

在语言与图像之间 ——DALL-E 3 在中文描述能力训练中的应用

王静

普林斯顿大学, 美国

摘要

“描述”是外语能力标准要求高级水平外语学习者掌握的三种语言功能之一。然而,通过分析美国两本具有代表性的中高级教材中的描述性语言,我们发现中文教材在话题选择和语言功能的分布上呈现出重叙议而轻描述的倾向;教学活动对学生描述能力的训练也相对缺乏。因此,本研究在美国大学高级中文课中引入 AI 图像生成工具 DALL-E 3,根据学生语言水平和课程内容选择具有文化内涵的图片,要求学生用中文与 DALL-E 3 进行互动,并不断调整描述语言,力求使 DALL-E 3 生成接近原图的图片。通过对学生使用的描述性文字和反思文章的量化及质性分析,我们发现学生的描述能力普遍低于叙述和议论的能力;华裔学生及具备一定中文背景的学生在描述能力上表现出明显优势。本文提出的基于 DALL-E 3 训练学生描述能力的教学活动是将 AI 工具引入中文教学的一次有益探索。

关键词

中文教学, 描述, DALL-E 3, 语言能力标准

1 引言

2022 年 OpenAI 公司发布了 ChatGPT3.5 版本。这种生成式人工智能工具能够使用多种自然语言与用户交流,处理文本生成、翻译、问答、编写代码等多种复杂任务。随着 ChatGPT4、图像生成工具 DALL-E 3、视频生成工具 SORA 以及其他具有类似功能的人工智能应用的推出和成熟,人工智能工具在自然语言处理方面的表现突飞猛进,能够在文字、图像、音频与视频等多模态之间自由转换、应对各种场景,这也为其在外语教学领域的应用提供了更多的可能性。在国际中文教育领域,已有部分研究探讨人工智能技术的具体应用,涵盖了利用 ChatGPT 进行词汇教学(孟凯, 2024)、写作反馈(高安娜, 2025)、少儿读物编写(宋飞, 2024)、批判性思维培养(Jiang et al., 2023)等诸多方面,并将 ChatGPT 的使用与项目式学习(Liang & Luo, 2024)、任务型教学(Jiang et al., 2023)、差异化教学(Li et al., 2024)等二语教学理论相结合,探讨了 ChatGPT 在不同中文水平、教学环境及课型中的表现。然而,外语教师除了可以利用 ChatGPT 在文本输出方面的能力把它作为教学辅助(刘利, 2023),也可以借助其他 AI 工具来聚焦特定语言能力的训练。本研究通过设计教学活动,利用 DALL-E 3 将语言转化为图像的功能来训练中文学习者的中文描述能力。这种将科技融入教学的尝试,是否能够训练学生的

描述能力开辟有效的新途径,同时提升语言学习的趣味性,培养学习者的自主学习与探索能力?在与人工智能互动的过程中,学生的学习体验对其学习态度可能产生怎样的影响?这些问题构成本研究的主要关注点与探索方向。

2 文献综述

2.1 关于文本类型的研究

在进入关于“描述”的研究综述之前,我们需要先明确什么是“描述”(describe)。The Britannic Dictionary(网络版)对“describe”一词的定义是:

to tell someone the appearance, sound, smell, events, etc., of (something or someone): to say what something or someone is like. 向他人说明某人或某物的外表、声音、气味、发生的事情等情况;说明某人或某物是什么样的。(作者翻译)

Pflugmacher (2005)指出描述“用于识别地方、物体或人物的属性。”在中文里,“描述”和“描写”这两个词都可以对应英文的“describe”。根据《现代汉语词典》(2016年第7版):“描写”指“用语言文字把事物的形象或客观的事实表现出来”,而“描述”是“形象地叙述;描写叙述”。可见,“描述”的对象既可以是静态的存在,也可以是动态的发展过程。

描述、叙述(narrative)和解释(explanation)是认知科学、语言学以及文学研究中的重要概念。他们“不仅可以被视为认知活动,而且也是交流形式,是根植于社会文化、制度和特定学科实践历史中的文本类型”(Herman, 2008)。早在20世纪70年代,语言学家就提出了“文本类型”(text type)的概念,作为“一种抽象与分类的原则,一种旨在捕捉文本或话语中的结构、功能以及其他常规化模式的分析性类别。”(Georgakopoulou, 2005)。

作为文本类型之一的“描述”被广泛应用于语言学研究的许多方面,例如在语言交流障碍的研究上,Vandenborre等学者(2017)比较分析了长期失语症患者与普通人对指定图片进行书面和口头描述的差异。在语用研究领域,Tseng(2015)通过分析亚洲生产和设计的创意产品的英文说明来探究在跨文化商务交流中获得正面反馈的要素。

文本类型的概念也被引入二语习得的研究和评估框架中。Nagao(2022)采用“基于文体的方法”(Genre-based approach)探讨如何有效进行学术写作中描述部分的教学,其研究对象为日本大学的初级和高级英文学习者。Wu(2024)从量化和质性两个角度分析了高级水平中文学习者的口语表达能力。在质性观察的部分,作者分析了不同水平的学习者在OPI口语测试中回答描述型问题时的表现,发现高级和优级学生在描述特定空间时,与中级水平学生存在显著差异:他们的描述内容更倾向于超越个人生活环境,描述口吻从第一人称向第三人称代词过渡。随着语言水平的提高,学习者在描述中能够提供更具体的内容,运用更复杂的词汇和话语结构,同时提高语言准确度,并减少对测试者后续问题的依赖。

