

# 自适应英语发音学习绘本的交互设计

卜瑶华, 鲁晓波

(清华大学 美术学院, 北京 100084)

**摘要:** **目的** 根据学习者的学习特征, 运用自适应学习技术, 结合数字化交互手段, 探究如何对儿童英语发音学习绘本进行交互设计, 以增加儿童对学习内容的吸引力, 并提升儿童的英语发音水平。**方法** 分析目前英语发音学习绘本存在的设计痛点, 从自适应学习、自适应训练、自适应测评和自适应反馈 4 个方面对儿童英语发音学习绘本进行设计实践与反思。**结论** 儿童自适应发音学习绘本利用了数字媒体的特性, 通过多感官交互设计, 激发了学习者的学习兴趣。此外, 儿童自适应发音学习绘本综合考虑了学习者的发音水平、学习动机、学习风格等因素, 为学习者提供了个性化的发音课程、情景互动式的发音训练、精准化的发音检测和具有教学表现力的发音纠正反馈, 实现了真正意义上的自适应学习。

**关键词:** 自适应学习; 人工智能; 个性化学习; 交互设计

**中图分类号:** TB472 **文献标识码:** A **文章编号:** 1001-3563(2022)04-0293-09

**DOI:** 10.19554/j.cnki.1001-3563.2022.04.037

## Interactive Design of Adaptive English Pronunciation Learning Picture Books

BU Yao-hua, LU Xiao-bo

(Academy of Arts & Design, Tsinghua University, Beijing 100084, China)

**ABSTRACT:** According to the learning characteristics of learners, using adaptive learning technology, combined with digital interactive methods, this paper explores how to interactively design picture books for children's English pronunciation learning to increase children's attraction to the learning content and improve children's English pronunciation. This paper analyzes the design pain points of the current English pronunciation learning picture books, and practice and reflect on the design of children's English pronunciation learning picture books from four aspects: adaptive learning, adaptive training, adaptive assessment and adaptive feedback. Children's adaptive pronunciation learning picture books take advantage of the characteristics of digital media and stimulate learners' interest in learning through multi-sensory interaction design. Furthermore, the children's adaptive pronunciation learning picture books take into account the learners' pronunciation level, learning motivation, learning style, and other factors to provide personalized pronunciation learning courses, situational interactive pronunciation training, accurate pronunciation detection, and expressive pronunciation corrective feedback, realizing adaptive learning in the real sense.

**KEY WORDS:** adaptive learning; artificial intelligence; personalized learning; interactive design

初级自适应学习研究起源于人工智能运用, 并在 20 世纪 70 年代开始流行。随着人工智能技术的爆发, 人工智能自适应学习研究成为了学术研究的热点, 机器学习之父 MITCHELL T 在 2018 全球 AI+智适应教

育峰会中表示:“最近这 10 年对自适应教学来说是最好的 10 年, 因为我们开始看到技术变得成熟。未来机器学习和人工智能, 将成为自适应学习的驱动式技术。”

收稿日期: 2021-11-21

基金项目: 国家重点研发计划“中国风格文化创意及智能产品设计技术集成与应用示范”(2019YFB1405700)

作者简介: 卜瑶华(1993—), 女, 湖南人, 清华大学美术学院博士生, 主要研究方向为交互设计、交互艺术等。

通信作者: 鲁晓波(1959—), 男, 湖南人, 清华大学美术学院教授, 博士生导师, 主要研究方向为信息艺术设计、工业设计教学, 以及相关领域的理论研究、艺术创作和设计实践等。

英语绘本是有效提升英语综合水平的儿童读物,将人工智能自适应学习技术应用在儿童绘本的研究较少,以提升儿童英语发音为目的的自适应学习绘本设计研究更是一个崭新的方向。发音学习绘本设计综合考虑了学习者的认知水平、情感因素、个性差异、学习风格等因素,通过不断挖掘学习者的学习需求,实现了为学习者推荐个性化的教育内容、多感官的交互学习模式、可持续的教学路径规划与学习效果的预测与反馈。自适应发音学习绘本通过多感官的交互手段及智能化的自适应学习模式,不仅培养了儿童学习者发音学习的兴趣,还提高了学习者的综合英文水平,如英语发音标准度、英语口语沟通能力、英文阅读能力等。因此,探究如何设计英语发音自适应学习绘本有着极强的学术研究与应用价值。

## 1 研究背景

### 1.1 儿童发音学习的特点与难点

儿童学习者相比成人在英语口语学习上具有很大的优势<sup>[1]</sup>,有观点认为,从大约2岁到青春期的关键时期内,儿童相比成人可以更轻松地学习第二语言<sup>[2]</sup>。还有研究表明,早期接触第二语言确实可以比晚期接触语言更准确地“感知”和“说”第二语言<sup>[3]</sup>。儿童对语言学习具有较强的敏感性,他们的语言感受力强、模仿能力优秀<sup>[4]</sup>、且具有可塑性。儿童在发音学习中善于模仿老师的发音动作和声音,有着极强的学习能力。同时,儿童擅长机械记忆,反复、多次的跟读训练对儿童发音学习有着促进作用。此外,儿童喜欢趣味性的课程机制,在课程中引入有趣的交互元素,更容易被孩子接受。研究发现儿童更喜欢具有教学表现力的发音教学<sup>[5]</sup>,如教师使用清晰的语音、夸张的口腔发音运动、配合丰富的表情及肢体动作来感染学习者。具有表现力的教学反馈也是教师纠正错误发音的有效策略,当儿童学习者出现发音错误时,老师可以通过夸张的发音动作和重音来凸显学习者的发音错误。ALGHAMDI N<sup>[6]</sup>等人已经证明,口腔发音运动的视觉夸张可以提高被试在感知训练中的视听语音识别能力。

