音学习高频输入、即时反馈与反复对比练习的实现提供了新的可能性 (Li et al., 2022)。

汉语作为一种典型的声调语言, 其语音系统具有显著的语义区分功能, 给来自非声调语言背景的学习者带来了独特挑战 (马秋武、赵永刚, 2017)。研究显示, 第二语言学习者在语音感知和产出方面常常受到母语迁移效应的影响 (Flege, 1995), 难以识别和区分普通话声调的细微差异 (邓丹、林雨箐, 2017)。传统语音教学依赖教师示范与重复练习, 受限于课堂时间和班级规模, 往往难以提供充分的个性化指导和实时反馈 (马秋武、翟海莹, 2024; 石锋, 2022; 石锋、温宝莹, 2023)。在此背景下, Praat 与 Audacity 等数字工具被引入语音教学, 通过语谱图、音高轨迹等可视化形式帮助学习者理解声学特征, 并在感知与产出对比中获得更明确的语音反馈 (邓丹、朱琳, 2019; 梅丽, 2025)。

* 通讯作者。联系电邮: wangxuanxuan@ncwu.edu.cn

从理论层面看,SLM 强调感知与产出之间的循环互动是二语语音范畴建构的核心机制 (Flege, 1995; Flege & Bohn, 2021)。在这一框架下,学习者通过持续的感知输入与产出实践逐步调整其语音表征,从而建立新的语音范畴。已有研究进一步表明,在二语语音训练中保持感知输入与产出练习之间的平衡,对于促进稳定的语音范畴结构具有重要作用 (Flege, 1995)。然而,在具体教学实践中,这一理论机制仍缺乏系统性的教学干预检验。尤其是在部分数字化或课堂语音训练工具中,训练活动往往偏重产出练习,而对感知输入的系统强化不足,这可能在一定程度上限制了感知—产出循环机制的有效发挥 (Flege & Bohn, 2021)。随着数字技术在语音教学中的广泛应用,数字化语音学习工具如何促进学习者的学习投入与持续性这一问题持续受到广泛关注。SDT 指出,当学习者的自主性 (Autonomy)、胜任感 (Competence) 与关联性 (Relatedness) 需求得到满足时,其学习投入与学习成就将显著提升 (Ryan & Deci, 2017)。这一理论视角对于解释在语音学习中因母语迁移而易产生挫败感的汉语学习者的学习坚持具有重要启示意义 (Lietaert et al., 2015)。尽管已有研究证实了数字工具在远程或课堂语音教学中的积极作用(如 Watanabe et al., 2020),但较少从 SDT 视角考察学习者在工具使用过程中自主性、胜任感与关联性等心理需求的满足情况 (Ryan & Deci, 2017)。

从实践层面来看,数字化教学工具的应用可以为汉语语音的感知与产出机制的操作化提供现实条件 (Dalman, 2025)。相较于以教师示范为主的传统课堂训练,基于移动终端的语音学习工具能够在课内外为学习者提供高频、可重复的感知输入,并通过即时可视化反馈支持学习者进行自主对照与修正 (Valk et al., 2010)。这种以学习者为中心的练习方式,有助于突破课堂时间与教学节奏的限制,使感知—产出循环得以在更高频率和更稳定的条件下发生。与此同时,数字化工具所提供的自主练习空间与即时反馈机制,可能在实践中增强学习者的胜任感与学习控制感,从而影响其学习投入与持续性 (Zhou et al., 2025)。但现有研究较少在真实教学干预中系统考察数字化语音工具如何同时作用于语音加工机制与学习者心理体验,尤其是在以声调为核心、感知—产出负荷较高的汉语语音学习情境中,此类整合性实证研究仍显不足 (Hao & Lv, 2023)。这也进一步凸显了从理论与实践结合的视角,探讨数字化语音教学效能驱动机制的必要性。

基于上述研究不足,本文以 SLM 与 SDT 为理论框架,聚焦数字化语音教学中感知—产出机制的有效性及其心理驱动路径,系统考察“汉音语声”小程序在汉语语音教学中的作用。重点聚焦以下两个问题:

1. 在“汉音语声”小程序的干预下,实验组短期与长期的语音学习效果是否优于对照组?
2. 实验组相比对照组,学习者的自主性、胜任感与关联性是否显著高于传统教学?这些心理需求的满足程度是否与语音学习进步显著相关?

为回答上述问题,本研究将采用为期 16 周的准实验设计,招募 70 名母语水平相近的中级汉语学习者,比较基于“汉音语声”的感知产出教学模式与传统教学模式在语音评估结果上的差异,验证该教学模式对汉语语音能力保持的影响。同时,通过对 SDT 的量表评估,还将探讨学习者在自主性、胜任感和关联性等维度上的心理满意度,以揭示理论驱动的技术干预如何同时提升学习成果与心理动机。

2 研究 1: 汉语语音学习效果的实验

研究 1 以 SLM 为依据,采用课前评估、课后测验与定期评估的测量方式,旨在通过教学实验系统检验“汉音语声”测评工具在提升汉语语音学习效果方面的有效性,探究其在真实课堂教学中对学习者的语音表现是否存在促进作用。

2.1 测量工具

研究选用“汉音语声”小程序作为核心测评工具，它能够自动评估学习者汉语发音水平，标注语音偏误，同时生成可量化的语音表现数据，能够为教学实验提供稳定、客观的测评依据。“汉音语声”以微信小程序为载体，以 Java 语言构建服务端，使用 vue、Uni-app 技术构建前端应用框架，CSS 技术构建小程序内容与表象样式，基于讯飞语音评测流式版 API 接口对接语音智能技术，自动评测发音水平、标注发音错误等内容。图 1 与图 2 是“汉音语声”的感知（听力）测试页面与产出（发音）测试页面，可以辅助汉语教师将感知与产出教学模式应用于汉语语音课堂，广泛收集汉语学习者课前、课中、课后的测评数据，便于汉语教师开展数据收集、量化分析、语图绘制，为调整课堂教学策略、平衡感知输入、产出练习的比重提供指导性意见；也可以帮助汉语学习者通过听力训练、发音训练提升辨认汉语语音的敏感度与声韵调的发音技巧，综合提升汉语语音的感知与产出能力。