上述两项研究表明无论在书面写作还是口语表达方面,外语学习者的描述能力都需要进行专门训练。然而,在语言学和二语习得领域对“描述”这一语言功能的关注仍较为有限。在外语教学中,关于如何提高学习者描述能力的实证研究更是匮乏。虽然Wu(2024)的文章中有专门的部分讨论口语测试研究对教学的启示,但作者提供的建议比较宽泛,诸如提高语速、加强产出性的词汇训练、以及从单句扩展到逻辑连贯的语段输出等。本研究试图在一定程度上弥补相关实证研究上的空白。

2.2 外语教学能力标准对“描述”的要求与规定

在美国外语教学委员会 (ACTFL) 制定的外语能力标准 (Proficiency Guidelines 2024) 中, 叙述 (narration)、描述以及议论或假设 (argument and hypothesis) 被明确列为三种基本的语言功能。中国教育部于 2021 年颁布并实施的《国际中文教育中文水平等级标准》(以下简称《标准》) 也对学习者使用中文进行描述的能力做出了具体规定。

ACTFL 标准将学习者的语言能力分为五个主要层级, 并对听、说、读、写四个技能进行了详细说明。以口语表达为例, 对“描述”功能的要求首次出现在中级水平的说明中。在展示模式 (presentational mode) 下:

Can express own thoughts and present information and personal preferences on familiar topics by creating with language, primarily in present time. May show emerging evidence of the ability to tell or retell a story and provide additional description. (p.19). 【学习者】应该能够灵活运用语言表达想法, 并在熟悉的话题上用现在时态传递信息和个人偏好。同时, 【学习者】可能会显示出讲述或复述故事以及提供额外描述性内容的初步能力。(作者翻译)

以上说明反映出学习者的口语表达能力从中级向高级阶段过渡的特征。在高级阶段, ACTFL 对学习者的“描述”能力的要求显著增加并进一步细化。学习者应该能够在所有时态下进行具体、直接的描述。高级阶段又细分为三个层级, 在从低级到高级的进阶中, 对描述能力的要求逐步提高: 从“叙述与描述相分离” (Their narrations and descriptions tend to be handled separately. p. 20) 到“叙述与描述相结合, 并交织在一起, 进行成段表达” (Narration and description tend to be combined and interwoven to relate relevant and supporting facts in connected, paragraph-length discourse. p. 21)。在从高级向优级过渡的阶段, ACTFL 标准指出学习者在进行议论和假设时, 可能会退回到简单的叙述和描述 (They may resort to simplification through the use of description or narration in place of argument or hypothesis. p. 21)。

由此可见, 在 ACTFL 的能力标准框架中, 议论和假设的难度和层级高于叙述和描述。在优级阶段, ACTFL 标准强调使用文化和历史典故的能力 (often using cultural and historical references to express meaning without extended description or explanation, p. 23)。如果学习者还需要通过详细的描述或解释来传情达意, 这往往被认为是口语表达能力不足的表现。简而言之, 叙述和描述是外语学习者在高级阶段发展的两种核心语言功能。学习者在达到高级水平时, 应能将二者相结合, 参与关于个人和公共兴趣话题的交流, 并通过成段表达处理复杂情况。

中国教育部制定的《标准》将“学习者运用中文完成某项语言交际任务时所表现出的语言水平”从低到高分三等九级, 不仅对音节、汉字、词汇和语法四个语言要素提出了量化要求, 还从听、说、读、写、译五个方面对不同能力等级做出了具体规定, 旨在培养和评估学习者“在不同情境下, 就不同话题, 用中文进行交际的能力。”(p.1) 就描述功能而言, 《标准》仅在“说”的方面对描述作出要求, 在其他四个方面没有相应规定。在从 4 级到 6 级的发展过程中, 学习者应该能够“描述较为复杂的情景”(p.4)、“较为详细地描述事件”(p. 5), “较为详细地描述事件和场景”(p. 6)。

中美两个标准都在语言学习的中高级阶段引入了对描述功能的要求, 期望学习者在高级阶段习得描述功能, 但同时也把这一功能限定在高级阶段。ACTFL 标准甚至把避免使用叙述和描述作为判断学习者语言能力是否达到优级的一项指标。

3 美国中文教材中的描述性文字

既然中美两国的语言能力标准都强调描述功能在高级阶段的重要性,那么目前美国高校所采用的中文教材是否反映了这一要求?在美国已出版的中高级中文教材中,《中文听说读写》第4册(Integrated Chinese, level 2 part 2)和《我看中国:现代汉语中高级读本》(Eyes on China: An Intermediate-Advanced Reader of Modern Chinese)是两本具有代表性的教材,不仅市场占有率高,而且所选话题都超越了个人生活而集中在社会议题上。

《中文听说读写》第四册共包含10课,课文长度在1000字左右。虽然每篇课文里描述性文字的篇幅因话题不同存在较大差异,但整体所占比重很低。描述的对象主要可以分为以下三类:

1. 空间:包括房间、城市、旅游景点和自然风景等。如第11课《中国的节日》中对小区环境以及房间布置的描述。
2. 人物:包括个人的性格、习惯,以及一个群体的特点。如第15课《男女平等》中对舅舅的描写。
3. 事件或场景:如第14课《生活与健康》中对晨练的描述。

《我看中国》共30课,课文篇幅介于800至1500字之间,话题丰富,语言的复杂度高。《我看中国》的课文以意见的表达为主,强调对学习者议论和假设能力的训练,因此描述性语言在每篇课文和整本教材中的占比极低。但另一方面,课文中的描述性文字所涉及的种类增加,除了与《中文听说读写》重合的空间、人物和事件,还扩展到动物、自然现象等。例如,第1课《雾霾》从雾霾的颜色、质感以及对人类的影响来描述这一空气污染现象;第2课《龙与熊猫》从样子、性情以及给人的印象等方面来描写这两种动物。

