

ChatGPT 辅助汉语二语写作反馈研究

高安娜

华东师范大学, 中国

摘要

随着以 ChatGPT 为代表的生成式大语言模型的不断发展, 人工智能辅助二语教学成为近年来的研究热点之一。人工智能辅助二语教学的相关研究涉及词汇习得、二语写作等多个方面, ChatGPT 辅助二语写作的相关研究成果也不断涌现, 但 ChatGPT 辅助汉语二语写作的研究成果仍相对较少。本文以某在华北美短期强化沉浸式中文项目为例, 收集了 62 篇中高级汉语二语学习者的作文文本, 对比并总结中文主讲教师和人工智能书面纠正性反馈类型的异同, 并从内容和形式方面对比中文主讲教师的书面纠正性反馈和 ChatGPT 的书面纠正性反馈特点, 并探究 ChatGPT 的书面纠正性反馈在词法、句法层面的正确性, 以期 ChatGPT 在汉语二语写作反馈方面的应用提供一定借鉴。

关键词

ChatGPT, 书面纠正性反馈, 国际中文教育

1 引言

长期以来, 二语写作反馈在培养二语习得者写作技能的过程中起到了重要作用, 受到了教师和学者们的关注 (王颖, 2007)。二语写作反馈按照主体可分为教师反馈、同伴反馈、计算机反馈三部分。随着计算机技术的发展, 自动生成写作反馈 (automated written corrective feedback, 简称 AWCF) 的相关工具不断被应用于二语教学 (Koltovskaia, Svetlana, 2020), 革新了二语教师书面反馈的传统方式。而 ChatGPT 作为生成式的大语言模型, 能依据输入的指令, 依靠算法和自然语言处理技术输出类似人类的语言 (Gatt, Krahmer, 2017), 与传统的自动生成写作评估 (automated writing evaluation) 工具相比, ChatGPT 能更有效地修改专业英语作文中的低级错误, 合适的指令也能有效提升 ChatGPT 的反馈质量 (Yushan Zhou, 2023)。ChatGPT 不仅能提供及时的、个性化的反馈, 还能检测并修正二语习得者写作中的语法错误, 还能二语写作的内容、结构、逻辑论证提出建议 (Kasneci, E. et al., 2023), 具备辅助教师进行二语写作书面反馈的功能 (Deng, J. & Y. Liu, 2022)。

与此同时, 现有研究表明, 二语教师花费了大量时间和精力为二语习得者提供书面纠正性反馈 (Ferris, 2013), 但二语教师书面纠正性反馈很难保证正确性和有效性, 教师需要接受更多关于二语书面纠正性反馈的专业训练和培训以提升书面纠正性反馈的有效性 (Lee, 2004)。有效的书面纠正性反馈也要求教师具备扎实的语言功底和专业技能, 并要求教师在课后花费相应的

时间和精力,在一定程度上,这可能加重二语教师的工作量。随着现代教育技术的不断发展,人工智能等现代教育技术能够跨越时空的局限,为学习者提供及时的反馈,能在一定程度上弥补二语教师书面纠正性反馈的不足,减轻教师反馈的负担(Jim Ranalli, 2018),这也启示教师应有效利用现代教育技术以提升教学质量和工作效率。

在国际中文教育领域,计算机辅助二语写作反馈的相关研究较为有限。人工智能能否提升教师的二语写作反馈效率、能否有助于二语习得者的自我反思、纠错还有待探讨。随着人工智能技术的发展,国际中文教育的大变革已经箭在弦上,如何利用人工智能技术赋能国际中文教育高水平发展也有待进一步研究(刘利、史中琦、崔希亮等, 2023)。由此,本文将以 ChatGPT 为例,探究二语教师书面纠正性反馈和 ChatGPT 书面纠正性反馈的异同,为二语教师和二语习得者合理利用 ChatGPT 辅助汉语为二语的写作提供一定经验和策略支持,以期国际中文教育的智能化发展提供一定启示。

2 研究现状

二语教师的写作纠错反馈可以分为书面和口头两种形式,教师书面纠正性反馈(written corrective feedback, 简称 WCF)是由教师在学生作文上写出修改建议或评语,口头纠错反馈通常借助师生之间的写作会议(oral conferences)进行,鉴于 ChatGPT 目前以书面反馈形式为主,本文仅对比 ChatGPT 和教师的书面写作纠错性反馈,不涉及非书面形式的纠错。

2.1 教师书面纠正性反馈

学界关于教师书面纠正性反馈的研究可分为四部分(Hyland、Hyland F, 2006),包括教师书面纠正性反馈的类型及策略(Ferris, D.、Roberts, 2001; Lee, 2004; Ellis, 2009)、教师对书面纠正性反馈的态度(Jamoom, 2016)、学生对教师书面纠正性反馈的看法(Chen et al., 2016)及教师书面纠正性反馈的有效性(Bitchener et al., 2004)。在教师书面纠正性反馈方面,现有研究多以英语作为第二语言为对象开展实证研究。在汉语作为第二语言的教学视阈下,现有教师书面纠正性反馈研究成果较少。

目前,学界对书面纠正性反馈的具体分类还未形成共识。Ellis (2009)将书面纠正性反馈策略分为直接反馈、间接反馈、元语言反馈、聚焦与非聚焦反馈、电子资源反馈、重述。Wilson 和 Czik (2016)从语言水平和形式两个维度进行分类,首先将书面纠正性反馈归纳为对低阶写作技能和高阶写作技能两部分的修正。低阶写作技能包含七个要素,即拼写、大小写、标点符号、句子结构、语法、格式和用词;高阶写作技能包括四个要素,即观点和阐述、逻辑结构、行文风格和自我,这四个要素都针对学习者的写作过程和经验,整体分类较为全面。此外,现有研究分析、总结了教师使用不同类型书面纠正性的频率及影响因素。教师会根据学生的偏误类型给出不同的反馈,针对不符合语法规则的偏误,教师更倾向于直接反馈(Ferris, 2006)。研究还指出,不同类型的反馈可能与教师是否为母语者有关,母语者更倾向于提供间接反馈(Han, 2017)。