在研究 1 中，“汉音语声”主要负责采集实验组与对照组在课前评估、课后测验、阶段性评估中的语音数据，用于呈现学习者在不同阶段的语音学习情况。此外，“汉音语声”还可以提供即时反馈、个性化练习与复习计划，便于实验组在课堂与课后自主训练中使用，而对照组在教学中则不使用该工具，仅依赖传统课堂活动。

图 1. “汉音语声”小程序感知（听力）测试页面

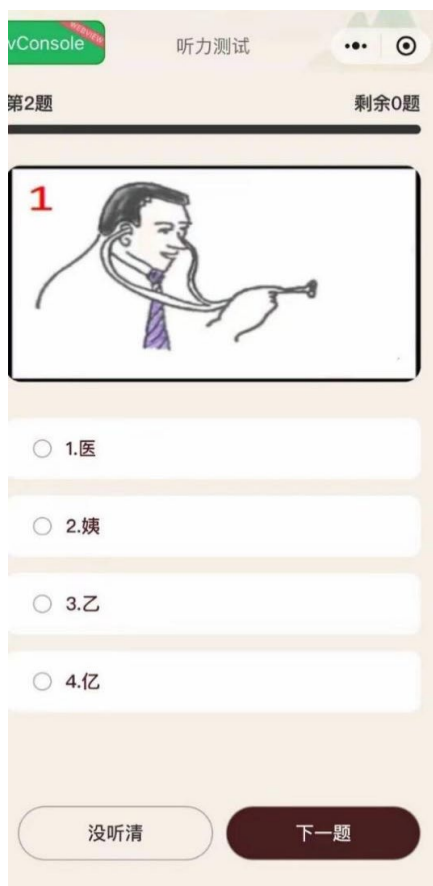


图 2. “汉音语声”小程序产出（发音）测试页面



2.2 实验对象

研究选择在一所拥有丰富留学生资源的高校进行。这所高校位于河南省郑州市，年均招收留学生 400 余名，同时承担了留学生本科、硕士及博士的培养工作。为保证实验的严谨性、公正性，我们在 97 名留学生中筛选出 70 名符合条件的参与者。他们抵达中国的时间范围在 1-1.5 年，未系统学习过汉语语音，且年龄区间在 19-22 岁。这 70 名被试虽然有中文文化课程的学习经验（选修），但他们并未系统学习过汉语语音训练与正音课程，尤其在声母、韵母及声调辨析方面仍存在较大提升空间。

随后，我们采用随机抽样的方法，将 70 名参与者随机分配至实验组 ($n=35$)、对照组 ($n=35$)。为验证实验组、对照组具有可比性，我们分别对性别、年龄、语言环境等因素进行差异性分析。结果显示，实验组、对照组在各变量上的差异均不显著 ($p>0.05$)，表明两组样本分布基本一致，具有可比性。

2.3 数据收集

研究以“汉音语声”作为核心评估工具，持续追踪实验组、对照组学习者在 16 周教学周期内的学习表现。为保证教学实验的可参照性，两组被试的上课频次相同，均为一周一次，每次 90 分钟。课程均由同一位授课教师授课，以确保教学内容、教学节奏等的一致性。教学实验开始之前，实验人员分别向实验组与对照组说明研究目的、实施流程及保密原则，使所有参与者充分了解并熟悉教学实验安排。为进一步呈现两组教学流程与教学强度的可比性，16 周课程的教学内容与实施流程安排如下表所示：

表 1. 教学内容与流程表

周次	教学模块	实验组	对照组
1-2	汉语拼音基础	使用“汉音语声”进行音节练习，系统会自动标注发音偏误并进行即时纠正	课堂听辨训练，教师口头纠音与示范练习
3-6	声调组合与辨析	“汉音语声”声调训练与阶段性评分反馈	课堂声调练习与教师口头反馈，模拟声调组合练习
7-12	句子练习	系统化句子朗读训练，阶段性巩固与复习，由“汉音语声”提供个性化反馈	课堂句子朗读练习，教师逐一指导，提供口头反馈与纠正
13-16	综合口语表达	综合句子与短文朗读，按照“汉音语声”计划，完成自主练习与强化	课堂综合口语输出，教师安排口头练习与反馈，布置复习任务但不进行系统复习工具辅助

教学周期内，两组被试将分别完成为期 16 周的汉语语音课程。实验组全程使用“汉音语声”进行即时反馈与自主练习，覆盖预习、课堂训练及阶段性复习。对照组虽未使用工具，但课堂活动安排、训练内容与课时量与实验组保持一致，包括听辨训练、发音练习、句子朗读及教师反馈。两组在课堂时间、活动环节及练习量上基本均衡，唯一区别在于实验组借助“汉音语声”工具获得即时反馈和自主练习支持，而对照组仅依靠传统教学指导。

值得说明的是，对学习效果的评测通过干预前评估、干预后评估、追踪测评三个阶段实现。干预前评估作为基线，用于初步测评实验组、对照组对汉语语音的熟悉程度、发音准确程度，为进一步对比干预后评估、追踪测评提供判断依据。

干预后评估是指教学周期（第 16 周）结束后，两组被试将分别完成第二轮评估。评估形式与干预前评估保持一致，实验组使用“问卷星”线上完成评估，对照组则使用纸质问卷接受汉语语音的感知与产出评估。值得说明的是，第一轮评估与第二轮评估的内容完全一致，但题号顺序、题项顺序均保持乱序，以避免记忆对评估结果造成干扰。

追踪评估阶段是在干预后评估两周后完成，实验组继续接受“汉音语声”的教学支持。具体而言，实验人员由“汉音语声”导出干预评估结果，由老师根据该结果制定严谨、完善的复习计划与练习题，录入“汉音语声”后台，实验组被试则继续完成复习与练习。对照组则继续保持汉语课堂的“教学惯性”，老师带领复习，布置、督促学生完成复习任务，同时反馈追踪评估结果。两周后，我们继续对实验组、对照组被试进行追踪评估，评估内容依然与前两次评估内容保持一致，题目顺序、题项顺序再次乱序。最终，实验人员逐一检查三次评估数据，将其录入 Excel 表格，对数据进行转化、清洗与分析。