由于《我看中国》的课文篇幅较长,并且特别重视深入讨论有争议性的社会话题,因此同一话题下往往能看到叙述、描述和议论的有机结合,且句段之间逻辑关系清晰,这种篇章结构为学习者提供了向优级阶段提升写作和口语表达能力时可以模仿的范例。这种混合模式也呼应了关于描述、叙述和解释这三种文本类型之间模糊边界的理论,如 Mosher (1991) 所提出的“descriptized narration”描述化叙事和“narratized description”叙事化描述。以第20课《正能量》为例,以下引用的这段话通过武汉水灾的例子,讨论了老百姓对正能量报道的质疑:

今年夏天,武汉发生水灾,很多地方都被淹了。武汉政府出动军队救灾。央视的微博这样描述士兵救灾的情形:“【在大雨里,他们不穿雨衣;为了运沙袋,他们浑身是泥,里外湿透;60斤的沙袋,每个人扛300个;雨水加馒头,就是他们的食物。】”很多人看到微博的反应是:“政府每年投入那么多军费,为什么救灾设备还这么差?让救灾的士兵吃雨水馒头,怎么可以这样对待我们的英雄。”最后在网友激烈的批评下,央视删掉了这条“正能量”微博。

从功能的角度来分析这个语段,开头和结尾下划线部分组成了叙述性的大框架。在这个框架下,既有方括号内对人物外表以及对动作和事件的描述性语言,也有以灰色标注的议论。在描述和议论之前,分别有两个过渡性语句(引文中加粗的部分)进行提示。整个篇章围绕正能量报道引发负面批评这个话题展开,把叙述、描述和议论有机结合在一起。

通过分析这两本中高级教材中描述性语言的数量和特点,我们发现不同功能的语言在课文中占比不平衡是个普遍存在的现象;描述性语言在两本教材中所占比例很低,描述相对于叙述

和议论来说没有受到足够的重视。基于这些教材的教学可能也相应地在不同程度上忽视了对学习者描述能力的训练。

4 使用 DALL-E 3 训练学生的描述能力

在“描述”能力的训练方面，除了传统的以口头或书面形式描述具体对象的课堂活动和课后练习以外，ChatGPT4 附带的 DALL-E 3 也可以用于写作练习和自主学习的设计。DALL-E 是 Open AI 公司于 2021 年推出的基于深度学习技术的人工智能工具，2023 年发布的第三代能够根据自然语言指令生成“极其关注细节、甚至超越真实的图像，模糊了人类想象力与机器精确度之间的界限。”（Picano, 2024）

Picano 在其 2024 年出版的专著 *Generating Creative Images With DALL-E 3* 中对这项技术进行了全面介绍，从解释底层工作机制，到如何完善改指令以得到满意的图像，从技术背后潜在的伦理和法律问题，到分享不同领域使用 DALL-E 3 的具体案例，展示了科技与创造性交汇所带来的无限可能性。书中提到的具体案例，涵盖了广告、科学教育、电脑游戏、建筑设计等多个行业。其他学者的研究也涉及医学（Adams et al., 2023）、法律（Smith, 2023）、文学（Mudrik & Charles, 2023）等领域。在这些应用中，使用 DALL-E 3 的核心目标和最终产品都是它所生成的图像。

有学者指出 DALL-E 在理解非英语语言指令时存在明显缺陷，这限制了它的用户群体，并在不同语言用户之间造成技术不平等。Mudrik & Charles（2023）提出了一个基于 DALL-E 模型的自动化框架，一方面改善了 DALL-E 对非英语语言歌曲和故事的可视化表现，如希伯来语和俄语；另一方面突破了 DALL-E 目前只能提供一次性生成选项的局限，实现了能够展示故事情节变化的连续图像。尽管该研究主要关注儿童基础教育，但对外语教学也有借鉴意义。

目前，在外语教育领域，笔者尚未发现专门针对 DALL-E 3 的研究，更不用说反其道而用之，以其为工具进行语言能力训练。作为一种新尝试，本研究旨在为希望把人工智能工具应用于外语教学的学者和教师提供一些参考。

4.1 研究对象

此次教学实验共有 39 名学生参加，其中 19 名为非华裔高级中文课的学生，他们的中文水平跟据 ACTFL 能力标准处于中级高（Intermedia High）到高级低（Advanced Low）之间；另外 20 名为当代中国文学与电影课的学生，其中 4 名非华裔学生，中文水平处于高级低到高级中；11 名华裔学生，中文水平在高级中到高级高；还有 5 名有中文背景的学生，他们在华语地区生活过或接受过初中级教育，语言水平接近母语者。

4.2 教学活动设计

本研究中的教学活动要求学生用中文描述指定图片，并通过不断调整描述语言，尽可能使 DALL-E 3 生成的图片接近原图。由于 DALL-E 3 具备强大的理解与生成能力，用户即使只提供一个字，它也能生成无穷多幅图片。如果让学习者随意跟 DALL-E 3 进行对话，教师将难以有效控制输入的文字或生成的图片，这对提高学习者的语言能力几乎是无效的。因此，要求学习者根据原始图片向 DALL-E 3 发出指令，目的在于使学生的描述性语言尽量准确、具体和丰富，