相关研究还揭示了影响教师进行书面纠正性反馈的因素,教师的工作量、学校制度要求、教师自身的专业度和对书面纠正性反馈的看法都会影响教师反馈的数量和类型(Lee, 2008)。

学生对书面纠正性反馈的态度各不相同。学生对教师的直接反馈,例如勾画偏误词汇、直接标注正确表达等,具有较为正面的评价,因为这能让学生在较短时间内看到自身的不足,但学生也在自我纠错的过程中有更多收获(Chandler, 2003)。

最后, 现有研究也衡量了教师书面纠正反馈的有效性, 前人或聚焦于对全文的纠正, 或聚焦于对特定语法点的纠正。例如 Chandler 的研究表明教师的直接反馈对学生书面写作有积极的影响作用, 能够减少学生的偏误 (Chandler, 2003)。Ellis 等学者 (Ellis et al., 2008) 则聚焦于英语冠词通过对照组实验验证了教师书面纠正性反馈能促进学生对某一语言点的掌握。

汉语和英语的特点有显著差异, 二语习得者的偏误也有一定差异, 本文将在总结前人研究的基础上, 结合现代化教育科技辅助二语写作反馈的情况优化现有的二语写作的书面纠正性反馈分类, 建构编码表, 为进一步的对比分析奠定基础。

2.2 自动生成写作反馈

随着教育技术的不断发展, 自动生成写作反馈 (automated written corrective feedback, 简称 AWCF) 的相关研究成果也逐步增多。自动生成写作反馈有助于减轻教师的反馈负担 (Jim Ranalli, 2018), 同时提供及时且全面的反馈、帮助二语习得者反思自己的写作成果 (Stephanie et al., 2022; Jessie S. Barrot, 2023), 反馈效率较高。此外, 二语教师在开展书面纠正性反馈时可能会因个体情绪而有失公允, 自动生成写作反馈不受个体的情绪影响, 能提供客观的修正反馈 (Jim Ranalli, 2018)。由此, 现代教育技术辅助二语写作反馈具有一定现实意义, 但在国际中文教育领域方面的可行性和有效性还有待探究。

在 ChatGPT 辅助二语教学方面, 现有国外和国内的研究成果主要分布在英语二语教学两方面, 例如 ChatGPT 辅助英语二语词汇学习 (Javaid M, 2023) 和英语二语写作 (Su Y, 2023; 毛延生等, 2024)。现有的实证研究成果较少, 多为对 ChatGPT 辅助二语写作的建议和构想。对于 ChatGPT 是否能有效辅助二语教学这一课题, 更为深入的定量、定性研究还有待进一步开展。在汉语作为第二语言的教学方面, 现有研究成果相对更少, 仅有一项研究表明聊天机器人 (chatbot) 有助于汉语二语习得者的写作 (Chen H, 2020)。ChatGPT 作为生成式大语言模型, 其应用前景和方式也仍待深入探究。由此, 本研究将聚焦于国际中文教学二语写作反馈, 探究 ChatGPT 辅助二语写作反馈的可行性, 以期在一定程度上填补现有的研究空白。

现有研究也指出自动生成写作反馈工具能在一定程度上提升学生的学习参与度, 但学生也可能盲目遵循机器或者工具的修改和建议 (Koltovskaia, 2020), 这也从侧面警示教师应谨慎、合理地运用相关工具辅助教学。由此, 本研究仅对比教师书面反馈和 ChatGPT 书面反馈, 在保证教学质量、不影响正常教学开展的前提下, 遵循质性研究范式开展个案研究, 收集整理相关语料进行分析, 不涉及引导学生利用 ChatGPT 或其他自动生成写作反馈软件或网站。

3 研究目标

本研究以某在北美沉浸式中文项目中的教师二语写作反馈为语料开展定性研究, 学生均为从北美来华的留学生。在国际中文教育视域中, 目前暂无 ChatGPT 辅助写作的相关实证性研究成果, ChatGPT 辅助中文为二语的写作成效也有待通过实践进一步检验。本文将通过个案研究描写 ChatGPT 和主讲教师书面反馈的特点, 并从语言学视角分析 ChatGPT 的语言生成和反馈能力, 探究 ChatGPT 辅助二语写作反馈的有效性。在此基础上, 本文将归纳并总结具体的 ChatGPT 辅助二语写作反馈策略, 以期为教师利用人工智能辅助以汉语为第二语言的教学提供实践经验为二语习得者能否运用 ChatGPT 辅助中文写作提供一定参考。

本文的研究问题可归纳为两方面。其一, 教师书面纠正性反馈和 ChatGPT 书面纠正性反馈类型和数量有什么异同? 其二, ChatGPT 书面纠正性反馈的用词、语法是否准确、逻辑是否清晰?

由此，本文亦期望对未来的二语写作反馈提出展望，内容包括教师如何合理利用 ChatGPT 辅助二语书面纠正性反馈、提升写作反馈质量、二语习得者能否运用 ChatGPT 辅助二语写作的自我修改。

4 研究方法

本研究以笔者执教的某在北美短期强化沉浸式中文项目中的高级班为例收集语料，该班级的学生均有两年以上的中文学习经历。该班级由笔者和另一位中文老师合作授课，笔者为中文班主任，负责讲授本班的大班课、周考、作业反馈和成绩记录。另一名中文主讲老师负责少部分大班课的讲授、一对一课的讲授和作业反馈。