2.4 数据分析

Cronbach 的 alpha 系数是衡量问卷内部一致性的重要指标，常用于检验量表或测试工具的信度 (Dorsah, 2026)。为考察本研究中感知与产出测量工具内部的一致性，本研究对相关数据进行了信度分析。结果显示，Cronbach's α 为 0.864，校正后的项目总相关 (CITC) 为 0.761，均达到较高水平。这表明本研究数据具有良好的内部一致性，各题项之间的相关性较强，能够有效反映学习者在感知与产出层面的实际表现。

为回应研究问题 1，我们使用 RStudio 构建了线性混合效应模型 (Linear Mixed-Effects Model, LMM) 与分段增长曲线模型 (Piecewise Growth Curve Model, PGCM)。研究将组别 (实验组与对照组) 和次数 (干预前评估为第一次、干预后评估为第二次、追踪评估为第三次) 设定为固定效应，以考察不同教学条件在不同时间点上的学习效果差异，并进一步检验两组学习者在短期 (干预后) 与长期 (追踪评估) 语音表现上的变化轨迹是否存在显著差异。此外，模

型还纳入了组别与时间的交互效应，以分析不同教学干预条件下学习进展模式的差异。与此同时，为控制个体间差异对结果的影响，本研究将学习者个体（participant）设定为随机效应，从而更准确地估计组别与时间对语音学习结果的影响。

进一步地，我们构建了分段增长曲线模型，以更细致地刻画两组学习者在三个评估时点上的变化趋势。该模型将“干预前 - 干预后”和“干预后 - 追踪评估”设定为两个独立的生长阶段，旨在分别检验“汉音语声”对短期学习提升与长期保持的影响强度，并比较实验组与对照组在两个阶段的生长斜率是否存在系统性差异。

2.5 研究结果

2.5.1 描述性统计

表 2 展示了实验组和对照组在三个评估点（干预前评估、干预后评估和追踪评估）的描述性统计数据。在干预前评估中，两组得分相近，整体处于同一水平，但实验组略高于对照组（实验组：M=1.771，SD=0.731；对照组：M=1.629，SD=0.770），表明两组在初始水平上具有可比性。在干预后评估中，接受“汉音语声”干预的实验组其平均得分高于对照组（实验组：M=4.829，SD=0.785；对照组：M=4.086，SD=0.812）。追踪评估结果显示，实验组得分仍然优于对照组（实验组：M=5.486，SD=0.658；对照组：M=4.800，SD=0.797）。

表 2. 实验组和对照组在三次评估中的测试分数描述性统计结果

名称	n	最小值	最大值	平均值	标准差	中位数
实验组 - 前测	35	0.000	3.000	1.771	0.731	2.000
实验组 - 后测	35	3.000	6.000	4.829	0.785	5.000
实验组 - 追踪	35	4.000	6.000	5.486	0.658	6.000
对照组 - 前测	35	0.000	3.000	1.629	0.770	2.000
对照组 - 后测	35	2.000	6.000	4.086	0.951	4.000
对照组 - 追踪	35	3.000	6.000	4.800	0.797	5.000

2.5.2 线性混合效应模型

为了检验“汉音语声”能否提升学习效果，我们构建了包含固定效应与随机效应的线性混合效应模型（LMM），用于探讨实验组、对照组在不同教学条件下的表现差异。其中，组别、次数设定为固定效应，用以预测学习者在追踪评估阶段的语音表现；学习者则作为随机效应，以控制个体差异带来的影响。

表 3. 汉语语音教学实验线性混合效应模型估计结果

预测变量	β	SE	SD	df	t	p	Cohen's d	
截距	4.955	0.636	-	52.37	7.788	0.001***	-	
固定效应	组别	-0.522	0.183	-	65.75	-2.840	0.006**	0.68
	前测	-0.083	0.127	-	3.33	-0.656	0.555	0.16
	后测	0.251	0.098	-	64.95	2.548	0.013*	0.61
	随机效应	-	-	0.092	-	-	-	0.009
残差	-	-	0.704	-	-	-	0.495	

注：*** $p < 0.001$ ，** $p < 0.01$ ，* $p < 0.05$

表3描述了两组被试的线性混合效应模型估计结果。其中,固定效应截距为4.955 ($SE=0.636$, $df=52.37$, $t=7.788$, $p<0.001^{***}$),组别效应显著 ($\beta=-0.522$, $SE=0.183$, $df=65.75$, $t=-2.840$, $p=0.006^{**}$; Cohen's $d=0.68$)。表明对照组在追踪评估中的表现显著低于实验组。这一结果验证了“汉音语声”在长期语音学习中的显著促进作用。其次,干预前评估的综合系数为-0.083 ($SE=0.127$, $df=3.33$, $t=-0.656$, $p=0.555$; Cohen's $d=0.16$),未达到统计显著水平,表明学习者在干预前的基础能力并不能有效预测追踪评估阶段的表现,两组间初始水平差异对最终结果没有影响。如表3所示,干预后评估则对跟踪评估具有显著预测能力 ($\beta=0.251$, $SE=0.098$, $df=64.95$, $t=2.548$, $p=0.013$; Cohen's $d=0.61$):干预后得分每提升1个单位,追踪评估平均得分提高0.251分。这进一步说明“汉音语声”在短期内提升的学习效果能够延续并转化为长期的学习收益,再次支持其作为二语语音教学工具的有效性。