对人机交互过程进行有效引导,使其更好地服务于语言教学;同时希望增强语言学习的互动性和趣味性。

鉴于参加教学活动的两组学生在语言水平和所学课程内容上存在差异,要求他们描述的图片也有所不同,以更好地匹配各自的学习需求和能力水平。以下图 1-3 是提供给学生进行描述的三幅图片。

图 1. 高级中文课学生需要描述的以人物为主题的原图



图 2. 高级中文课学生需要描述的以救灾为主题的原图



图 3. 中国文学与电影课学生需要描述的原图



高级中文课学生可以从图 1 和图 2 中任选一幅进行描述。图 1 是专为中高级水平学生设计拍摄的。无论是描述空间还是人物的外表和动作,其语言水平都能胜任。图 2 人物较多,背景复杂,描述难度显著高于图 1。这张照片源自网络,反映的是《我看中国》第 20 课《正能量》中提到

的武汉水灾事件，是当时央视微博报道所配发的新闻图片。学生在参与这个教学活动前，已学完这一课，熟悉事件背景，并掌握了描述所需的一些核心词汇。

图 3 是为中国文学与电影课的学生提供的图片，名为《善斗》，选自中国当代画家刘晓东创作的《二流子》油画系列作品。该画被普林斯顿大学艺术馆收藏。刘晓东曾应邀访问普大，并为这门课的学生做了一次讲座，详细介绍了该作品的主题和创作过程，因此，学生对画作中的元素、创作背景及文化含义已有一定了解。经画家本人授权，这幅油画被用作此次教学实验的原图之一。选择央视新闻图片以及中国当代画家的作品作为描述素材，旨在通过语言训练，让学生接触更多的中国社会与文化元素。

4.3 数据收集

39 名学生都在课下独立完成了与 DALL-E 3 的互动。虽然部分学生是第一次接触该 AI 工具，但都成功通过描述所选图片的细节，指挥 DALL-E 3 生成了一系列 AI 版本的新图片。学生需要提交他们跟 DALL-E 3 的完整交互记录，包括他们在不断调整描述的过程中输出的所有文字以及 DALL-E 3 生成的所有图片，并从中选出自己最满意的作品。活动结束后，每个学生都写了一篇反思文章，分享在使用 DALL-E 3 的过程中遇到的问题和挑战，评价该教学活动对提高他们中文水平的有效性，并讨论人工智能对外语学习的影响。

4.4 研究结果

4.4.1 量化分析

作者首先对两组学生提交的文本进行了字数统计，图 4 是高级中文课的统计结果，图 5 是中国文学与电影课。图中蓝色代表描述文本的字数，橘色代表反思文章的字数。两图中横轴上的数字分别代表两组参加教学实验的 19 及 20 个学生。两图中的数据根据描述部分的字数从低到高排列，提交多次互动记录的学生，择其字数最多的描述进行计算。中国文学与电影课的数据还按照学生的语言背景进行归类。

图 4. 高级中文课学生输出的文本字数

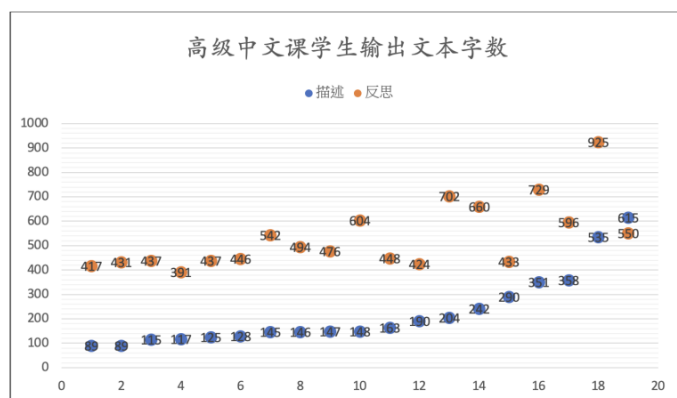
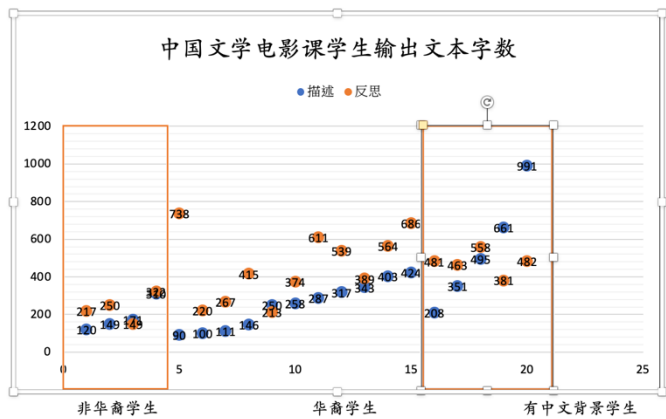


图 5. 中国文学与电影课学生输出的文本字数



比较两图中的数据，从字数的角度分析，无论学生的语言水平如何，反思文章的字数，即叙述和议论性语言的数量，普遍高于对图片的描述。学生对叙述和议论的主动输出意愿更强。然而，在高级中文课的非华裔学生群体中有一个例外，即 19 号学生，该学生曾在电影行业实习，因此在文章中提供了关于相机视角和物品颜色的详细描述。在电影与文学课中，非华裔学生、华裔学生及有中文背景学生输出的描述文字的平均字数分别为 188、248、541，可见有中文背景的学生在描述能力上表现出明显的优势。

从互动轮次来看，92% 的学生在反思文章里提到他们跟 DALL-E 3 进行了多轮互动，但可惜的是，并非所有学生都提交了完整的互动记录。在保存记录的文档中，互动的次数在 3 次到 9 次之间。