本研究共收集 62 篇美国来华大学生的作文，每篇作文均保留了主讲教师的修改和批注，其中大部分为电子版作文，少部分手写作文由研究者转化为电子版，总字数为 21534 字。作文中涉及的学生姓名、班级等可识别性隐私信息均用随机字母代替。首先，中文班主任修改并保存初次修改文档，另一名中文主讲老师审核，如有问题，双方协商修改后保存文档，在此基础上发给学生。本研究共收集 62 篇老师修改后的作文，保持原有修改痕迹。而后，研究者运用 ChatGPT4.0 版本（后文简称 ChatGPT）修改学生的作文，保存 ChatGPT 修改后的作文，标注出学生作文和 ChatGPT 生成作文不同的部分，储存为 ChatGPT 的修改文档。

鉴于本教学项目根据美国 ACTFL 外语能力指导方针（2012 版）制定了标准化书面和口语反馈要求，详细标准要求见附录 1。教师的所有反馈均根据项目制定的标准化要求开展，本文根据本项目的标准化课程要求和 ChatGPT for higher education and professional development: A guide to conversational AI (Atlas, 2024) 设计了向人工智能输入的具体指令，即“你是一名中文老师，根据以下要求修改这篇美国大学生的中文作文，学生水平为中高级，注意词汇不要超出学生的语言水平：首先，准确使用词汇和语法点、汉字书写无误，其次，进行成段叙述，逻辑性强，条理清晰，有一定深度”。

本研究的研究过程可分为五步。第一步，构建反馈类型编码表，本研究主要参考 Wilson 和 Czik (2016) 对书面纠正性反馈的分类进行初步编码，依据语言水平将书面纠正性反馈分为初阶和高阶两大类。鉴于教师反馈包括直接、间接、元语言反馈等形式，而受限于输入指令的要求，ChatGPT 提供的均为直接反馈，本文从反馈内容方面参考结合汉语的特点细化下位分类，总结归纳反馈编码表及具体阐释每一类的定义。第二步，本文对教师书面作文反馈进行编码，保留具体的教师反馈类型及典型示例，检验并修改编码表，形成教师书面反馈编码表，如表 1。第三步，本文对 ChatGPT 书面反馈进行编码，保留具体的 ChatGPT 反馈类型及典型示例，检验并修改编码表，最后确定 ChatGPT 书面反馈编码表，如表 2。第四步，本文分别统计教师和 ChatGPT 书面反馈类型的频次，得出教师书面反馈和 ChatGPT 书面反馈类型表及典型示例，并请两位有十五年教学经验的国际中文教师核查书面反馈编码以提升准确度，确定最终的书面反馈编码结果。第五步，根据具体的反馈案例深入对比、分析教师书面反馈和 ChatGPT 书面反馈的异同，从词汇、语法、篇章结构方面总结 ChatGPT 书面纠正性反馈的正确性，并在对比书面纠正性反馈之后为教师反馈提出实操性较强的建议。值得注意的是，本文的研究重点是对比纠正性反馈，同时也并未要求 ChatGPT 进行鼓励性反馈，因此本研究收集的教师表扬、鼓励式评语并未纳入编码统计的范围。在编码过程中，教师反馈和 ChatGPT 反馈有时会涉及多种类型，本文仅以核心反馈内容为依据进行标注。

表 1. 教师书面反馈编码表

| | 书面反馈类型 | 具体说明 |
|----|---------|--|
| 初阶 | 汉字反馈 | 圈画、修改错别字, 不改变句子结构和意思 |
| | 词汇反馈 | 针对词汇问题进行反馈, 例如近义词混用、口语词和书面词误用等 |
| | 词汇替换 | 行文不构成语法错误, 但进行同义词替换, 替换为更高级或更得体的词汇 |
| | 语法反馈 | 针对句子的语法结构、语序进行反馈, 修改有错误的部分 |
| | 标点、格式反馈 | 圈画、修改标点符号和作文格式错误 |
| 高阶 | 语义反馈 | 对文章句子、段落内容的直接或者间接修改, 旨在让表意更明确, 是对文章整体内容的评价和建议 |
| | 语篇衔接 | 对文章结构方面的衔接问题进行反馈, 例如关联词运用、段落前后衔接问题 |
| | 文体风格 | 对文体风格提出建议, 例如用更书面的表达代替口语表达 |
| | 思维逻辑 | 对文章思维逻辑的反馈, 旨在提升逻辑的连贯性和论证的合理性, 是对文章逻辑、思想的评价和建议 |
| | 身份认知 | 教师对学生学习和生活行为、思考的反馈, 例如以书面形式探讨社会问题、个人思考成果 |

表 2. ChatGPT 书面反馈编码表

| | 书面反馈类型 | 具体说明 |
|----|---------|------------------------------------|
| 初阶 | 汉字反馈 | 直接修改错别字, 不改变句子结构和意思 |
| | 词汇反馈 | 针对词汇问题的直接修改, 例如近义词混用、口语词和书面词误用等 |
| | 词汇替换 | 行文不构成语法错误, 但进行同义词替换, 替换为更高级或更得体的词汇 |
| | 语法反馈 | 针对句子的语法结构、语序直接修改, 修改有错误的部分 |
| | 结构替换 | 行文不构成语法错误, 但用其他的结构重述 |
| | 标点、格式反馈 | 对标点符号误用、作文格式错误进行直接修改或重述 |
| 高阶 | 语义反馈 | 对文章句子、段落内容的直接修改, 让表意更明确 |
| | 语篇衔接 | 对文章结构衔接问题的直接修改, 例如关联词运用、段落前后衔接问题 |
| | 文体风格 | 用更书面的表达直接替换口语表达 |
| | 思维逻辑 | 对文章思维逻辑的直接修改, 提升逻辑的连贯性和论证的合理性 |

5 案例分析

5.1 反馈类型统计

本文标注的教师反馈共计 339 条, ChatGPT 反馈共计 1137 条, ChatGPT 的反馈数量远大于教师反馈数量。编码后的反馈类型合并总结如下表。具体示例见附录。