发音教学在儿童英语教学中是至关重要的,如果在儿童阶段没有养成一个良好的发音习惯,可能会造成学习者长期性的发音错误,从而限制了学习者英语综合水平的提高。英语发音学习、纠正的困难主要有2个方面,一方面,双语学习者在学习第二语言时更加倾向于母语的发音,譬如母语为粤语的第二语言学习者倾向于发“th”音为“f”音,因为“th”音在粤语中是缺失的,这种现象被称为语言迁移的负影响<sup>[7]</sup>,这会导致学习者难以意识到自身的发音错误,使其纠音很困难。另一方面,缺少专业的英语为母语的口语教师。根据英国文化协会提供的统计数据,全球虽然大约有

15亿英语学习者,但只有25万母语人士有资格担任英语教师。据报道,加拿大67%的英语教师没有接受过发音教学培训,而在欠发达地区,母语为英语的教师更加缺乏<sup>[8]</sup>。

### 1.2 儿童用户阅读电子绘本的特征

儿童发育学习绘本主要针对4~8岁的儿童,这个年龄段儿童获取知识的方式,主要来源于对生活中的认识,以及对世界万物的观察模仿,因此电子绘本成为了儿童获取认知的一个重要窗口,绘本所传达的爱与正直的观点,将影响到儿童的身心发展。根据儿童善于语言模仿与表达的特点,可以鼓励儿童在看绘本故事的同时,与绘本本身进行语言沟通,从而培养儿童理解、模仿和表达故事的能力。儿童在绘本阅读中喜欢多感官的互动,电子绘本通过视听触觉上的合理感官刺激,可以有效地激发儿童阅读的兴趣,提升阅读的关注度。此外,由于4~8岁儿童的整体认知水平不够成熟,所以需要给予儿童交互引导,辅助儿童流畅且直观地完成学习、阅读等操作。儿童喜欢趣味性的元素,当电子绘本融入具有吸引力的游戏化元素时,将会对儿童产生用户黏性。然而,如果趣味性元素没有很好地与学习内容结合,将无法达到提升儿童的学习效果。

### 1.3 人工智能自适应学习的定义与特点

学习是一个复杂且隐性的过程,由于学习主体的多样性和差异性,基于人工实现的自适应学习或基于简单计算机规则实现的初级自适应学习,已无法满足信息社会对教育发展的新需求。随着物联网、云计算、大数据、人工智能、数字孪生等信息技术的快速发展,以及我国对教育信息化发展的政策支持,人工智能自适应教育遇到了难得的历史发展机遇。人工智能自适应学习被应用于不同的教育领域,并不断细分发展,取得了丰富的研究成果,如根据学生学习和认知风格的个性化推荐系统<sup>[9]</sup>,自我学习评价系统<sup>[10]</sup>,在线自适应学习异常检测系统<sup>[11]</sup>,协同学习系统<sup>[12]</sup>等。基于人工智能算法的自适应学习不仅是对传统自适应学习的升级,也是对新型学习交互方式的探索,从而满足日益复杂的远程教育的新需求,平衡教育资源分布不公平的社会现状,改善新冠疫情导致的线下教育受阻的现状,实现学校、家庭和社会“三位一体”的教育网络。

自适应(Adaptive),是指自我调节和匹配。自适应学习(Adaptive Learning)<sup>[13]</sup>是基于人工智能技术,根据学习者的知识水平、认识风格等特征,为每位学习者安排并最优、最佳的学习内容,并对学习者的学习状态和学习结果进行实时检测和可持续的分析,以提供自适应的教学反馈。自适应学习通过对学习者的行为数据的深入挖掘,实时调整学习内容、评估方式和知识序列,真正实现了个性化的教育理念<sup>[14]</sup>。

人工智能自适应教育的实现过程包括：确定设计目标与搜集学习数据、构建学习模型和输出学习反馈。搜集大数据是指采用人工智能技术，如分类树、贝叶斯网络、深度学习、神经网络、遗传算法等技术分析学习行为的大数据(如学习者的语音、做题结果、学习时间、学习时的情感状态等)。构建学习模型是人工智能自适应实现的重点，学习模型以教学目的分类主要包括专家模型、教学方法与风格模型、学生模型等。模型的训练结果将输出到学生反馈模块，包括给学生推荐自适应的学习教材、输出学习掌握程度的测评、预测学生学习效果及规划后期学习知识点等。一个完整的人工智能自适应学习系统应包括如下组成部分：提供教学内容、实施教学策略和提供评估学生学习进度的机制。在理想的情况下，自适应学习系统为学生提供大规模高质量学习为目的，帮助学习者开展基于任务的学习，精准评估学习者的表现，帮助教师、学校和教育机构动态调整课程，检测与预防学习者的异常心理<sup>[15]</sup>，并将教学重点放在真正需要之处。

## 2 自适应英语发音学习系统的相关研究

### 2.1 自适应英语发音学习系统

关于自适应学习系统的分类，主要包括 2 个类别：智能教学系统(Intelligent Tutoring Systems, ITS)和自适应超媒体(Adaptive Hypermedia Systems, AHS)<sup>[16]</sup>。ITS 系统通过利用人工智能技术的优势模拟教师和学生之间的一对一学习过程，为学生提供及时且智能的学习指导与反馈。完整的 ITS 系统应包括

如下组成部分：呈现学习内容、实施教学策略及评估学生的学习进度。自适应超媒体系统(AHS)将超媒体和自适应教学系统结合起来，自适应超媒体系统应该满足 3 个标准：属于超文本或超媒体系统；具有用户模型；利用用户模型进行自适应推荐。