4.4.2 对描述文字的质性分析

以下选取了两组有代表性的描述片段及 DALL-E 3 生成的对应图片，从内容、语法、词汇、篇章结构及输出单位的角度对文本进行分析，以反映不同水平和背景的学生在语言质量上的差异。两组片段所选择的都是主题和内容相近的描述性文字，因而具有一定的可比较性。原文中的错误用星号标出，下划线部分则是超出本年级学习内容或学生语言水平的表达，这些表达可能是学生在写作过程中借助词典或人工智能工具完成的。

片段 A1（高级中文课 18 号学生）

他是中国人，30 岁左右。他的头发又短又黑，他戴着眼镜，眼睛的框子很薄，几乎看不到。男人好 * 想笑着，但是我们不看到他的嘴巴因为他戴着 * 的口罩，是黑色的 KN-95。男人穿着深蓝的毛衣，是 Ralph Lauren 的 * 名牌所以在毛衣的前面缀着黄色的打马球的人。毛衣也有咖啡色的护肘。右边的手里拿着咖啡色的手机，他在打电话。至于它的左边的胳膊和手，我们都看不到，因为它们在打开的苹果电脑的后面。男人一边打电话，一边看电脑。

图 6. 高级中文课 18 号学生提交的图片



片段 A2（高级中文课 1 号学生）

一个中国男人；他穿着白色的运动鞋，牛仔裤和海军的毛衣；他也戴着眼镜；他也 * 穿着口罩；他坐在椅子上；他一边打电话一边看着电脑；他的左腿交叉在右腿上；他的电脑 * 藏在腿上；他差不多三十五岁。

图 7. 高级中文课 1 号学生提交的图片



在高级中文课 19 名学生中，10 名选择了描述人物图片。上文引用的两个例子都聚焦于人物的外貌和姿势，学生主要依靠课堂上学过的词汇和语言结构完成对原图的描述。通过分析比较这两段文字，我们发现学生的描述能力存在明显差异。A1 片段使用了更多的词汇来描述细节，并通过使用如“几乎”、“好像”这样的副词，以及如“又 A 又 B”、“一边……一边”这样的语法结构让句式更复杂。此外，关联词“但是”、“至于”和“因为”的使用让句与句之间形成比较密切的逻辑关系。在未引用的部分，该学生还仔细描述了人物所处的环境，为 DALL-E 3 生成图片提供了更具体的信息。相比之下，A2 片段的描述完全由彼此毫无关系的单句堆砌而成。无论从词汇还是话语单位来看，都没有达到高级水平的要求。

中国文学与电影课的学生提交的描述和图片在不同样本之间的差异比高级中文课学生更明显。这主要是因为一方面学生需要描述的原图更复杂，另一方面学生的语言背景和水平差异更大。以下引用的三个片段分别来自一位在中国完成初中教育后移民加拿大的学生，一位华裔学生和一位非华裔学生。

片段 B1（中国文学与电影课第 20 号学生，在中国完成初中教育）

这幅油画的背景描绘了一个宁静的乡村景色。天空占据了画面上方三分之一的空间，呈现出一片淡蓝色，并点缀着散布的白云。画面的中部是一片丘陵地带，覆盖着层次分明的绿色植被，一些地方露出土黄色的土壤。在这些丘陵之上，可以看到远处的绿色山脉轮廓。

图 8. 中国文学与电影课第 20 号学生提交的图片



片段 B2（中国文学与电影课第 10 号学生，美国华裔）

画的中间是两个蹲着的男生，左边的穿着黑蓝色的外套，右边的穿着黑红色的外套。这群年轻人在农村，身边都是杂草和花。他们的后面能看到山还有蓝天白云。往山后看可以看到一个寺庙在画的右边，一个小塔在画的中间，还有零零散散的三个高楼在画的左边。

图 9. 中国文学与电影课第 10 号学生提交的图片



片段 B3 (中国文学与电影课第 2 号学生, 非华裔)

这油画的背景是布满岩石的绿色山坡。油画的设定是春天, 蓝天。有 * 萨诺摩天大楼在 * 破山的上面。这三个摩天大楼在画的左远端。这些摩天大楼非常小。而且, 这三个 * 摩天是公寓。一座庙宇在 * 破山的右远端。这也非常小。活动的中心是六个高中生在油画的前面。两个高中生是女生。四个高中生是男生, 两个穿古奇衣服。

图 10. 中国文学与电影课第 2 号学生提交的图片



以上三个片段都以背景描写为主。无论是文字描述, 还是 DALL-E 3 生成的图片与原图的差异, 都反映出这三个学生对画作的观察和语言水平上的差距。B1 片段的语言表达相当出色, 主题句引领全段, 篇章结构分明, 语言得体优美。B2 片段没有明显的语法错误, 但整段话以基本句式结构的单句组成, 例如使用系动词“是”的判断句和使用“在”的存现句, 句式重复且缺乏变化。就词汇来说, 大多数词语根据《标准》的分类属于 4 级或以下词汇, 个别词如“蹲”、“蓝天”以及“寺庙”等名词属于 6 级或更高级词汇。从内容来看, 除了列举空间中的事物外, 缺乏对具体细节的描述。虽然篇章遵循了从人物到布景, 由近及远的顺序, 但句子之间因为没有关联词而显得松散。B3 片段则存在较多的语法和词汇错误, 对空间和人物的描述还停留在简单的前后位置说明和人物服装的介绍。输出单位为单句, 而且句与句之间缺乏关联词或用于组织篇章的连接成分。短短一段文字中, 大多数句子都用来提供新信息, 缺乏话题链的建构和发展。一些偶尔出现的高级词汇, 如“布满”、“岩石”、“山坡”、“庙宇”等, 并非学生自己语言储备的一部分, 混杂在其他的基础词汇和句式中, 显得突兀且不协调。