表 3. 二语作文书面纠正性反馈类型统计表

| | 书面反馈类型 | 教师反馈 | | ChatGPT 反馈 | |
|----|---------|--------|-----|------------|-----|
| | | 数量 / 个 | 百分比 | 数量 / 个 | 百分比 |
| 初阶 | 汉字反馈 | 43 | 13% | 19 | 2% |
| | 词汇反馈 | 73 | 22% | 73 | 6% |
| | 词汇替换 | 13 | 4% | 325 | 29% |
| | 语法反馈 | 115 | 34% | 141 | 12% |
| | 结构替换 | 0 | 0% | 323 | 28% |
| | 标点、格式反馈 | 14 | 4% | 17 | 1% |
| 高阶 | 语义反馈 | 32 | 9% | 97 | 9% |
| | 语篇衔接 | 18 | 5% | 78 | 7% |
| | 文体风格 | 15 | 4% | 50 | 4% |
| | 思维逻辑 | 9 | 3% | 14 | 1% |
| | 身份认知 | 7 | 2% | 0 | 0 |

由上表数据可知，教师反馈和 ChatGPT 反馈均以初阶反馈为主，教师的初阶反馈注重修正语法和词汇偏误，语法反馈占比 34%，词汇反馈占比 22%。在学生表达不构成明显错误的情况下，教师没有进行语法结构的替换，结构替换的数量为 0，教师仅进行了少部分的词汇替换，占比仅 4%。而 ChatGPT 的初阶反馈以词汇和结构的同义替换为主，替换型反馈共占比 57%，相比之下，ChatGPT 针对词汇和语法的反馈占比明显少于同类的教师反馈占比，整体而言，ChatGPT 对词汇和语法的反馈针对性不强，以同义替换为主，这要求教师细化输入指令，教师可以要求人工智能进行元语言形式的反馈，指出具体的语法结构和例句并让人工智能说明理由。

在汉字和标点格式反馈方面，教师反馈的占比比 ChatGPT 反馈的占比大，且教师的汉字、标点格式反馈以直接的圈画批注为主要形式，ChatGPT 反馈以重述为主要形式。当 ChatGPT 重述时，学生作文的汉字、格式等都有一定程度的变动，在标注时，研究者仅选取最主要的修改类型标注。这也受限于 ChatGPT 的反馈形式，目前 ChatGPT 只能通过文本或图片形式进行反馈，很难运用“圈画批注”的反馈形式。

在高阶反馈中，教师身份认知的反馈占比 2%，ChatGPT 反馈并未体现对身份认知的反馈，本类型的标注数量为零。在高阶反馈中，教师和 ChatGPT 在语义反馈和文体风格反馈方面的占比相同，在语篇衔接和思维逻辑方面的反馈占比较为接近，但整体而言，ChatGPT 高阶反馈的数量更多。这体现出 ChatGPT 在文体风格、语篇结构修正方面的优势，也启示国际中文新手教师在反馈时也应并重词汇语法和篇章结构。

5.2 教师和 ChatGPT 反馈的特点

5.2.1 教师反馈的特点

首先，教师反馈形式较多样。一方面，教师的反馈用语包括汉语和英语，教师会根据内容的难易程度批注部分英语帮助学生理解教师的书面反馈。另一方面，教师的反馈主要包括直接反馈、间接反馈和元语言反馈，教师会标注具体的语法结构，提示学生复习学过的语法结构，教师也会采用简洁的批注符号，让学生能更明确地看出自己的错误，引导学生进行自我修正。

其次,教师反馈能在一定程度上激发学生的批判性思考并提供情绪价值。例如附录2中展示的示例,在学生提到 Duck Syndrome 时,主讲教师通过书面批注的形式针对这一社会现象发表了自己的看法,引导学生不需要假装自己很轻松,可以和家人朋友多沟通、缓解压力。值得一提的是,以本文收集的语料为例,主讲教师不仅提供纠正性反馈,还在纠正具体语言错误后多次附加正向的评论。因这部分评论为正向鼓励,因而并未计入上文的书面纠正性反馈中。除此之外,教师也提供了非常规的反馈,学生在作文中提到自己在中国留学遇到的问题和文化冲击,教师运用英文评论了自己的看法,鼓励学生不断提升语言水平以积极乐观的态度应对跨文化交往中的误解。学生也提到自己使用中国软件时遇到了许多困难,于是,教师制作了中国软件使用纸质手册,介绍四款中国常用软件的功能和界面,补充常用短语和界面操作流程,分发给所有学生。这在一定程度上启示国际中文教师,正向书面反馈能以更为多元、实用且有效的形式呈现,这也能有助于进一步构建和谐和谐的师生关系。

最后,教师反馈以学生为中心。首先,教师的书面纠正性反馈语言难度以学生现有的中文能力为基准,在反馈时,教师有意识地避免过于复杂的语言,优先采用学生可理解的书面表达。其次,教师的反馈除了书面纠正性反馈外,还包括正向鼓励性反馈,除了上文所述示例外,教师有整体性的正向书面评价,也有针对细节的正向反馈,这有助于激发学生的信心,在一定程度上鼓励学生。例如教师在作文的细节修改后添加总评论“我很喜欢你的观点和结构,但要注意复杂结构的表达。”

5.2.2 ChatGPT 反馈的特点

首先,ChatGPT 的文体风格反馈数量较多,修改后的作文较正式,有较强的参考价值。ChatGPT 能识别大部分学生作文中的口语化表达,并通过直接修改或重述的方式改为书面性表达。同时,ChatGPT 在进行词汇替换、结构替换和语篇衔接纠正的过程中也会将词汇替换为书面色彩更浓厚的词汇,修改后的文章比学生原文更为正式。相比之下,主讲老师仅针对最明显的口语化词汇或表达进行修正,在文体风格方面的反馈较有限。

其次,ChatGPT 语篇反馈数量较多且有效性较强。相比于教师反馈,ChatGPT 反馈对文章结构方面的修改更多。ChatGPT 在反馈中加入更多关联词或过渡句,让学生行文更流畅。本文收集的教师反馈均为新手教师的反馈,这也从侧面警示新手教师应该提升语篇反馈的能力,新手教师语篇反馈数量较少也可能与新手教师纠错能力及语篇反馈意识不强有关。词汇、语法固然重要,但篇章衔接的错误也不容忽视。特别是对于中高级二语习得者,篇章衔接是写作的难点和重点,这启示新手教师应该有意识地提升这一方面的纠错能力。