自适应学习在英语发音学习领域发展较早，拥有大量的英语学习材料、标准化测评模型和英文语料库。在 20 世纪 70 年代引入的计算机辅助英语发音训练系统是最早的、具有低自适应水平的学习系统。关于计算机辅助英语学习的智能教学系统的开创性的出版物，是 SWARTZ M L 等人于 1992 年编辑的《外语学习智能教学系统》<sup>[17]</sup>，该研究从多学科视角进行研究，结合了智能教学系统和自然语言处理来构建语言学习的教学框架，讨论语法检查、错误分析、用户建模等几种智能方法，探索如何对技术进行调整和组合来实现更有效的语言学习系统。2000 年左右，英语发音学习系统侧重于语音识别技术的研究。随着人工智能的发展，英语发音自适应学习系统逐步引入了多种技术，包括自然语言处理(Natural Language Processing, NLP)、自然语言生成(Natural Language Generation, NLG)、自动语音识别(Automatic Speech Recognition, ASR)、错误发音检测与诊断(Mispronunciation Detection and Diagnosis, MDD)、表现力语音合成、可视语音合成等。英语发音自适应学习系统的研究注意力放在了英语发音学习、训练、检测和反馈的全过程，并加入多样化的交互手段，如利用虚拟现实技术搭建的沟通场景等。英语自适应系统的相关研究成果，见表 1。

表 1 英语自适应学习系统相关研究成果  
Tab.1 Research results of English adaptive learning system

文献来源	系统描述	教学方法	技术
系统名称 “Talk to Me”	程序与用户进行对话，程序提出问题，用户语音回答问题，如果程序识别出用户回答了正确答案，将进入“对话”的下一阶段	对话沟通训练	自动发音识别技术
SHEI C C 等人 (2001 年) <sup>[18]</sup>	基于学习者输入的文本自动生成适合学习者特定需求的语言课程	个性化课程	用户模型与自适应学习
UTHER M 等人 (2005 年) <sup>[19]</sup>	旨在帮助日语使用者在感知上区分/r/和/l/音素，因此设计了移动端的自适应语言学习程序	移动自适应学习	移动自适应学习
FISHER T 等人 (2009 年) <sup>[20]</sup>	比较了使用纸质词典书、词典电子书和自适应学习词典电子书对英语词汇习得的影响，自适应学习词典电子书为学生提供一种引人入胜的方式，从而提高他们的外语水平	自适应学习	自适应学习
YUEN K W 等人(2011 年) <sup>[21]</sup>	计算机辅助发音学习的移动网络平台，包含检测和诊断发音错误的语音识别器，1 个语音合成器和 1 个视动画制作器	预定义的课程 自我实践课程	自动发音识别技术， 表现力音视频合成
SFENRIANTO S 等人(2018 年) <sup>[22]</sup>	系统对学习者进行测试以确定学习者的熟练程度(包括初级、中级和高级)，系统根据学习者的熟练程度提供相匹配的英语学习材料	自适应课程	自适应学习
BU YAOHUA 等人(2021 年) <sup>[23]</sup>	具有夸张表现力的听视觉多模态融合的发音识别与发音反馈虚拟老师纠音系统	参与式戏剧教学 夸张表现力教学	错误发音识别技术， 夸张表现力音视频合成

## 2.2 自适应英语发音学习绘本

文献调研发现英语发音领域的自适应学习研究较多,但是自适应学习与电子绘本结合的研究相对较少<sup>[24]</sup>,以英语发音学习为目的的自适应学习绘本的研究更是一个全新的方向,具有广阔的研究空间。以下调研了几款应用于国内学习市场的英语发音自适应学习绘本,其在系统设计中应用了不同程度的自适应学习技术和灵活的交互手段。

“点读精灵”是一款配合纸本英文阅读绘本的智能点读笔,点读笔可以识别英文文字,引领孩子跟读英文绘本,并对儿童跟读的发音进行打分,让孩子在大胆开口说英语的同时纠正发音。“点读精灵”依托于香港中文大学的语音纠错技术,可精准识别音素级别的发音反馈,但是该产品从学习者体验角度仍需要完善,譬如如何增加英文跟读过程的趣味性。

“魔岛英语”是一个专为4~8岁孩子设计的英语语伴应用,让孩子在实际对话任务中学习英语表达。故事虚拟人在对话任务中主动和学习者进行英文语音交流,然后对儿童的回答进行关键词识别,从而判断儿童是否成功完成沟通对话任务。该应用通过互动情景设计,很好地激励了儿童英语沟通的积极性,但是该系统的自适应学习程度还有进一步优化的空间。

“多纳学英语”是专为3~7岁儿童设计的英语启蒙应用,它的功能设计内容包括:分级阅读、情景英语、英文儿歌、双语绘本、场景化游戏课程等。该应用具有完整的“学、练、用、测”环节,并应用语音识别技术检测儿童英语发音的准确度,是一款功能较全面的自适应学习绘本。

“伴鱼绘本”对绘本难度从学龄前到小学毕业进行了分级,实现了分级推荐学习内容。该绘本还采用了一些功能化设计包括:支持孩子自己给绘本配音,创造自己的专属绘本;采用口语测评人工智能模型给儿童的英文配音打分,根据儿童读音特征设计评分标准。该工作实现了相对较高程度的自适应学习标准。

以上的应用产品设计对英语发音学习绘本的研究有着较高的参考价值。然而,现有的大部分自适应

学习绘本设计忽视了一些问题,具体包括以下2个方面。

1) 大部分的自适应学习绘本在设计上注重给人留下深刻印象的交互方式,通常是为了趣味性和吸引大众市场<sup>[25]</sup>,而没有重视如何在教学上安排合理的学习内容,提升儿童对绘本所传达知识的理解能力。