4.5.1 关于反思文章的讨论

参与教学活动的学生在反思文章中分享了他们使用 DALL-E 3 的心得并讨论了人工智能对外语学习的影响。与 DALL-E 3 的互动让学生们意识到人工智能工具在理解和处理中文材料时的局限及其伦理偏见, 验证了 Mudrik & Charles (2023) 指出的 DALL-E 3 在根据非英语语言指令生成图片时的缺陷。以高级中文课学生对救灾新闻图片的描述为例, 学生们在反思文章中提到的问题包括: 首先, DALL-E 3 无法理解或有效执行中文指令, 如图片中的人被强行戴上口罩。学生反复强调的一些细节, DALL-E 3 无法准确执行, 比如出现在画面中的人物数量, 鞋子的颜色等。其次, DALL-E 3 对特定群体存在偏见, 比如坚持给图片中的中国士兵佩戴头盔和步枪。最后, 图片中出现不切实际的创造, 例如雨伞拿在第三只手上或悬浮在头顶。对这些现象, 学生们也提供了自己的解释。例如有学生指出 DALL-E 3 使用最近更新的数据库来生成图片, 所以图片上的军人都带着口罩, 也有学生用“人工智能的幻觉”来解释图片中的错误。无论这些解释的深度和可信度如何, 学生能积极主动地探索并思考, 用中文清楚地表达自己的意见, 提供自洽的解释, 这就已经达到了本次教学实验的目标。以下图 11 和 12 反映了学生们提到的这些问题。

图 11. 戴口罩的中国士兵及悬浮的雨伞



图 12. 第三只手及佩戴头盔和武器的士兵



所有参与此次教学实验的学生一致认为这是个既有趣又有意义的活动，其中 54% 的学生在反思文章中明确指出这个活动对提高他们的中文能力有帮助。为了复原原图，学生不得不“费劲脑筋想各个形容词”。“在大多数课文和上课的讨论中，学生使用的生词都跟政治和社会问题有很密切的关系。不过，做这个功课的时候，我不得不使用描述性的生词和语法，而且学习一些跟颜色、身体动作、服装有关的生词。”这个活动“不但让学生练习非常详细地描写一个东西或者一个环境——大到衣服的颜色，小到墙和地的质地，相机的视角，等等，而且它生动地说明 ChatGPT 和别的 AI 的坏处。”同时，92% 的学生指出因为人工智能存在明显的缺陷，只能在有限的范围内充当外语学习的工具，但无法取代老师的角色。当然也有学生对 AI 在外语学习中的作用持否定态度，一个学生表示，在他跟 DALL-E3 互动的时候，“我不太注意我的中文，主要注意 ChatGPT 的错误。再加上我练习说中文的时候想跟中文是母语的人谈话，所以无论作为学习中文的人还是作为 ChatGPT 的用户，我都不太满意。”

如何在语言教学及其他学科中有效使用人工智能工具，并对学生的行为进行规范，已经成了近两年高等教育界最关注的话题之一。在高级中文课学生中，有位法文专业的博士。她在反思文章中表示，一些老师“不但禁止 ChatGPT 的使用，而且跟学生们不断地讨论 AI 的危险性，它给社会会带来什么问题，等等。我固然同意，可是我觉得更有效的做法就是向中文老师学习：在一个活动里使用 ChatGPT，鼓励学生一方面评说它的好处，另一方面质疑它的有效性。这会让他们独立思考！”学生的正面反馈证明了该活动在提高学生学习兴趣，进行描述训练以及培养学生自学能力方面的有效性。

本次教学实验中学生对使用 DALL-E 3 的反馈在一定程度上印证了孙加玲等人（2025）的研究结论。该研究的一部分通过访谈问卷的方式，考察了上海高校留学生在用 ChatGPT 辅助中文学习过程中的“感知易用性”及“态度、感受与思考”。学生们均认可人工智能工具所提供的各种便利，同时也指出其相对于人类伙伴缺乏互动性等问题。但由于学生背景、所使用的 AI 工具及其输出形式、研究视角等方面的不同，两项研究中的学生反馈也呈现出一些差异。例如，在人工智能是否会取代人类教师，对人工智能伦理问题的关注等方面，学生的态度并不一致。两项研究可以互为补充，为今后在该领域的进一步探索提供参考。

5 讨论

让学生通过与 DALL-E 3 互动的方式练习用中文描述人物、场景与事件，虽然学生输出的描述文字中不可避免地存在不少文字重复，语法错误和表述不当之处，但学生的语言输出数量是传统的作业形式无法相提并论的。语言质量的提升则通过老师对学生所提交文本的反馈和修改建议来实现。同时，这是一个学生可以独立掌控并自主学习的过程。描述到底细致具体到什么程度，完全是由学生自行决定的，学生会主动查他找所需要却还未习得的词汇和表达法。

对两组不同水平学生所提交的描述文本和反思文章的分析表明,学生在描述、叙述和议论三种语言功能上的发展是不平衡的。描述能力并不会随着学生整体语言能力的提高有明显进步,尤其是对非华裔学生而言,需要专门训练。造成学生能力失衡有多方面的原因,除了观察能力的培养、词汇积累和写作训练等因素外,我们的教学材料和方法对描述的忽视也要负一定责任。