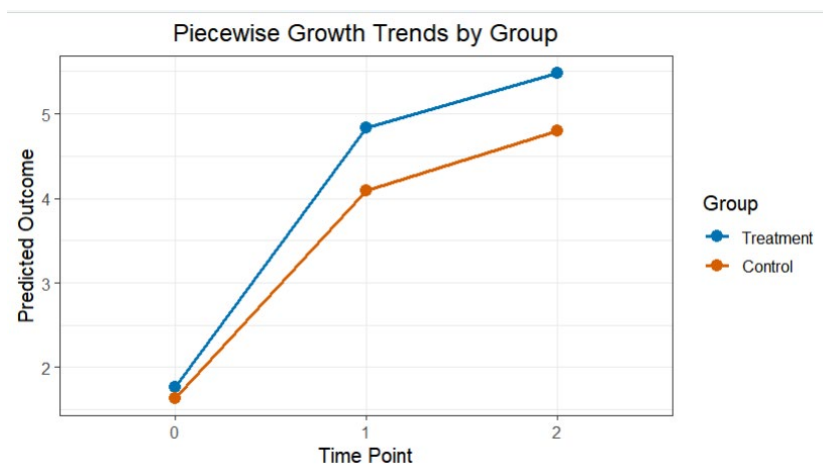
2.5.3 分段增长曲线模型

图3直观展示了实验组与对照组在三个时间点上的预测增长路径,其趋势与分段增长曲线模型的统计结果高度一致。如表4所示,截距值为1.77 ($SD=0.137$, $t=12.879$, $p<0.001$),表明两组在基线阶段(时间点0)的得分显著高于零,且处于相近水平。组别主效应并不显著 ($\beta=-0.143$, $SE=0.195$, $t=-0.734$, $p=0.465$),进一步说明实验组与对照组在干预前不存在显著差异,这也与图1中两条曲线在时间点0几乎重合的情况相符合。时间阶段主效应显著, ($\beta=3.06$, $SE=0.184$, $t=16.615$, $p<0.001^{***}$),说明随着时间推移,两组学习者的语音表现整体均显著提升。组别与时间阶段的交互项系数为-0.600 ($SD=0.260$, $t=-2.306$, $p=0.024$),表明随着时间推移,实验组与对照组在学习效果的保持程度上存在显著差异,实验组相较对照组在干预后测评阶段与追踪测评阶段展现出更快、更具备持续性的增长。

表4. 分段增长曲线模型参数结果

参数	值	标准差	DF	t	p
(截距)	1.771	0.137	68	12.879	0.000***
组别	-0.143	0.195	68	-0.734	0.465
时间阶段	3.06	0.184	68	16.615	0.000***
组别:时间阶段	-0.600	0.260	68	-2.306	0.024*

图3. 实验组与对照组分段增长趋势图



3 研究 2：基于 SDT 的学习者心理需求与语音学习表现

研究 2 以 SDT 视角为切入点，尝试通过组间比较与相关分析的方法，探讨实验组、对照组汉语语音学习者的内在心理机制。重点关注自主性、胜任感、关联性三种基本心理需求与语音学习表现的联系，旨在揭示不同教学模式在心理需求满足方面的差异，厘清这些心理因素在语音学习成效中的作用路径。

3.1 测量工具

本研究采用基于 SDT 的基本心理需求满足量表 (Basic Psychological Need Satisfaction Scale, BPNS) 作为测量工具，用以评估学习者在汉语语音学习情境中的自主性、胜任感与关联性 (Ryan & Deci, 2017; Sheldon & Niemiec, 2006)。值得注意的是，该量表已被证实具有良好的结构效度与内部一致性 ($\chi^2/df=1.68$, CFI=0.97, RMSEA=0.07, Cronbach's $\alpha=0.86$)。

鉴于本研究聚焦于数字化语音教学情境，为增强量表与研究任务之间的情境契合度，在不改变原有理论维度与条目结构的前提下，研究团队邀请三位具有汉语语音教学与教育心理学背景的专家对量表条目进行审阅，并对部分题项的表述进行了情境化调整，使其更贴合学习者在汉语语音感知、发音练习及课堂互动中的真实体验。量表共包含 9 个条目，分别测量自主性 (3 项)、胜任感 (3 项) 和关联性 (3 项)，采用 7 点 Likert 量表计分。九个测量条目的具体内容如表 5 所示：

表 5. SDT 心理需求量表

维度	条目编号	测量条目
自主性 (Autonomy)	A1	在汉语语音学习过程中，我的学习选择基于自身的兴趣和目标
	A2	在汉语语音练习中，我能够按照自己的节奏和方式完成学习任务
	A3	我在语音学习中的学习选择能够真实反映我的学习意愿
胜任感 (Competence)	C1	在汉语语音学习过程中，我能够成功完成具有挑战性的发音任务
	C2	我认为自己能够应对汉语语音学习中的困难与挑战
	C3	在进行语音练习时，我对自己的发音能力感到有信心
关联性 (Relatedness)	R1	在汉语语音学习过程中，我感到与教师或同伴保持良好的互动
	R2	在语音课堂或练习中，我感到与对我重要的人建立了积极的学习联系
	R3	在汉语语音学习过程中，我在较大程度上依赖教师的反馈与指导

注：所有条目均采用 7 点 Likert 量表计分 (1=非常不同意, 7=非常同意)。

总体而言，量表得分越高表示学习者在相应心理体验上的感知程度越强；但对于反映学习支持来源或依赖程度的条目，其解释需结合具体教学情境与支持方式加以讨论。

3.2 实验对象

研究 2 的实验对象与研究 1 保持一致，包括实验组 (n=35)、对照组 (n=35)。为确保心理需求测量能够全面反映被试对整个教学周期的整体体验，本研究选择在研究 1 追踪测评结束后统一派发 SDT 量表。

3.3 数据分析

为探讨自主性、胜任感与关联性在“汉音语声”支持下的心理作用,本研究主要采用两类统计方法。首先,通过独立样本 *t* 检验比较不同教学模式下学习者心理需求满足程度的差异,以评估数字化语音教学的潜在动机效益。其次,在实验组内部采用 Pearson 相关分析检验基本心理需求与语音学习表现之间的关联程度,为理解心理需求在学习过程中的作用提供实证依据。

3.4 研究结果

3.4.1 独立样本 *t* 检验

为比较不同教学模式下学习者的自主性、胜任感与关联性差异,本研究对实验组 ($n=35$) 与对照组 ($n=35$) 的量表得分进行了独立样本 *t* 检验。结果如表 6 所示,实验组在九个测量条目中有八项显著高于对照组,效应量均达到中等至极大程度,显示“汉音语声”对心理需求满足具有明显促进作用。

表 6. 实验组和对照组在自主性、胜任感和关联性方面的独立样本 *t* 检验结果

	组 (均值 ± 标准差)		<i>t</i>	<i>p</i>	Cohen's <i>d</i>
	1.0 ($n=35$)	2.0 ($n=35$)			
A1	6.43 ± 0.61	4.74 ± 0.85	9.527	0.000**	2.28
A2	6.31 ± 0.68	5.17 ± 0.89	6.046	0.000**	1.45
A3	7.00 ± 0.00	4.77 ± 1.17	11.314	0.000**	3.81
C1	6.23 ± 0.81	4.60 ± 0.55	9.844	0.000**	2.37
C2	6.26 ± 0.66	5.03 ± 0.71	7.532	0.000**	1.79
C3	6.29 ± 0.57	5.49 ± 0.51	6.189	0.000**	1.48
R1	6.20 ± 0.80	3.89 ± 0.87	11.627	0.000**	2.77
R2	6.34 ± 0.68	5.11 ± 0.68	7.560	0.000**	1.81
R3	3.91 ± 0.45	4.54 ± 0.51	-5.520	0.000**	-1.31