2) 人工智能自适应技术与绘本阅读结合是一个新兴的领域,现有的大部分自适应学习绘本没有很好地应用人工智能技术来满足自适应教育的需求。面向儿童发音学习的自适应学习绘本的设计应根据学习者的学习需求和学习特点,提供合适的学习与反馈,以实现真正意义上的自适应学习目标。

## 3 自适应英语发音学习绘本的交互设计

从自适应学习、训练、测评和反馈的角度,分别阐述如何实现儿童自适应发音学习绘本的交互设计框架,见图1。其中,学习模块安排了个性化的发音学习内容;训练模块完成了情景式口语沟通的场景搭建;自适应测评模块实现了实时发音检测与发音数据分析;自适应反馈模块提供了具有教学表现力的虚拟老师教学反馈与长时间的发音学情报告。

### 3.1 自适应学习

#### 3.1.1 学习材料设计

关于自适应学习绘本的文字设计,需要设计符合儿童的认知水平,有助于提高口语能力的,通俗易懂的语句或对话作为绘本内容。绘本的单词难度不宜过难,故事文本的生词数应根据学生的英语水平而设置。关于英语学习绘本故事的句型,可以设计重复的句型结构,便于学生理解和重复记忆。如英文著名童谣《The More We Get Together》里重复出现的单词“together”,不仅方便读者吟唱,而且渲染了温馨的氛围。另外还可以选择具有韵律美的语言来描述绘本故事,使用押韵的短句来描绘故事。如我国著名语言学家、“现代语言学之父”赵元任先生,于1930年在美国写的一篇同音文《施氏食狮史》,文章较多地使



图1 应用于英语发音学习的儿童自适应学习绘本的交互设计框架

Fig.1 Interactive design framework of children's adaptive learning picture books for English pronunciation learning

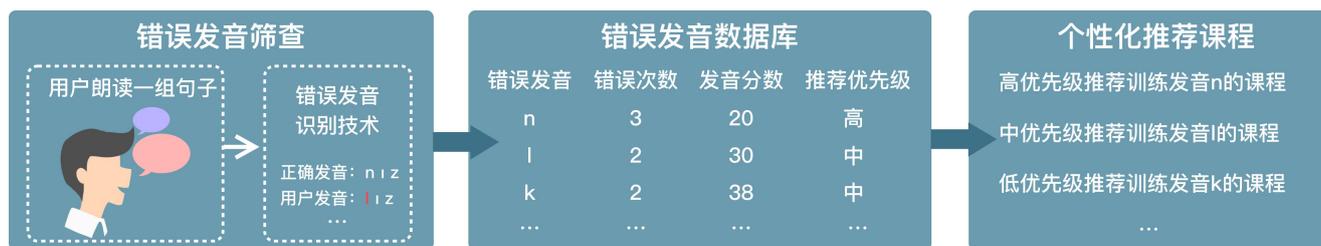


图 2 根据学习者的发音错误推荐个性化课程内容

Fig.2 Personalized course recommendation according to learners' pronunciation errors

用一组同音的字母来设计故事的文本,不仅使儿童跟读起来琅琅上口,同时具有针对性和趣味性。关于语言长度的选择,针对低年级的孩子主要选用简洁的单词或词组,针对高年级的孩子可选择多个句子组成的短文。

关于自适应学习绘本的交互元素设计,需要设计大量符合儿童审美、生动活泼的图片和动画。利用电子绘本的优势与特性,可以增加大量多感官交互元素,充分调动儿童学习者的视触听觉感官,在吸引儿童在阅读的过程中潜移默化学习知识,获取对世界的认知。如儿童学习绘本《牛顿和我》,精美的绘本画面使它成为 Arbordale 出版社非常受欢迎的作品。故事主人公是一个男孩,当男孩与他的宠物狗牛顿玩耍后,他发现了的日常活动中力和运动的规律。学习者可以动手操控画面中的小球发生运动,在通过触觉控制小球的过程中加深儿童对基础物理概念的认知。

关于自适应学习故事的故事结构设计,尽可能选择简洁且清晰的框架。绘本主题以贴近生活、轻松快乐的儿童主题为主,譬如可以讲述童话故事或生活中的趣事。若选择的故事主题与目标阅读学习者的文化生活环境有很大的差异时需要慎重选择,并在儿童阅读前提供文化背景介绍。此外,绘本的故事要满足发音训练和口语沟通的目的,譬如英语发音学习绘本的故事内容针对某组易混淆的发音而进行设计。

### 3.1.2 自适应内容推荐

儿童发音自适应学习绘本具有极强的针对性,为了符合儿童的认知水平,提升儿童发音学习的效率,需要根据儿童的学习需求、行为、能力、心态和偏好等情况,个性化推荐合适的学习路径和学习内容。以儿童发音学习绘本的设计为例,绘本对发音错误进行筛查,推荐具有针对性的发音学习材料作为绘本课程。具体来说,对首次使用绘本的学习者,系统在筛选模块检测学习者的发音情况,根据学习者错误发音的程度,按照优先级匹配相应的绘本内容。如系统识别到学习者出现的严重发音错误——把发音“n”读成发音“l”,因此优先推荐训练易混淆音“n”和“l”发音的绘本内容,见图 2。学习者每次的发音错误都将被记录,对长期使用绘本的学习者来讲,系统可根据学习者的历史表现,更准确地推荐英语发音课程。