进行细致复杂的描述所需要的语言能力并不亚于议论和假设。对不同对象进行具体描述需要强大的词汇储备、细致的观察能力,甚至需要了解目标语文化在描述特定种类对象时约定俗成的表达方式,如中文中成语和典故的使用。

此外,任何语言功能都具有复杂度的差异,随着学习者整体语言能力的提高,应呈现出持续发展提升的过程。一方面由于中文不同于西方语言,例如没有明显的时态标记,在进行疑问和假设时无需改变句子的语序和动词形式,因此不少学习者在中高级阶段即可进行结构完整、意义清晰的议论和假设。另一方面,在高级阶段之后的学习过程中,因为中文词汇和语体的正式度有单双音节的区分及古汉语的影响,所以使用中文进行具体且正式的叙述和描述反而不是一件易事。

任何一种语言功能的训练都应是一个持续发展的过程,而非阶段性的目标。就描述能力而言,未来的研究可以尝试按照难易程度和常见度对描述对象进行分类,从最基本的空间,如房间、教室、图书馆,逐步扩展到人物外表、物品和常见动物的简单描写,从日常生活中的常见活动和场景,如节日的庆祝,发展到复杂的事件和情况,如空气污染、交通堵塞。对每一类别的描述又有由简入繁、从白描到工笔的进阶过程。虽然细致的描述往往涉及大量词汇,但不同语言中对某一类对象的基本描述往往遵循固定的范式。这些基本范式是可以通过教学传授并被学生习得的。这一构想还需要在未来的研究和实践中进一步细化,并验证其可行性和实际教学效果。

6 结论与展望

描述、叙述和解释是评估学生语言水平的重要指标,其中,“描述”无论在理论研究还是教学实践中受到的关注都比较少。尽管中美两国的外语能力标准都强调“描述”在中高级水平学生语言能力中的重要性,但美国市场上的中文教材在话题选择和语言功能分布上明显重叙、议而轻描述,缺乏典范的描述性语言。这种倾向使基于这些教材的教学实践对描述的训练往往少之又少。为了填补描述功能在研究与教学中的空白,笔者设计了一项通过人工智能工具 DALL-E 3 训练学生描述能力的教学活动,把图像跟语言表达有机结合,同时突破了 DALL-E 3 图像生成功能的传统应用场景。通过分析学生对指定图片的描述性文字和反思文章,研究发现学生在描述、叙述、议论这三种语言功能上的发展不平衡。华裔学生及具备一定中文背景的学生在描述能力上表现出显著的优势。这一发现为进一步探索描述功能的教学策略和教材编写提供了实践依据。

尽管学生指出了 DALL-E 3 的局限性,并认为人工智能工具无法取代人类教师,但我们应正视人工智能工具的巨大潜力,在批评 ChatGPT 及其他类似生成性人工智能工具负面影响的同时,更应充分发挥其优势,将其积极引入中文教学,创造性地利用 AI 为教学赋能。

参考文献

ACTFL Proficiency Guidelines (2024). https://www.actfl.org/uploads/files/general/Resources-Publications/ACTFL_Proficiency_Guidelines_2024.pdf.

- Adams, L.C., Busch, F., Truhn, D., Makowski, M.R., Aerts, H.J.W.L., & Bressemer, K.K. (2023). What Does DALL-E 2 Know About Radiology? *Journal of Medical Internet Research*, 25(1), e43110-e43110. <https://doi.org/10.2196/43110>
- Chou, C.P., Liu, J.C., & Zou, X. (2019). *Eyes on China: An Intermediate-Advanced Reader of Modern Chinese* 我看中国：现代汉语中高级读本. Princeton University Press.
- Gao, Anna (高安娜). (2025). ChatGPT 辅助汉语二语写作反馈研究 [A Study on ChatGPT-assisted Written Corrective Feedback in Teaching Chinese as a Second Language]. *国际汉语教学学报 [International Journal of Chinese Language Teaching]*. 6(2), 30-43.
- Georgakopoulou, A. (2005). Text-type Approach to Narrative. Herman, D., Jahn, M., and Ryan, M.L., (Eds.), *Routledge Encyclopedia of Narrative Theory*. Routledge. 594-596.
- Herman, D. (2008). Categories and the Cognitive Foundations of Discourse Competence. *Poetics Today*, 29(3), 437-472. <https://doi.org/10.1215/03335372-073>
- Jiang, X.L., Li, J., & Chen, C.H. (2023). Enhancing Critical Thinking Skills with ChatGPT-Powered Activities in Chinese Language Classrooms. *国际汉语教学学报 [International Journal of Chinese Language Teaching]*. 5(1), 47-73.
- Lexicographic Editorial Office, the Institute of Linguistics, Chinese Academy of Social Sciences (中国社会科学院语言研究所词典编辑室). (2016). 现代汉语词典第7版. [Modern Chinese Dictionary 7th edition]. 商务印书馆 [Shangwu yinshu guan].
- Li, Q.C., Zhang, J.F., & Cai, W. (2024). Utilizing ChatGPT to Implement Differentiated Instruction. *国际汉语教学学报 [International Journal of Chinese Language Teaching]*. 5(1), 74-89.
- Liang, X., & Luo, J. (2024). Integration of ChatGPT into Project-based Learning: A Course Design Framework. *国际汉语教学学报 [International Journal of Chinese Language Teaching]*. 5(1), 29-46.
- Liu, Li (刘利). (2023). ChatGPT 给国际中文教育带来的机遇与挑战 [Opportunities and Challenges Brought by ChatGPT to International Chinese Education] *世界汉语教学 [Chinese Teaching in the World]*. 37(3), 291-315.
- Liu, Y.H., Yao, T.C., Bi, N.P., Ge, L.Y., & Shi, Y.H. (2010). *Integrated Chinese: Level 2 Part 2. 3rd edition*. Cheng & Tsui.
- Meng, Kai (孟凯). (2024). ChatGPT 在国际中文词汇教学中的表现与应用策略 [The Performance and Application Strategies of ChatGPT in CSL Vocabulary Teaching]. *国际汉语教学学报 [International Journal of Chinese Language Teaching]*. 5(1), 3-14.
- Mosher, H.F. (1991). Toward a Poetics of “Descriptized” Narration. *Poetics Today*. 12(3), 425-445. <https://doi.org/10.2307/1772646>
- Mudrik, N., & Charles, A. S. (2023). Multi-Lingual DALL-E Storytime. *2023 IEEE Integrated STEM Education Conference (ISEC)*. IEEE. 326-332. <https://doi.org/10.1109/ISEC57711.2023.10402311>
- 中华人民共和国教育部国家语言文字工作委员会 (National Language Commission under the Ministry of Education of the People's Republic of China). (2021). 国际中文教育中文水平等级标准 [Chinese Proficiency Grading Standards for International Chinese Language Education].
- Nagao, A. (2022). A Genre-based Approach to Teaching Descriptive Report Writing to Japanese EFL University Students. *TESL-EJ: Teaching English as a Second or Foreign Language*. 26(3). doi: 10.55593/ej.26103a13.
- Pflugmacher, T. (2005). Description. *Routledge Encyclopedia of Narrative Theory*. Routledge.
- Picano, H. (2024). *Generating Creative Images With DALL-E 3*. Packt Publishing.
- Smith, B.W. (2023). DALL-E Does Palsgraf. *Journal of Law, Technology and the Internet*. 14(1), 87+ <https://link-gale-com.ezproxy.princeton.edu/apps/doc/A742003570/LT?u=prin77918&sid=summon&xid=5dea5656>