注: * $p<0.05$ ** $p<0.01$

在自主性方面,实验组得分显著高于对照组 (A1: 6.43 ± 0.61 vs 4.74 ± 0.85 , $t=9.527$, $p<0.01$, $d=2.28$; A2: 6.31 ± 0.68 vs 5.17 ± 0.89 , $t=6.046$, $p<0.01$, $d=1.45$; A3: 7.00 ± 0.00 vs 4.77 ± 1.17 , $t=11.314$, $p<0.01$, $d=3.81$)。这些结果表明实验组学习者在学习节奏掌控、自主探索及课堂参与方面具有更强的自主感。在胜任感方面,实验组同样显著优于对照组 (C1: 6.23 ± 0.81 vs 4.60 ± 0.55 , $t=9.844$, $p<0.01$, $d=2.37$; C2: 6.26 ± 0.66 vs 5.03 ± 0.71 , $t=7.532$, $p<0.01$, $d=1.79$; C3: 6.29 ± 0.57 vs 5.49 ± 0.51 , $t=6.189$, $p<0.01$, $d=1.48$)。这说明实验组学习者在语音学习的掌握度、任务完成感与进步感方面有更高的主观体验。在关联性方面,实验组在两项指标上显著高于对照组 (R1: 6.20 ± 0.80 vs 3.89 ± 0.87 , $t=11.627$, $p<0.01$, $d=2.77$; R2: 6.34 ± 0.68 vs 5.11 ± 0.68 ,

$t=7.560, p<0.01, d=1.81$)。然而, 对于“学习过程中依赖教师反馈”的条目(R3), 对照组得分更高(4.54 ± 0.51 vs $3.91 \pm 0.45, t=-5.520, p<0.01, d=-1.31$), 提示两组在社交互动类型或反馈渠道偏好上可能存在差异。需要指出的是, 条目 R3 所测量的并非心理需求满足程度的高低, 而是学习者在语音学习过程中对教师反馈这一外部支持来源的依赖程度。因此, 该条目并不遵循“得分越高代表更优学习状态”的线性解释逻辑。对照组在 R3 上得分更高, 反映出其在传统课堂教学情境中更依赖教师主导的纠错与反馈; 而实验组在“汉音语声”支持下, 能够通过系统反馈与自主练习获得学习支持, 对教师反馈的依赖程度相对降低。这一结果表明, 两种教学模式在学习支持结构与反馈渠道上存在显著差异, 而非实验组在关联性需求满足方面的不足。

总体来看, 实验组在自主性与胜任感维度上系统性优于对照组, 在关联性维度的大多数条目上亦具有优势。这些结果支持“汉音语声”能有效提升学习者的心理需求满足水平, 符合 SDT 所强调的技术支持型学习环境对学习动机的促进作用。

3.4.2 皮尔逊相关系数分析

为进一步考察自主性、胜任感与关联性三项基本心理需求与语音学习表现之间的关系, 本研究在实验组内部计算了 SDT 总分及其三个维度与学习成效指标的 Pearson 相关系数, 结果如表 7 所示。SDT 总分与干预后评分显著正相关($r=0.426, p<0.01$), 并与短期进展亦呈显著相关($r=0.278, p<0.05$), 表明学习者在学习过程中体验到的心理需求满足程度越高, 其语音表现与短期提升越显著。然而, SDT 总分与长期进展之间的相关性不显著($r=0.012$), 提示心理需求更主要作用于即时学习阶段。

表 7. SDT 结构与学习成果之间的相关性

变量	干预后评估分数	短期进展	长期进展
SDT 评分	0.426**	0.278*	0.012
自主性评分	0.381**	0.255*	-0.028
能力评分	0.413**	0.306**	-0.081
相关度得分	0.267*	0.229	0.026

在自主性维度, 自主性得分与干预后评分显著正相关($r=0.381, p<0.01$), 并与短期进展呈中等强度的正相关($r=0.255, p<0.05$)。这表明能够自主调节学习节奏、主动参与任务的学习者在短期语音提升上表现更好, 但自主性并未显著预测长期保持($r=-0.028$)。胜任感维度, 胜任感与干预后评分呈显著正相关($r=0.413, p<0.01$), 与短期进展的相关性最为突出($r=0.306, p<0.01$), 说明学习者在学习过程中感受到的掌握感与成就感对其语音提升具有重要推动作用。同样, 胜任感与长期进展无显著关联($r=-0.081$)。在关联性维度, 关联性得分与干预后评分呈显著正相关($r=0.267, p<0.05$), 与短期进展呈弱相关趋势($r=0.229$, 未达显著水平)。关联性与长期进展的关系同样不显著($r=0.026$)。这一模式表明, 课堂中的互动支持和社会连结可能在学习表现中具有一定作用, 但影响更偏向即时阶段。

心理需求与语音学习表现之间呈现出一致模式: 自主性与胜任感对短期语音提升具有显著预测意义, 而长期保持不受三类需求显著影响。上述结果显示, 在“汉音语声”支持的数字化语音学习情境中, 心理需求满足主要作用于学习的即时收益阶段, 对学习保持效应的影响相对有限。

4 讨论

本研究通过准实验与纵向追踪,考察了“汉音语声”小程序在汉语语音习得中的应用效果,并试图从语音加工机制与学习动机两个层面解释其可能的作用路径。总体来看,数字化干预在短期提升与后期保持两个阶段均表现出一定优势,且学习者在使用过程中呈现出较为积极的心理体验。以下,将分别围绕研究问题一、研究问题二展开综合探讨。

4.1 研究问题一: 汉音语声对短期与长期语音学习效果的促进作用

研究一旨在通过准实验设计与纵向追踪,验证汉音语声小程序在 L2 语音习得中的效能。为回应研究问题一,我们将 LMM 与 PGCM 的统计结果,置于 SLM 的理论框架下进行综合诠释,以探究感知—产出模式相对于传统教学的优势。