## 3.2 自适应训练

发音教学研究表明,将发音机制与实际语境或沟通目标联系起来是十分可取的<sup>[26]</sup>。通过倡导“情境式沟通场景”,设计沉浸式交互剧情,可以让儿童学习者在沟通的情境中潜移默化地学习知识。在情景式沟通场景训练中,“互动参与式戏剧”(Interactive Participant Drama, IPD)方法和“游戏化任务机制”很适合引入沟通场景中。“互动参与式戏剧”的好处是可以充分调用学习者的感官,激励学习者参与更多感知发音(听)和产出发音(说)的训练。学习者在预先编程的场景中和虚拟人进行沟通,沟通的问题为精心设计好、具有诱导性的问题,见图 3。其中,虚拟人问学习者,图中的 Mike 小朋友搭乘的是什么工具?该问题的目的是激发学习者在看绘本画面的过程中与系统进行英语沟通。“游戏化任务机制”被引入情景式沟通场景训练,将增加发音学习的目的性、趣味性与挑战性。游戏任务机制针对学习者的沟通表现进行判定。若学习者正确回答了虚拟人物的提问,系统利用 MDD 技术,进一步判断学习者的发音是否标准:若发音正确,虚拟人与学习者进行下一场景的语音沟通;若发音错误,发音纠正反馈模块被引入来帮助学习者纠正发音。实现情景式沟通场景涉及搭建一个多维度的情景对话系统技术框架,见图 4。虚拟人与学习者进行语音对话,需要通过语音识别和语言理解技术,分别对学习者语音内容进行识别与理解,然后根据自然语言生成技术生成虚拟人的对话文本,最后根据语音合成技术完成语音合成。

## 3.3 自适应测评

自适应测评是指根据学习者的学习表现,进行准确、客观、高效的教学效果测评。例如,托福口语测评的 Speech Rater 引擎是英语口语测评最广泛的应用之一,旨在对发音者的口语表现进行评分。Speech Rater 的评分范围是开放的,包括言语表达(如发音和流利程度)、语法能力、语言主题的连贯性和思想进步等相关更高水平的评价。自适应测评在教学中提供实时反馈,一方面极大地减轻了教师负担;另一方面为教学决策和教学改革提供真实可靠的依据。以发音训练为目的的自适应学习绘本设计,利用 MDD 技术对儿童的发音情况进行音素级、单词级和句子级别



图3 情境式沟通场景设计  
Fig.3 Scene design of situational communication scenarios

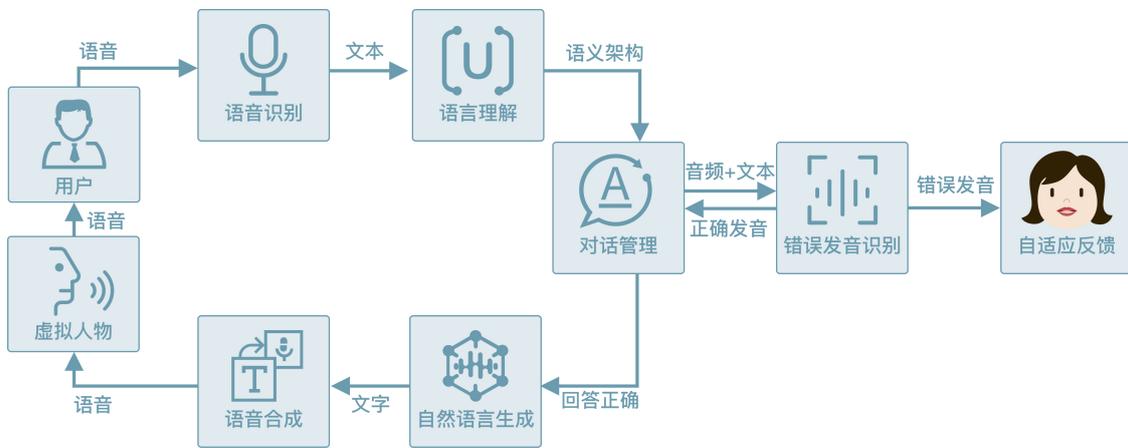


图4 情景式沟通场景的技术框架  
Fig.4 Technical framework of situational communication scenarios

的多维度评价,有助于保障孩子掌握的每一步学习都有学习检测与反馈。针对儿童出现的发音错误,专注发音的细节,即时纠正;针对长时间学习者,对学习者的发音错误进行累积。

### 3.4 自适应反馈

#### 3.4.1 具有教学表现力的实时发音反馈

斯坦福国际研究院学习技术中心战略研究与创新主任菲尔·瓦伊(PHIL VAHEY)提出,无论是个人或协同的学习系统,自适应反馈都是极其重要的。在发音学习中,自适应反馈可以帮助学习者清晰地了解自身的发音情况,通过合理的方式与方法纠正发音。通过调研发现,现有的发音反馈主要通过展示发音教学图(图片)、播放发音(音频)或标注音标(文字)的方法来纠正发音,但是以上的方法对儿童学习者具有局限性。例如,发音教学标准图与具有教学表现力的虚拟老师口腔发音教学示范见图5。其中,发音教学标准,见图5a,对儿童来说缺乏感官吸引力,同时,复杂且专业性强的发音教学图也可能导致儿童学习者识别困难。音频反馈的方式被广泛应用在发音教学中,学习者可以通过对比标准发音和自身发音的

区别了解发音错误,但是对儿童学习者来说,单一的音频反馈同样是缺乏感官吸引力的。此外,更多的发音反馈是通过音标的方式来展示发音,音标标注的方式可以准确地告知学习者哪个发音位置出现了发音错误,但是这样的方法对尚未掌握音标的儿童是无效的。