最后,ChatGPT 的重述性转写量过大,ChatGPT 对学生正确的表达进行了大量的同义替换。在学习任务繁重、学生语言水平不足的情况下,这样的修正是否有效也值得商榷。相比之下,教师精简化、针对性强的书面纠正性反馈或许更有利于引导学生进行有效反思和复习。大量的重述和转写也对使用人工智能的二语习得者的语言水平有更高的要求,二语习得者需要有较强的二语能力才能理解所有人工智能的反馈。

5.3 教师和 ChatGPT 反馈的差异

首先,教师反馈和 ChatGPT 反馈的数量差异较大,教师反馈数量远少于 ChatGPT 的反馈数量。在本研究中,教师有侧重地修改学生最主要的语法、词汇等错误,并没有纠正每一个细节,因此教师反馈数量远小于 ChatGPT 的反馈数量。而 ChatGPT 根据研究者输入的指令针对学生的

所有错误表达进行了修正，并进行了大量的词汇和语法结构替换，重述学生正确的表达，词汇和结构替换占比 57%。

其次，教师反馈和 ChatGPT 反馈的形式不同，在本文收集的教师反馈中，教师的反馈按照形式可分为直接反馈、间接反馈、元语言反馈和观点讨论，观点讨论即教师针对学生的某一观点进行延伸评论，该形式主要体现于身份认知反馈类型。而 ChatGPT 根据研究者输入的指令直接修改了学生的作文，输出修改后的文本，仅包括直接反馈和重述。教师反馈和 ChatGPT 反馈的形式也与学生的错误类型有关，在学生出现明显语法错误时，教师会结合元语言反馈形式直接指出语法结构的问题，提示学生正确的结构。但 ChatGPT 的反馈受输入指令的影响以重述和直接修改形式为主，直接替换词汇和结构，而非基于学生实用的语法结构修改。这体现出指令对人工智能反馈的重要影响，如果需要更多元语言反馈，教师或学生应该输入相关的详细指令，得到相应的内容，学生在使用过程中，受语言水平限制，也可以考虑加入更详细的指令，要求人工智能提供翻译或拼音。

再次，教师反馈和 ChatGPT 反馈的语言难度不同。教师反馈的语言以学生当前的语言水平为基准，主讲教师会避开过于复杂的中文表达，以可理解性反馈为主，遵循“i+1”原则，在遇到较难的表达时，教师会有侧重地使用英文批注，以降低难度。而 ChatGPT 反馈的语言难度超过学生现有的语言水平，且大量的同义替换在一定程度上会加重学生的认知负担，反馈效果也会因此受到影响。教师能够较快甄别人工智能反馈的有效性，但冗长且语言难度不匹配的反馈对非母语者的效用是较为有限的。

最后，教师反馈和 ChatGPT 反馈的个性化程度不同。主讲教师能基于学生现有语言水平、生活学习情况进行反馈，教师不仅巧妙运用正向反馈激发学生的学习动力，也在细节处下功夫，运用多样化的反馈展现教师对学生的关心，有助于维系良好和谐的师生关系。而 ChatGPT 反馈更为标准化，缺乏个性特点，但这一缺点也有可能通过输入更为详细的指令改善，这需要教师具备相应的教育科技应用能力，从而引导学生设计、输入详细且明确的指令，提高学生使用人工智能的能力。就反馈的效率而言，教师基于自身对学生的了解，能够更高效地提出个性化的修改建议，在特定情况下，教师的书面反馈也能展现出人文关怀，这一价值是人工智能无法代替的。

5.4 ChatGPT 反馈的局限性

在编码过程中，本研究共识别出 55 处 ChatGPT 的错误反馈，分为 8 小类，这 55 处错误反馈并未纳入表 2 的统计。一方面，ChatGPT 在现有的算法基础上会生成一部分有明显语法错误的中文句子，同时，ChatGPT 会误删关键信息，导致行文不畅，语义不明。另一方面，ChatGPT 很难准确识别学生所有的逻辑错误，在部分情况下，其反馈会遵循学生错误的逻辑表达，无法精准识别学生的逻辑错误。具体的错误反馈及示例总结如下。

5.4.1 句式杂糅

学生：Shake Shack 在上海同时面临挑战和机遇。

ChatGPT 反馈：Shake Shack 面临着既有挑战也有机遇。

学生的表述没有错误，但 ChatGPT 加入了连词“既……又……”，反馈的句子有明显的语法错误，句式杂糅，为无效反馈。

5.4.2 词汇误用

学生：通过看兵马俑，我看到了中国悠久的历史。

ChatGPT 反馈：透过观赏兵马俑，我感受到了中国悠久的历史。

不可否认的是，学生的表述有提升的空间，但并未出现需要修正的错误。而 ChatGPT 用“透过”替换“通过”，混用近义词，为错误反馈。

5.4.3 误删改

学生：和上海相比，西安有更悠久的历史。

ChatGPT 反馈：相比之下，西安拥有更为悠久的历史。

学生例句为段落的首句，本段意在对比西安和上海，“和上海相比”有助于说明本段的核心内容，但 ChatGPT 误删了“上海”这一关键信息，以“相比之下”替换“和上海相比”，修改后的内容失去了对比的关键词，让段落语义表达不完整，为错误反馈。