- Song, Fei (宋飞), Tan, Yulu (谭钰璐), & Pu, Yu (浦钰). (2024). 生成式人工智能辅助编写国际中文教育少儿读物的应用与实践——以 ChatGPT 与 Midjourney 为例 [Application and Practice of Generative Artificial Intelligence Assisted Writing of Children's Reading Materials in International Chinese Language Education Field]. *国际汉语教学学报 [International Journal of Chinese Language Teaching]*. 5(1), 15-28.
- Sun, Jialing (孙加玲), Wang, Yanyan (王艳艳), & Qian, Ziang (钱子昂). (2025). 中文学习者 ChatGPT 接受度的混合方法研究——技术接受视角 [A Mixed-Methods Study on ChatGPT Acceptance Among Chinese Learners: Technology Acceptance Perspective]. *国际汉语教学学报 [International Journal of Chinese Language Teaching]*. 6(2), 100-115. The Britannica Dictionary. <https://www.britannica.com/dictionary>
- Tseng, M.Y. (2015). Describing Creative Products in an Intercultural Context: Toward a Pragmatic and Empirical Account. *Journal of Pragmatics*, 80, 52-69. <https://doi.org/10.1016/j.pragma.2015.02.004>
- Vandenborre, D., Visch-Brink, E., Dun, K.V., Verhoeven, J., & Mariën, P. (2017). Oral and Written Picture Description in Individuals with Aphasia. *International Journal of Language & Communication Disorder*, 53(2), 294-307. <https://doi-org.ezproxy.princeton.edu/10.1111/1460-6984.12348>
- Wu, Y. (2024). Exploring Oral Proficiency Profiles of Advanced L2 Chinese Speakers: A Mixed-Methods Study. Kecskes, I. & Zhang, H., (Eds.). *Chinese as a Second Language Research from Different Angles*. Brill. 265-292.

投稿: 2024 年 9 月 8 日; 接受: 2025 年 2 月 17 日; 出版: 2025 年 6 月 11 日

作者简介

王静, 美国威斯康辛大学麦迪逊分校博士, 现为普林斯顿大学东亚系高级讲师, 中文项目主任。主要研究领域为国际中文教育及汉语的现代化。

Between Language and Images: Utilizing DALL-E 3 for Enhancing Descriptive Training in Teaching Chinese as a Second Language

Jing Wang

Princeton University, U.S.A.

Abstract

“Description” is one of the fundamental language functions for advanced level second language learners, as required by foreign language standards. However, an analysis of descriptive language in two representative intermediate-advanced Chinese textbooks in the U.S. reveals a tendency in these materials to emphasize narration and argument while ignoring description. As a consequence, instructional practices based on these materials provide relatively little training in the descriptive function. To address this issue, this study integrates the AI image generation tool DALL-E 3 into university-level advanced Chinese courses. Pictures with cultural significance were selected according to students’ language proficiency and course content, and students were tasked with interacting with DALL-E 3 in Chinese. They refined their descriptive language to enable DALL-E 3 to generate images as close as possible to the original pictures. Quantitative and qualitative analyses of students’ descriptive texts and reflective essays revealed that students’ descriptive abilities were generally weaker than their narrative and argumentative skills. Using DALL-E 3 to train students’ descriptive skills represents a valuable attempt to incorporate AI tools into Chinese language teaching.

Keywords

Teaching Chinese as a second language, description, DALL-E 3, language proficiency standards

Jing Wang (Ph.D), Senior Lecturer in the Department of East Asian Studies at Princeton University and serves as the Director of the Chinese Language Program. She received her Ph.D from the University of Wisconsin-Madison in Chinese Literature. Her research interests include Teaching Chinese as a Second Language and the modernization of the Chinese language.