根据实验一的研究结果可知,实验组在语音习得的各个阶段均展现出显著优势。LMM 分析显示,组间效应显著 ($\beta=-0.522, p=0.006$),实验组在干预后及追踪评估中的表现均显著优于对照组。这一实证结果高度契合 SLM 的预测,即通过针对性的感知训练能够克服第一语言 (L1) 语音表征对 L2 习得的限制 (Flege, 1995)。

SLM 理论认为,二语语音习得的难点在于学习者往往通过“不完整的母语感知过滤器”来处理目标语语音 (Flege, 1995)。传统语音教学往往较多依赖教师示范与重复模仿,对听觉辨识训练的系统性关注相对不足。这种教学取向在一定程度上可能影响学习者对声学线索的敏感性,使母语缺乏声调的学习者在区分普通话声调(如二声与三声)时更难捕捉其细微的声学差异。反之,“汉音语声”通过“感知—产出”的整合模块直接干预了这一机制。在课前感知练习中,学习者需识别样本录音中的声调错误;在产出练习中,系统提供的实时 AI 反馈与可视化音高轮廓 (Pitch Contours) 充当了关键的“认知脚手架”。这种将隐性听觉信息显性化的过程,帮助学习者建立了精确的 L2 语音范畴,从而规避了 L1 的负迁移。感知与产出之间的双向互动是 SLM 的核心假设 (Flege & Bohn, 2021)。实验组的数据表明,这种互动在数字化环境中得到了显著增强:准确的感知引导了更精确的产出,而基于产出错误的即时反馈又反过来提升了感知敏感度。值得注意的是,“汉音语声”不仅实现了 SLM 的理论预设,更通过分段增长曲线模型的结果展示了其在“对抗遗忘”方面的独特优势。虽然两组在基线水平无显著差异,但在“干预后—追踪”阶段,实验组展现出更具持续性的增长轨迹,而对照组的生长斜率则明显放缓。这表明,感知—产出训练帮助学习者将课堂内的显性知识转化为内隐的程序性知识,实现了从“依赖教师纠音”到“自主语音监控”的转变。

Flege 与 Bohn (2021) 认为,感知与产出的结合有助于促进较为稳定的二语语音表征的形成。本研究的结果进一步显示,数字环境提供的个性化反馈(如对比声学分析)在一定程度上弥补了传统课堂中难以实现高频听辨训练的不足。实验组在追踪阶段的表现也提示,在可扩展的数字化条件下运用 SLM 的核心原则,可能不仅有利于短期技能的提升,也有助于支持语音能力的较长期保持。

4.2 研究问题二: 心理需求满足对语音学习表现的作用机制

研究 2 以 SDT 为分析框架,探讨学习者在使用“汉音语声”过程中的心理体验及其与语音表现的关系。结果表明,数字化干预并不仅仅提供了技术支持,其在一定程度上也满足了学习者的基本心理需求,从而影响了语音学习的投入与表现。

独立样本 t 检验显示, 实验组在自主性 (Cohen's $d=2.28$) 和胜任感 (Cohen's $d=2.37$) 上的得分显著高于对照组, 并与干预后的语音成绩呈正相关 (自主性 $r=0.381$, 胜任感 $r=0.413$)。这一结果与已有研究的观点一致: 在传统课堂中, 学习者可能由于焦虑或担心在公开场合犯错而减少发音尝试; 而数字化环境为学习者提供了更低压力的练习空间, 有助于降低情感过滤效应。值得注意的是, 胜任感与短期进步之间的相关性较为突出 ($r=0.306$), 提示可视化反馈和等级式提示可能增强了学习者对自身能力的认知, 使其更愿意在短时间内投入更多精力进行调整与练习。

在关联性维度上, 本研究也观察到一个值得进一步讨论的现象: 尽管实验组的关联性总体得分高于对照组, 但在“依赖教师反馈”这一条目上, 对照组得分更高。这可能反映出两组学习者在学习过程中依赖对象的差异。对照组更倾向于依赖教师的直接反馈, 而实验组则更多利用系统提供的即时提示来进行自我监控。需要指出的是, 目前的数据尚不足以说明这种依赖对象变化所带来的长期影响, 其背后的心理机制仍需在后续研究中结合质性材料进行更详细的探讨。

尽管 SDT 总分及其各维度与短期表现有关, 但均未能显著预测长期保持。换言之, 动机因素可能主要影响学习的起始阶段, 尤其在帮助学习者克服初期困难、建立信心方面作用明显; 但语音能力能否在较长时间内保持, 可能更依赖练习的持续性及加工过程的自动化。这一结果提醒我们, 在数字化教学的后期阶段, 培养良好学习习惯、增强汉语语音练习可能比提升学习动机更为关键。

5 结语

研究围绕汉语语音习得中的关键难点, 采用准实验与纵向追踪的综合研究设计, 系统检验了“汉音语声”小程序在汉语语音教学中的应用效果, 并结合 SLM 与 SDT 探讨其效能背后的认知与心理机制。研究结果显示, 数字化工具在短期提升与长期保持两个维度均优于传统教学, 其优势主要体现在提供更明确的声学线索、更及时的反馈以及更充分的自我监控机会。对学习过程的分析表明, “汉音语声”将传统课堂中难以捕捉的声学差异以可视化方式呈现, 并通过即时纠错与多轮尝试有效降低母语感知过滤的影响, 推动学习者逐步建立稳定的目标语音范畴; “感知—产出”循环训练在一定程度上弥补了常规教学中个性化指导不足的缺陷, 使学习路径更具连续性与可追踪性。动机层面, 学习者在自主性与胜任感等方面的积极体验提高了其持续投入的意愿, 而实验后期语音表现的稳定提升, 则更多受到内化的练习节奏与加工习惯的支撑。总体来看, 数字化语音工具的价值不在于替代传统课堂, 而在于为学习过程提供更具结构性的支持, 并在不同学习阶段发挥不同作用, 其效果更宜被理解为一种促进学习迁移与能力保持的长期支持机制。