在线下发音教学中,当学习者难以意识到自身的发音错误时,教师在教学中广泛使用口腔发音动作示范,如清楚的发音(听觉),配合清晰的发音动作(视觉),展示如何正确发音。针对出现严重发音错误且难以纠正的学习者,教师使用更具有表现力的,略带夸张的教学方法来纠正发音:包括口腔发音动作放慢,夸张或停留(视觉),以及语速放慢,重读或改变音调(听觉)来强化错误发音的位置,帮助学习者分别从视听觉的角度意识到自身的发音错误。英语发音学习的儿童绘本在自适应反馈模块可参考老师线下教学的经验,创建虚拟老师来模拟真实老师的发音纠正方法。并针对不同学习者的不同发音水平(如发音准确度、出现同类发音错误的次数等),提供具有不同表现力的发音动作示范。为实现自适应的发音反馈,可通过机器学习的方式,“学习”真人老师对不同发音水平的学生,进行不同程度的表现力教学的视频

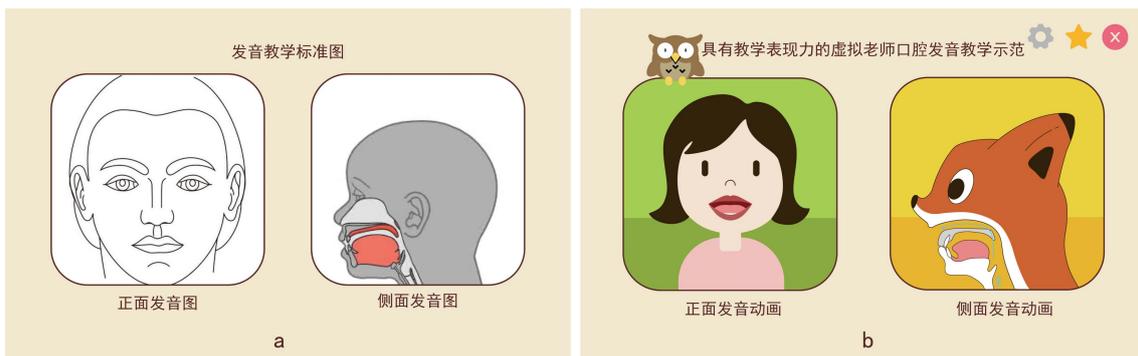


图 5 发音教学标准图与具有教学表现力的虚拟老师口腔发音教学示范  
Fig.5 Pronunciation teaching standard chart and the articulatory movement demonstration of the virtual teacher with teaching expressiveness

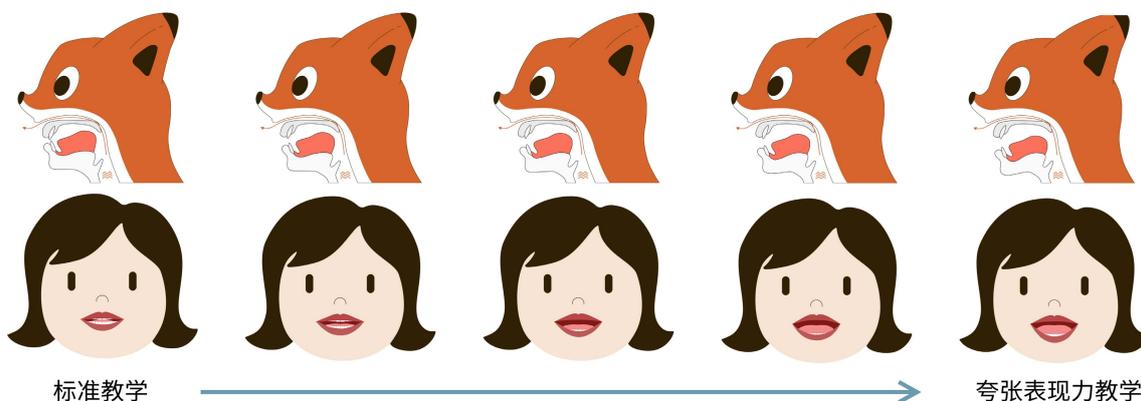


图 6 具有教学表现力的虚拟老师展示口腔发音动作（以张大嘴型为例）  
Fig.6 Articulatory movement of the virtual teacher with teaching expressiveness (open the mouth as an example)

数据，建立不同发音水平的学习者与不同程度的表现力发音教学之间的匹配关系。譬如针对错误程度低，错误频率低的学习者，虚拟老师提供标准的口腔发音动作教学。具有教学表现力的虚拟老师展示口腔发音动作（以张大嘴型为例）见图 6，演示了虚拟老师如何通过逐步夸张的发音动作，来示范张大嘴型的发音动作。具有不同教学表现力的虚拟老师被应用于纠正学习者的发音错误，这能够提升学习者的发音水平<sup>[23]</sup>。值得注意的是，具有表现力的发音动作示范是在符合正确发音运动规律的基础上进行的合理夸张，严格遵循了语音学、音韵学的理论。

关于具有教学表现力的虚拟老师的形象设计，采用儿童喜欢的、具有鲜明色彩且具有亲和力的人物外观，见图 5b，虚拟老师的形象可以是平面或立体的，或采用新颖的 VR 或 AR 来展示。正面角度需要重点展示嘴巴、牙齿和舌头的运动状态，侧面角度需要重点展示重要的口腔发音器官（包含舌头、上下嘴唇、上下牙齿、下巴、小舌）的运动状态。为了让儿童理解发音运动，发音器官的设计需要尽可能简化。

### 3.4.2 长时间的学习报告

为了让学生和家长了解长期的学习效果，包括学习的进度、学习的成果及学习中遇到的难点等，学习报告被设计应用在系统中。学习报告的核心在于识别

学生正在学习什么，哪些没有学会，哪些已经掌握，以及对未来哪些内容需要掌握，从而帮助学生不断评估，直到达到特定的学习目标。在具体的英语发音学习报告中，需要呈现优秀发音的音素、单词和句子；需要改进发音的音素、单词和句子；需要取得进步的音素、单词和句子。发音报告还需要对学习者的未来学习趋势进行预测，鼓励学习者保持积极的学习态度。