5.4.4 漏改

学生：受西方文化影响，更国际化。

教师：受西方文化影响，上海更国际化。

ChatGPT 反馈：受到西方文化影响，更加国际化。

学生在对比上海和西安时缺失了主语“上海”，表意不明，ChatGPT 并未修改这一关键错误，仅进行了同义词替换。

5.4.5 搭配不当

学生：我们调查了中国消费者对肯德基的态度，可以为 Shake Shack 的本土化策略提供参考。

ChatGPT 反馈：我们通过分析中国消费者对肯德基的态度，获得了宝贵的市场洞察，这对 Shake Shack 的本土化策略极为有用。

针对学生的句子，ChatGPT 添加了部分信息，让语义表达相对更完整，表明调查中国消费者的态度这一行为能对 Shake Shack 的本土化发展提供参考，但“获得市场洞察”搭配不当，为错误反馈。

5.4.6 成分残缺

学生：中国的教育比较重视学语言和你专业的课程。

ChatGPT 反馈：中国教育更注重语言学习和专业课程的深入。

上述学生例句不太正式，且语义表述不够明确，ChatGPT 进行了一定程度的修正，但其反馈出现了明显的语法错误，“深入”后的成分被遗漏，应改为“中国教育比较重视语言学习和专业课程的学习”。

5.4.7 逻辑错误

学生：家长急于让学生提高成绩。学生也逼家长上补习班。

ChatGPT 反馈：家长迫切希望孩子提高成绩，甚至导致学生逼迫家长参加补习班。

教师反馈：谁逼谁？

在上述案例中，教师并未直接给出正确的表达，只指出了句子的逻辑问题，为间接反馈，这有利于学生自主修改。但 ChatGPT 仅重述了学生的原句，并未修改逻辑问题，并且 ChatGPT 替换了大部分学生用的词语，整体的语言更为书面正式，但逻辑漏洞明显，并未达到理想的效果。

5.4.8 歪曲原意

学生：我觉得美国人对流浪汉的印象很消极。因为，有的流浪汉利用人们的好心，但是这看情况。

ChatGPT 反馈：我认为美国社会对流浪汉的印象很消极，因为一些流浪汉利用人们的善意，尽管情况并非如此。

学生的本意是“流浪汉是否利用人们的好心得看情况”，但 ChatGPT 的反馈则表示“流浪汉不利用人们的善意”，与原本意义不符，为错误反馈。由此可见，ChatGPT 能识别大部分学生表达中的问题并进行修正，但目前 ChatGPT 的中文表达仍会出现一定偏误，这也警示师生不能过度依赖 ChatGPT，需要鉴别 ChatGPT 反馈的有效性，这也对使用人工智能的二语习得者的语言能力有更高要求，初级水平的中文学习者很难理解所有人工智能的反馈。

6 总结与展望

在科技辅助教学层面，本研究对师生如何合理使用 ChatGPT 有一定启示价值。结合上文可知，ChatGPT 的书面纠正性反馈比教师的书面纠正性反馈更详尽、更正式，尽管 ChatGPT 有部分偏误，但国际中文教师，特别是新手教师，可以合理利用 ChatGPT 裨补阙漏、修正学生的作文。首先，教师在修改作文时会受限于教学涉及的词汇和语法，而 ChatGPT 能在一定程度上补充教师忽视的内容，为教师反馈提供有效参考。因此，教师可以借助 ChatGPT 修正学生的作文，有针对性地选择 ChatGPT 修改后的内容，提升二语写作反馈的准确度和专业度。其次，教师在使用 ChatGPT 时要仔细甄别 ChatGPT 用语的准确性和规范性，不能照搬，要结合学生的实际情况采用 ChatGPT 的反馈，避免采用 ChatGPT 的错误反馈。最后，国际中文教师需要提升自己使用人工智能的能力，可以有意识地调节指令让 ChatGPT 提供英文翻译、元语言反馈甚至是拼音，这能让 ChatGPT 的输出内容有更强的可理解性，能在一定程度上提升 ChatGPT 反馈的有效性。

结合上文可知，ChatGPT 的反馈相对教师反馈更为全面，ChatGPT 的反馈量更多，学者 Bitchener and Ferris (2012) 指出这可能会在认知上加重学习者的学习负担，本研究也通过个案分析对这一观点有一定佐证。与教师反馈相比，ChatGPT 的反馈内容更多，且词汇的难度和学生的汉语水平并不相符，这也反映出现有的人工智能难以准确评估中文等级水平，为未来的教育类人工智能提出发展方向。由此，二语习得者应慎用 ChatGPT。首先，ChatGPT 辅助二语习得者修改作文的有效性需要视学习者的汉语水平而定。对于高水平的汉语学习者而言，ChatGPT 的反馈或许会有效，但对于初中级的汉语学习者，ChatGPT 的反馈很可能是不可理解的无效反馈，甚至会加重学习者的自我修正负担。本文未涉及对不同水平的二语习得者使用 ChatGPT 有效性的考察，这也可以作为未来的一个研究方向。其次，二语习得者需要根据 ChatGPT 的优势针对性地使用 ChatGPT 辅助自学。以本文所得数据为例，ChatGPT 在书面用语、逻辑结构、语篇反馈方面的优势为二语习得者的词汇学习提供了技术支持，二语学习者可以利用 ChatGPT 拓展汉语词汇、了解词汇的常用搭配并要求 ChatGPT 提供情景和例句。相比于用 ChatGPT 辅助二语写作，或许学习者运用 ChatGPT 辅助词汇学习的效果会更好，这也为未来的研究提供了一定案例支持。

另一方面,在前人的研究中,纠正性反馈为书面反馈研究的重点,教师的正向反馈形式和内容的质量并未得到重视,但在实践过程中,教师可以进一步探索正向反馈的具体实践形式。国际中文教师应不仅限于“好棒”“真不错”类型的泛泛空谈,而应针对学生的具体行为、学习成果等进行正向反馈,这或许能为学生提供一定的情绪价值,正向书面反馈对师生互动情况、师生关系的影响也可以作为研究者们未来的一个研究方向。总而言之,在合理利用人工智能的同时,国际中文教师也应坚守人的主体性,发挥主观能动性,合理利用人工智能提升自我专业能力。