尽管如此, 本研究仍存在一定局限。受样本规模与学习者水平的限制, 其结果在其他母语背景及不同熟练度群体中的适用性仍需进一步验证; 研究关注的重点在感知与产出两个维度, 对于语音能力能否在真实交际中实现迁移尚未涉及; 纵向追踪的时间跨度相对有限, 难以呈现语音保持的长期变化。此外, 学习过程中的行为数据并未纳入分析, 这些因素可能与学习成效存在重要关联。未来, 研究可进一步扩大样本范围, 比较不同背景学习者的反应差异, 并将语音能力置于任务型或交际情境中考察其迁移效果; 同时结合学习行为轨迹, 构建更细致的过程模型, 以深化对数字化语音教学作用机制的理解, 为国际中文教育中语音教学的数字化探索提供更具深度的理论与实践依据。

致谢

本研究得到 2025 年河南省哲学社会科学规划项目“河南高校教师生成式人工智能素养构建与提升机制研究”(2025CJY074) 和 2025 河南省国际中文教育研究与实践课题“数字技术赋能汉语语音教学工具的开发与应用研究”(HZC2025KT042) 的支持。

参考文献

- Deng, Dan (邓丹). (2014). 跨语言语音相似度与日本学习者对汉语 /ts/ /tsʰ/ /t/ 三组辅音的感知和产出研究. [A Study on Cross-Linguistic Phonetic Similarity and the Perception and Production of the Three Consonants /ts/, /tsʰ/, and /t/ in Mandarin by Japanese Learners]. 世界汉语教学 [Journal of Chinese Language Teaching], 28(3), 393-410. <https://doi.org/10.13724/j.cnki.ctiw.2014.03.009>
- Deng, Dan (邓丹), Tan Kunming(谭坤明). (2023). 母语背景对二语学习者汉语轻声感知的影响研究 [The Influence of Native Language Background on L2 Learners' Perception of Chinese Neutral Tone]. 中国语音学报 [Chinese Journal of Phonetics], (2), 3-12.
- Deng, Dan (邓丹), Lin, Yujing (林雨箐). (2017). 感知训练方法在汉语语音教学中的应用研究 [A Study on the Application of Perceptual Training Methods in Chinese Phonetics Teaching]. 云南师范大学学报 (对外汉语教学与研究版) [Journal of Yunnan Normal University (Teaching and Research on Chinese as a Foreign Language)], 15(3), 21-27. <https://doi.org/10.16802/j.cnki.ynsddw.2017.03.003>
- Deng, Dan (邓丹), Zhu, Lin (朱琳). (2019). 二语学习者汉语普通话轻声的感知与产出 [Perception and Production of the Neutral Tone in Mandarin Chinese by L2 Learners]. 语言教学与研究 [Language Teaching and Linguistic Studies], 5, 13-24. <https://doi.org/10.13724/j.cnki.yyjx.2019.05.002>
- Jiang, Liping (姜丽萍), Wang, Li (王立). (2021). 智慧教育视域下中文学习平台的构建: 特征、功能与实现路径 [Construction of a Chinese Learning Platform from the Perspective of Smart Education: Characteristics, Functions and Implementation Paths]. 国际中文教育 (中英文) [International Chinese Language Education (Chinese-English Bilingual)], 4, 91-99.
- Li, Baomin (李宝敏), Gong, Lingling (宫玲玲), & Zhu, Zhiting (祝智庭). (2018). 在线学习力测评工具的开发与验证 [Development and Validation of an Online Learning Competency Assessment Tool]. 开放教育研究 [Open Education Research], (3), 77-84, 120. <https://doi.org/10.13966/j.cnki.kfjyyj.2018.03.009>
- Ma, Qiuwu (马秋武), Zhao, Yonggang (赵永刚). (2017). 音系学、语音学与语音教学 [Phonology, Phonetics and Phonetics Teaching]. 北京第二外国语学院学报 [Journal of Beijing International Studies University], 4, 40-55, 131. <https://doi.org/10.12002/j.bisu.106>
- Ma, Qiuwu (马秋武), Zhai, Haiying (翟海莹). (2024). 动态语音教学: 国际汉语语音教学的有效手段 [Dynamic Phonetics Instruction: An Effective Method in International Chinese Phonetics Teaching]. 汉语学习 [Chinese Language Learning], (1), 73-83.
- Mei, Li (梅丽). (2025). 第二语言语音感知训练研究的回顾与展望 [A Review and Prospect of Research on Second Language Speech Perception Training]. 华文教学与研究 [Teaching Chinese to Speakers of Other Languages Studies], (4), 96-105. <https://doi.org/10.16131/j.cnki.cn44-1669/g4.2025.04.002>
- Shi, Feng (石锋). (2022). 作为第二语言的汉语语音教学研究 [Research on the Teaching of Chinese Phonetics as a Second Language]. 国际中文教育 (中英文) [International Chinese Language Education], (3), 3-4.