## 4 反思与启示

### 4.1 如何降低系统误判造成的教学负面效应

随着技术的发展，运用人工智能手段来检测学习者的学习效果的准确率一直在提高，但现有的自适应学习系统依然会遇到一个问题，若不小心误判了学习者的学习效果，将如何降低错误判断后的教学负面效应。以 MDD 技术为例，尽管现有的 MDD 技术的准确性已经非常高，但反馈仍然可能包含 2 种错误：一种错误是“错误被接受”（False Accepts, FA）和一种错误是“错误被拒绝”（False Rejects, FR）。其中 FA 表示学习者不正确的发音被识别成正确，FR 表示学习者发音正确但被系统认定为错误。显然，应该在 FA 和 FR 之间找到一个平衡的关系，尤其是对 FR 要极其谨慎，因为 FR 可能会误导学生。

以本工作的解决方案为例,为降低FR导致的不利后果的设计,教学与设计方法被引入——在反馈单元设计一个具有教学表现力的虚拟老师“负责”发音纠正。一方面,为了保持儿童的学习积极性,虚拟老师并不会直接告诉儿童出现了哪些发音错误(如给予发音评分),而是鼓励儿童跟随虚拟老师进行发音练习。另一方面,虚拟老师的发音反馈改变了音标、音频或图示等传统纠音方式,而是依据儿童学习者对动画感兴趣的特点,给不同发音水平的儿童学习者呈现不同教学表现力程度的口腔发音动画教学。以上方法不仅强化了儿童学习者对错误发音的视听感知度,还提升了纠音反馈教学的趣味性,并有效降低了系统误判带来的负面影响:如果系统犯了FA的错误,学习者会收到不具有任何表现力的标准发音教学示范;如果系统犯了FR错误,学习者会收到具有表现力的教学反馈。

#### 4.2 如何提升儿童学习者兴趣的同时提升学习效果

应用于发音学习的自适应学习绘本的研究,如何通过有效的交互方式,在增加儿童吸引力的同时,提升儿童的发音水平,是需要持续研究的难点。多样化的交互方式如果运用不当,很可能降低学习的目的性,甚至导致儿童沉迷游戏本身而无法达成学习效果。在未来的工作中,需要对儿童的认知心理和发音教学方法进行更加深入的理论研究,探索儿童在阅读绘本中对哪些交互方式的接受程度更高,交互方式要怎么运用才能提升儿童的兴趣,交互方式要怎么设计才能提升儿童的学习效果等。

## 5 结语

儿童阶段是进行发音训练的重要时期,研究以面向儿童的发音学习为目标,人工智能自适应技术为基础,完成了自适应发音学习绘本的交互设计。该绘本分别从自适应学习、自适应训练、自适应测评和自适应反馈4个阶段依次实现了个性化的课程推荐、场景化的口语训练环境、精准化的发音识别诊断与具有教学表现力的发音纠正反馈。人工智能自适应技术与电子绘本结合是一个全新的研究领域,从交互设计角度研究电子绘本的设计具有较强的实践价值。针对发音学习绘本的交互设计,需要考虑儿童学习者的认知水平,以提升儿童对绘本所传达知识的认知理解力为目标,提供自适应的学习内容,并针对儿童学习者的认知特点,实现交互手段与学习内容之间的密切配合,通过合理的信息呈现方式来提升学习者对内容的感知与学习效果。

#### 参考文献:

[1] KUHL P K, WILLIAMS K A, LACERDA F, et al.

Linguistic Experience Alters Phonetic Perception in Infants by 6 Months of Age[J]. *Science*, 1992, 255(5044): 606-608.

[2] LENNEBERG E H. The Biological Foundations of Language[J]. *Hospital Practice*, 1967(12): 59-67.

[3] Flege JE, MacKay IR. Perceiving vowels in a second language[J]. *Studies in second language acquisition*. 2004, 26(1): 1-34.

[4] 朱吉虹, 赵越超. 基于儿童认知发展的学龄前儿童APP界面交互设计研究[J]. *包装工程*, 2020, 42(10): 55-61.

ZHU J, ZHAO Y. Research on APP Interface Interaction Design for Preschool Children Based on Children's Cognitive Development[J]. *Packaging Engineering*, 2020, 42, 4(10): 55-61.

[5] GRAMATKOVSKI B, KOCHOSKA J, RISTEVSKA M, et al. The Expressive Reading as an Opportunity for Improvement of the Preschool Children's Speech[J]. *International Journal of Innovation and Research in Educational Sciences*, 2017(1): 99-104.

[6] ALGHAMDI N, MADDOCK S, BARKER J, et al. The Impact of Automatic Exaggeration of the Visual Articulatory Features of a Talker on the Intelligibility of Spectrally Distorted Speech[J]. *Speech Communication*, 2017(95):127-136.

[7] MENG H, LOY Y, LAN W, et al. Deriving Salient Learners' Mispronunciations from Cross-language PHONOLOGICAL Comparisons[C]. *Piscataway: Automatic Speech Recognition & Understanding. IEEE Workshop on. IEEE*, 2008.

[8] DERWING T M, MUNRO M J. Second Language Accent and Pronunciation Teaching: A Research - Based Approach[J]. *TESOL Quarterly*, 2005, 39(3):379-397.

[9] VANDEWAETERE M, DESMET P, CLAREBOUT G. The Contribution of Learner Characteristics in the Development of Computer-based Adaptive Learning Environments[J]. *Computers in Human Behavior*, 2011, 27(1): 118-130.