附录 1

ChatGPT 书面纠正性反馈指令的参考标准

| 项目参考标准 | ACTFL 参考标准 |
|---|--|
| 1. 内容: 作文有逻辑、条理清晰、有一定深度。 2. 准确度: 汉字正确, 能准确使用词汇和语言点。 3. 成段表达: 能较为准确地使用关联词, 进行成段叙述。 | Advanced-mid Writers at the Advanced Mid sublevel are able to meet a range of work and/or academic writing needs. They are able to narrate and describe with detail in all major time frames with good control of aspect, and to write straightforward summaries on topics of general interest. Their writing uses a variety of cohesive devices in texts up to several paragraphs in length, and exhibits good control of frequently used syntactic structures and a range of general vocabulary. Thoughts are usually expressed clearly and supported by some elaboration. Writers at the Advanced Mid sublevel may combine organizational features from different languages, and their writing may at times resemble oral discourse. |

附录 2

二语作文书面纠正性反馈示例表

| 书面反馈类型 | 教师修改示例 | ChatGPT 修改示例 |
|------------|---|--|
| 汉字反馈 | 学生: 小部份人选择自由职业。 教师: 勾画“份”, 写出“分”。 | 学生: 所以出现心理健康的问题。 ChatGPT: 并可能导致心理健康问题。 |
| 词汇反馈 | 学生: 所以我觉得了巨大的压力。 教师: 勾画“觉得”, 标注“感到”。 | 学生: 他们认为我是中国人。 ChatGPT: 他们误以为我是中国人。 |
| 词汇替换 | 学生: 他们应该找有兴趣的活动。 教师: 勾画“找”, 标注“(We learned a new word participate.) which is better than 找”。 | 学生: 我有大的压力。 ChatGPT: 我承受着巨大的压力。 |
| 初阶 语法反馈 | 学生: 你想回收什么就回收什么。 教师: 勾画本句, 标注“sb1 想 do sth 就 do sth; sb1 想 do sth sb2 就 do sth”。 | 学生: 我对这种现象持反对。 ChatGPT: 我对这种教育现象持反对态度。 |
| 结构替换 | 无 | 学生: 我重视获得好成绩。我自己有很高的期望。 ChatGPT: 我非常重视成绩, 给自己设定了很高的标准。 |
| 标点、格式反馈 | 学生: 例如有 Trump 的言论, 2016 总统大选对美国社会的影响, 和日益激烈驴象之争的内容。 教师: 勾画“,”, 标注“、”。 | 学生: 我们有三种回收方式, 快递, 上店, 或者上门。 ChatGPT: 我们提供三种回收方式: 快递、到店自送或上门回收。 |

| | | | |
|------|---|--|--|
| 语义反馈 | <p>学生：招聘者问了我儿童教学的经验，比如如何发展孩子的情感。</p> <p>教师：标注 “What do you want to say? 如何调节孩子的情绪？”</p> | <p>学生：他们，尤其小城市的人，很喜欢聊天。</p> <p>ChatGPT：特别是在小城市，居民通常很喜欢聊天，这将有助于你的语言实践。</p> | |
| 语篇衔接 | <p>学生：如果你只有本科学历，找工作很难。</p> <p>教师：标注 “如果……就……”</p> | <p>学生：肯德基在中国分店实施了本土化策略。</p> <p>ChatGPT：与此同时，肯德基在中国的分店也实施了本土化策略。</p> | |
| 文体风格 | <p>学生：还有在周末学校让学生上补课。</p> <p>教师：勾画 “还有”，标注 “better and more formal：除此之外”</p> | <p>学生：所以这个很难。</p> <p>ChatGPT：虽然这在中国颇有难度。</p> | |
| 高阶 | 思维逻辑 | <p>学生：虽然这样的本土化策略让喜欢吃洋快餐的人增加，还有重要的劣势。一方面是随着人们的收入增加，关注健康和营养的人越来越多了。为了改变垃圾食品的形象，这一家快餐品牌需要满足健康和营养的需求。因此它打破了以油炸为传统的传统，所以肯德基增加了沙拉和蔬菜汤等健康食品。</p> <p>教师：这段话的逻辑有问题。The disadvantage of localization strategy is not releasing healthy food. For your reference: 最后，随着人们的收入增加，关注健康和营养的人越来越多了。为了改变垃圾食品的形象，这一家快餐品牌需要满足健康和营养的需求。因此它打破了以油炸为传统的传统，所以肯德基增加了沙拉和蔬菜汤等健康食品。除此之外，本土化策略可能带来了一些负面的影响，过度本土化会让品牌失去美国特色。</p> | <p>学生：美国的许多学校培养学生的创造力，而且提供课外活动。</p> <p>ChatGPT：许多学校不仅提供课程学习，还鼓励学生参加各类课外活动，以培养其创造力。</p> |
| 身份认知 | <p>学生：人人应该看起来非常轻松，好像一切都很容易，这个现象叫 “Duck Syndrome”。</p> <p>教师：Duck Syndrome 这个现象越来越普遍，大部分人叫它“狂鸭症”。中国还有一句话“你只有非常努力，才能看起来毫不费力。”（You have to work really hard to look effortless.）</p> <p>在我看来，努力是必不可少 indispensable 的，但我们也不需要假装自己很轻松，如果有负面的情绪或者感到很累，我们需要找人沟通、发泄我们的负面情绪，之后，我们才能真的“毫不费力”。</p> | 无 | |

参考文献

- Atlas, S. (2024). ChatGPT for higher education and professional development: A guide to conversational AI[OL]. https://digitalcommons.uri.edu/cba_facpubs/548/Accessed 01/03/2024
- Bitchener, J., Young, S. & Cameron, D. (2004). The effective of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 14: 191-205.
- Bitchener, J. & D. R. Ferris (2012). *Written corrective feedback in second language acquisition and writing*. Routledge.
- Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of second language writing*, 12(3): 267-296.