- Shi, Feng (石锋), Wen, Baoying (温宝莹). (2023). 试论第二语言语音教学的原则和方法 [On the Principles and Methods of Second Language Phonetic Instruction]. *天津师范大学学报 (社会科学版)* [Journal of Tianjin Normal University (Social Science Edition)], (1), 14–20. <https://doi.org/10.20247/j.issn1671-1106.2023.01.003>
- Cook, J. (2010). Mobile phones as mediating tools within augmented contexts for development. *International Journal of Mobile and Blended Learning*, 2(3), 1–12. <https://doi.org/10.4018/jmbl.2010070101>
- Chiu, T. K. (2022). Applying the self-determination theory (SDT) to explain student engagement in online learning during the COVID-19 pandemic. *Journal of Research on Technology in Education*, 54(sup1), S14–S30. <https://doi.org/10.1080/15391523.2021.1891998>
- Dalman, M. (2025). Second Language (L2) Learners' Perceptions of Online-Based Pronunciation Instruction. *Languages*, 10(4), 62. <https://doi.org/10.3390/languages10040062>
- Dorsah, P. (2026). The Use of Cronbach's Alpha Reliability in Educational Research: A Systematic Review. *European Journal of Contemporary Education and E-Learning*, 4(2), 39-50. [https://doi.org/10.59324/ejceel.2026.4\(2\).04](https://doi.org/10.59324/ejceel.2026.4(2).04)
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 233–277). <https://doi.org/10.12691/education-11-2-2>
- Flege, J. E., & Bohn, O. S. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3–83). Cambridge University Press. <https://doi.org/10.1017/9781108886901.002>
- Hao, C., & Lv, M. (2023). Review of the experimental research on tones in Chinese as a foreign language in the past two decades. *Modern Linguistics*, 11(1), 326–334. <https://doi.org/10.12677/ML.2023.111046>
- Li, L., Valcke, M., Badan, L., & Anderl, C. (2022). Chinese as a foreign language (CFL) teachers' pedagogical content knowledge in teaching Chinese pronunciation. *Language Teaching Research*. <https://doi.org/10.1177/13621688221117605>
- Li, X., & Alharbi, W. M. H. (2025). Impact of digital feedback, self-efficacy, and autonomy on motivation and general English performance in online courses. *Learning and Motivation*, 90, 102121. <https://doi.org/10.1016/j.lmot.2025.102121> [Get rights and content](#)
- Lietaert, S., Roorda, D., Laevers, F., Verschueren, K., & De Fraine, B. (2015). The gender gap in student engagement: The role of teachers' autonomy support, structure, and involvement. *British Journal of Educational Psychology*, 85(4), 498–518. <https://doi.org/10.1111/bjep.12095>
- Mushangwe, H. (2015). Using voice recognition software in learning of Chinese pronunciation. *The Journal of Language Learning and Teaching*, 5(1), 52–67.
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development and wellness*. Guilford Press. <https://doi.org/10.1521/978.14625/28806>
- Sinyagovskaya, D., & Murray, J. T. (2021). Augmented reality in Chinese language pronunciation practice. In *ISMAR-Adjunct* (pp. 403–408). IEEE. <https://doi.org/10.1109/ISMAR-Adjunct54149.2021.00092>
- Sheldon, K. M., & Niemiec, C. P. (2006). It's not just the amount that counts: Balanced need satisfaction also affects well-being. *Journal of Personality and Social Psychology*, 91(2), 331–341. <https://doi.org/10.1037/0022-3514.91.2.331>
- Shi, L., Li, S., & Xing, J. (2025). Exploring Chinese secondary EFL students' self-regulated learning and task engagement in AI-assisted classrooms: A latent growth curve modelling study. *European Journal of Education*, 60(4), e70241. <https://doi.org/10.1111/ejed.70241>

- Valk, J. H., Rashid, A. T., & Elder, L. (2010). Using mobile phones to improve educational outcomes: An analysis of evidence from Asia. *International Review of Research in Open and Distributed Learning*, 11(1), 117-140. <https://doi.org/10.19173/irrodl.v11i1.794>
- Watanabe, Y., Omae, T., & Odo, S. (2020). Teaching online Chinese pronunciation with pronunciation training software: An empirical study. In *Online Chinese Teaching and Learning in 2020* (p. 234).
- Zhang, L., & Liu, S. (2020). Acoustic analysis of Chinese tone production by Thai-speaking learners of L2 Chinese. *Journal of Second Language Studies*, 3(2), 180–204. <https://doi.org/10.1075/jsls.19035.zha>
- Zhou, T., Chen, Y., Li, X., Yang, L., & Zhao, Z. (2025). The relationship between online learning self-efficacy and learning engagement: the mediating role of achievement motivation and flow among registered nurses. *Frontiers in Psychology*, 16, 1629174. <https://doi.org/10.3389/fpsyg.2025.1629174>

投稿: 2025 年 12 月 13 日; 接受: 2026 年 3 月 8 日; 出版: 2026 年 4 月 15 日

作者简介

石柳, 香港理工大学硕士、香港教育大学博士, 现为华北水利水电大学外国语学院汉语国际教育系副主任, 讲师。主要从事数字技术赋能国际中文教育的综合性研究。

赵文君, 博士, 华北水利水电大学外国语学院教学办公室副主任、讲师。研究方向为韩国文化、国际中文教育。

王譔翹, 博士, 华北水利水电大学外国语学院党政办公室主任、讲师。研究方向教育学、国际中文教育。

Research on the Driving Mechanism of Chinese Pronunciation Learning Effectiveness Based on the *HanPhonic* Mini Program

Liu Shi

Wenjun Zhao

Xuanxuan Wang

North China University of Water Resources and Electric Power, China

Abstract

Grounded in the perception–production teaching paradigm and guided by Self-Determination Theory (SDT), this study examines the effectiveness of the Hanyin Yusheng mini program in enhancing second language (L2) learners' acquisition of Chinese phonetics. Seventy learners of Chinese were randomly assigned to an experimental group (n=35) and a control group (n=35) and participated in a 16-week quasi-experimental design comparing mini program–integrated instruction with traditional face-to-face teaching. Results from linear mixed-effects models, piecewise growth curve analyses, and moderated mediation analyses showed that the experimental group significantly outperformed the control group in both short-term performance and long-term retention, with a lower rate of decay. Mechanistic analyses further revealed that autonomy moderated the relationship between tool use and immediate progress ($\beta=1.409$, $p=0.006$), relatedness facilitated sustained progress ($\beta=1.623$, $p<0.001$), and competence was strongly associated with short-term gains ($r=0.306$, $p<0.01$). Overall, the findings validate the perception–production linkage proposed by the Speech Learning Model (SLM) in a digital learning context and highlight the mediating role of SDT's core psychological needs in shaping learning outcomes, offering theoretical and practical implications for technology-enhanced Chinese phonetic instruction.

Keywords

Han Phonic Mini program, perception–production teaching, phonetic learning effectiveness, driving mechanism

Liu Shi holds a Master's degree from The Hong Kong Polytechnic University and a PhD from The Education University of Hong Kong. She is currently a lecturer and Deputy Director of the Department of Teaching Chinese to Speakers of Other Languages at the School of Foreign Languages, North China University of Water Resources and Electric Power. Her research focuses on digital technology–empowered International Chinese Language Education.

Wenjun Zhao, PhD, is a lecturer and Deputy Director of the Teaching Affairs Office at the School of Foreign Languages, North China University of Water Resources and Electric Power. Her research interests include Korean culture and International Chinese Language Education.

Xuanxuan Wang, PhD, is Director of the Party and Administrative Office and a lecturer at the School of Foreign Languages, North China University of Water Resources and Electric Power. Her research interests include pedagogy and International Chinese Language Education.