[10] DUFFY M C, AZEVEDO R. Motivation Matters: Interactions between Achievement Goals and Agent Scaffolding for Self-regulated Learning within an Intelligent Tutoring System[J]. *Computers in Human Behavior*, 2015, 52(11): 338-348.

[11] DONG L I, SHULIN L I U, ZHANG H. A Method of Anomaly Detection and Fault Diagnosis with Online Adaptive Learning under Small Training Samples[J]. *Pattern Recognition*, 2017(64): 374-385.

[12] JONG B S, CHAN T Y, WU Y L. Applying the Adaptive Learning Material Producing Strategy to Group Learning. In *International Conference on Technologies for E-Learning and Digital Entertainment*. Springer, Berlin, Heidelberg, 2006(4): 39-49.

[13] PARAMYTHIS A, LOIDL REISINGER S. Adaptive Learning Environments and E-learning Standards[J]. *Electronic Journal of e-Learning*, 2004, 2(1): 181-194.

- [14] 菅保霞, 姜强, 赵蔚, 李勇帆. 大数据背景下自适应学习个性特征模型研究——基于元分析视角[J]. 远程教育杂志, 2017, 35(4): 87-96.  
JIAN B, JIANG Q, ZHAO W, et al. Research on Adaptive Learning Personality Model in the Context of Big Data: from the Perspective of Meta-analysis [J]. Journal of Distance Education, 2017, 35(4): 87-96.
- [15] 景春晖, 支锦亦. 基于大数据的智能教育系统体验设计研究[J]. 包装工程, 2020, 41(2): 21-27.  
JING C, ZHI J. Research on Experience Design of Intelligent Education System Based on Big Data[J]. Packaging Engineering, 2020, 41(2): 21-27.
- [16] ELENA V, REGUERAS L M, MARÍA J V, et al. Is Adaptive Learning Effective? A Review of the Research[C]. Piscataway: In 7th WSEAS Int Conf on Applied Computer & Applied Computational Science, World Scientific and Engineering Academy and Society, 2008: 710-715.
- [17] SWARTZ M L, YAZDANI M, LINGUISTIK, et al. Intelligent Tutoring Systems for Foreign Language Learning[J]. Nato Asi, 1992: 80.
- [18] SHEI C C. Follow You: An Automatic Language Lesson Generation System[J]. Computer Assisted Language Learning, 2001, 14(2): 129-144.
- [19] UTHER M, SINGH P, UTHER J. Mobile Adaptive CALL (MAC): an Adaptive s/w for Computer Assisted Language Learning[C]. New York: ICPS'05. Proceedings of International Conference on Pervasive Services, IEEE, 2005.
- [20] FISHER T. Mobile Learning of Vocabulary from Reading Novels: a Comparison of Three Modes[C]. Florida, Proceedings of 8th World Conference on Mobile and Contextual Learning, Orlando, Florida, 2009.
- [21] YUEN K W, LEUNG W K, LIU P F, et al. Enunciate: An Internet-accessible Computer-aided Pronunciation Training System and Related User Evaluations[C]. Piscataway: In 2011 International Conference on Speech Database and Assessments (Oriental COCODSA), IEEE, 2011: 85-90.
- [22] SFENRIANTO S, HARTARTO Y B, AKBAR H. An Adaptive Learning System based on Knowledge Level for English Learning[J]. International Journal of Emerging Technologies in Learning (IJET), 2018, 13(12): 191.
- [23] BU Y, MA T, LI W, et al. PTeacher: a Computer-Aided Personalized Pronunciation Training System with Exaggerated Audio-Visual Corrective Feedback[C]. New York: In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, ACM, 2021: 1-14.
- [24] TAKANO K, LI K F. An Adaptive Learning Book System Based on User's Study Interest[C]. Piscataway: In Proceedings of 2011 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, IEEE, 2011(8): 842-847.
- [25] NATALIA K, FRAUENFELDER U H. On the Effects of 12 Perception and of Individual Differences in 11 Production on 2 Pronunciation[J]. Frontiers in Psychology, 2014(5): 1246.
- [26] PENNINGTON M C. Computer-aided Pronunciation Pedagogy: Promise, Limitations, Directions[J]. Computer Assisted Language Learning, 12(5): 1999: 427-440.

(上接第257页)

- [9] 王中, 高放, 王曦阳, 等. 基于灰色关联分析的互联网产品用户体验度量方法[J]. 包装工程, 2018, 39(24): 141-145.  
WANG Zhong, GAO Fang, WANG Xi-yang, et al. Measurement Method of Internet Product User Experience Based on Gray Correlation Analysis[J]. Packaging Engineering, 2018, 39(24): 141-145.
- [10] 潘伟杰, 范灵秀, 吕健. 面向多维意象的产品设计方案评估研究[J]. 工程设计学报, 2017, 24(3): 250-255.  
PAN Wei-jie, FAN Ling-xiu, LYU Jian. Research on Product Design Scheme Evaluation for Multi-Dimensional Imagery[J]. Chinese Journal of Engineering Design, 2017, 24(3): 250-255.
- [11] 杨印生, 谢鹏扬, 李洪伟. 基于DEA的加权灰色关联分析方法[J]. 吉林大学学报: 工学版, 2003, 33(1): 98-101.  
YANG Yin-sheng, XIE Peng-yang, LI Hong-wei. Weighted Gray Correlational Analysis Method Based on DEA[J]. Journal of Jilin University(Engineering and Technology Edition), 2003, 33(1): 98-101.
- [12] 景荣. 外观创新在装备制造业中的设计研究[D]. 广州: 广东工业大学, 2012.  
JING Rong. Research on the Contour Innovation and Design in Equipment Manufacturing Industry[J]. Guangdong University of Technology, 2012.