- Canagarajah, A. S. (2012). Teacher development in a global profession: An autoethnography. *TESOL Quarterly*, 46: 258-279.
- Chen, S., Nassaji, H., & Liu, Q. (2016). EFL learners' perceptions and preferences of written corrective feedback: A case study of university students from mainland China. *Asian-Pacific Journal of Second and Foreign Language Education*, 1(5): 1-17.
- Chong S W. (2019). A systematic review of written corrective feedback research in ESL/EFL contexts. *Language Education & Assessment*, 2(2): 70-95.
- Chen H L, Vicki Widarso G, Sutrisno H. (2020). A chatbot for learning Chinese: Learning achievement and technology acceptance. *Journal of Educational Computing Research*, 58(6): 1161-1189.
- Deng, J. & Y. Liu. (2022). The benefits and challenges of ChatGPT: An overview[J]. *Frontiers in Computing and Intelligent Systems*, (2): 81-83.
- Ellis, Rod & Sheen, Young H & Murakami, Mihoko & Takashima, Hide. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, 36: 353-371.
- Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, 63: 97-107.
- Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes. *Journal of Second Language Writing*, 10(3): 161-184
- Ferris, D. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland, & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 81-104). Cambridge, UK: *Cambridge University Press*.
- Ferris, D.R., Liu, H., Sinha, A., & Senna, M. (2013). Written corrective feedback for individual L2 writers. *Journal of Second Language Writing*, (22): 307-329.
- Gatt, A. & E. Krahmer. (2017). Survey of the state of the art in natural language generation: Core tasks, applications and evaluation. *Journal of AI Research*, 61: 75-170
- Han, Y. (2017). Mediating and being mediated: Learner beliefs and learner engagement with written corrective feedback. *System*, 69: 133-142.
- Hyland, K. & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching*, 39(2): 83-101
- Jamoom, O. (2016). Teachers' beliefs and practices of feedback and preferences of students for feedback in university level EFL writing classrooms. [Unpublished Ph.D. dissertation]. University of Southampton.
- Jim Ranalli. (2018). Automated written corrective feedback: how well can students make use of it?. *Computer Assisted Language Learning*, 31(7): 653-674.
- Jessie S. Barrot. (2023). Using automated written corrective feedback in the writing classrooms: effects on L2 writing accuracy. *Computer Assisted Language Learning*, 36(4): 584-607.
- Javaid M, Haleem A, Singh R P, et al. (2023). Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, 3(2): 100115.
- Koltovskaia, Svetlana. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*. 44: 100450.
- Kasneji, E. et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103: 102274.
- Lee, I. (2004). Error correction in L2 secondary writing classrooms: The case of Hong Kong. *Journal of Second Language Writing*, (13): 285-312
- Lee, I. (2008). Understanding teachers' written feedback practices in Hong Kong secondary classrooms. *Journal of second language writing*, 17(2): 69-85.

- Su Y, Lin Y, Lai C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57: 100752.
- Stephanie Link, Mohaddeseh Mehrzad & Mohammad Rahimi. (2022). Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. *Computer Assisted Language Learning*, 35(4): 605-634.
- Wilson, J., & Czik, A. (2016). Automated essay evaluation software in English Language Arts classrooms: Effects on teacher feedback, student motivation, and writing quality. *Computers & Education*, 100: 94-109.
- Yushan Zhou. (2023). The Effectiveness of Automated Written Corrective Feedback on L2 Learners' Revision Outcomes: A Case for ChatGPT. *International Journal of New Developments in Education*, 5(25): 58-65.
- 刘利、史中琦、崔希亮等（2023）ChatGPT 给国际中文教育带来的机遇与挑战——北京语言大学与美国中文教师学会联合论坛专家观点汇编，《世界汉语教学》，3：291-315。
- 毛延生、王一航、邢艳茹（2024）ChatGPT 辅助高中英语写作反馈的实证研究，《教育测量与评价》，1：3-13。
- 牛瑞英、张蕊（2018）二语写作教师书面反馈焦点、策略及成效个案研究，《解放军外国语学院学报》，3：91-99+160。
- 王颖（2007）二语写作中的反馈研究述评，《山东大学学报（哲学社会科学版）》，5：48-52。
- 魏爽、李璐遥（2023）人工智能辅助二语写作反馈研究——以 ChatGPT 为例，《中国外语》，03：33-40。

投稿：2024年4月31日；接受：2024年9月20日；出版：2024年10月18日

作者简介

高安娜，华东师范大学在读硕士生，研究方向科技辅助教学、二语习得，先后于浙江师范大学《国际中文教育研究》和厦门大学《海外华文教育》发表独作论文，现为美国 Vassar college language fellow，国际中文教学时长超过 1000 小时。

A Study on ChatGPT-assisted Written Corrective Feedback in Teaching Chinese as a Second Language

Anna Gao

East China Normal University, China

Abstract

With the continuous development of generative large language models such as ChatGPT, AI-assisted second language learning has become one of the hot research topics in recent years. The research on AI-assisted second language education involves various aspects such as vocabulary acquisition and second language writing. There has been an emerging numbers of research on ChatGPT-assisted second language writing, but the research on ChatGPT-assisted writing in Chinese as a second language is still relatively scarce. This paper collects 62 essays from intermediate and advanced American college-level learners of Chinese as a second language from an intensive immersion Chinese learning program located in China. It compares and summarizes the similarities and differences between the written corrective feedback types provided by the Chinese instructors and AI, and contrasts the content and form of the Chinese instructors' written corrective feedback with that of ChatGPT. It analyzes the accuracy of the written corrective feedback of ChatGPT from the lexical and syntactic perspectives, in the hopes of providing insights for the application of ChatGPT in the feedback of Chinese as a second language writing.

Keywords

ChatGPT, written corrective feedback, Teaching Chinese as a Second Language

Anna Gao, master's student at East China Normal University, with the research focus on technology-assisted language teaching and second language acquisition. She has published two papers in International Chinese Language Education Research and Journal of Overseas Chinese Language Education. Currently, she is the Chinese language fellow at Vassar College, with over 1000 hours of teaching in TCSL.