

# 智能评改系统在高校英语写作教学中的探索

毕一飞

上海理工大学外语学院, 上海

收稿日期: 2023年3月7日; 录用日期: 2023年4月4日; 发布日期: 2023年4月11日

## 摘要

随着人工智能技术发展, 智能评改系统在英语写作教学中的应用越来越多。本研究旨在通过对三十位英语专业学生一学期的学习追踪, 探索当下学生对智能评改系统的接受度、对自我英语写作能力的认知状况以及比较国内外两种智能评改系统的差异。研究表明学生对智能评改系统的接受度低于预期; 对自己的英语写作能力的认知情况有一定提高, 这也是一定程度上肯定了本学期的英语写作教学方法; 两种智能评改系统在评改风格等方面差异性较大。在之后的英语评改研究中, 授课教师应考虑智能与人工评改相结合的方式, 探讨学生可接受、并相对能达到最大效力的评改模式。

## 关键词

英语写作, 智能评改系统, 批改网, Grammarly

# A Study of Intelligent Assessment System in Teaching English Writing for University Students

Yifei Bi

College of Foreign Languages, University of Shanghai for Science and Technology, Shanghai

Received: Mar. 7<sup>th</sup>, 2023; accepted: Apr. 4<sup>th</sup>, 2023; published: Apr. 11<sup>th</sup>, 2023

## Abstract

With the development of artificial intelligence technology, intelligent assessment systems are increasingly used in English writing teaching. This study aims to explore the current students' ac-

ceptance of the intelligent assessment system, the status of their perception of their English writing ability, and to compare the differences between two intelligent assessment systems at home and abroad by tracking thirty English-major students for one semester. The results show that students' acceptance of the intelligent assessment system is lower than expected. The cognitive status of their English writing ability has improved, which is a certain affirmation of the method used in English writing teaching. Two intelligent assessment systems have significant differences in their assessment styles. For future studies, researchers could consider the combination of intelligent and manual assessments and explore the assessment mode that is acceptable to students to achieve maximum effectiveness.

## Keywords

English Writing, Intelligent Assessment System, Pigai, Grammarly

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

英语专业学生的英语写作教学是高校教师自评难教的一门课之一，一直以来也是高校教师教学与研究的重点。国内学者从英语教学的各个角度进行了大量的研究，比如在英语写作教学与课程思政融合方面很多学者进行了总结并提出了切合实际的课堂应用方法[1] [2]；在具体的教学法方面，对比国外的任务型教学法(Task-based Language Teaching, TBLT)，国内学者文秋芳教授提出了更适合国内学生学情的产出导向法(Production-oriented Approach, POA) [3]。然而，作为一种语言输出的重要手段，几乎所有的英语写作教育研究都不可避免地要考虑到写作的有效批改与评价。批改和评价学生作文是英语教学过程中的重要阶段之一，也是高校教师一直面临的挑战，比如教师对学生的写作评价存在费时低效的现象。具体表现为，教师批改与评价任务重，批改作文耗费大量的时间与精力；教师批改的效果有局限性，一些写作问题教师不能全部指出或者学生忽略教师的评价、误用教师反馈等。

针对教师目前批改与评价现状以及随着人工智能技术的发展，特别是语料库与云计算不断完善，国内外出现了一些写作智能评改系统，旨在有效缓解教师费时低效的评改现状以及提高学生的学习效果。国外目前较为成熟的写作智能评改系统包括 Grammarly, ProWritingAid, Ginger, stylewriter, whitesmoke, NOUNPLUS, Hemingway 等。国内相比国外起步略晚，近年来较为成熟的智能评改系统包括批改网、iWrite、冰果写作等。其中，批改网在我国英语写作教学领域用户群体庞大，以批改网为载体探索大学英语写作教学效果的研究也越来越多。以“批改网”、“大学英语写作(教学)”为主题的研究在知网上截至目前有三百多篇相关文章发表。比如宋皓[4]对比了批改网智能评分和教师人工评分的信度，经过统计分析后得出结论为批改网的打分信度较高。而以国外的 Grammarly 或其系统为载体探索我国大学英语写作教学的研究却非常少，据笔者在知网搜索，目前还没有相关研究，并且国内外智能评改系统的差异也无从考证。不同评改系统对一份作文是否打分一致，英语专业学生对评改网的接受度如何还有待进一步研究。

针对以上背景，据我们目前了解，本研究是第一个以国内使用较多的批改网和国外使用较多的 Grammarly 两个智能评改系统为例，探索评改网的打分统一性以及追踪英语专业学生对评改网的接受度，为智能评改系统在高校英语写作教学中的应用提供参考。

## 2. 研究设计

### 2.1. 研究对象

本研究选取了笔者任教的英语专业大二学生共 30 人进行了为期一个学期的学习追踪。

### 2.2. 研究问题

具体研究问题包括: 1) 英语专业学生对智能评改系统的接受度如何? 2) 学生对自我写作能力的认知状况如何? 3) 国内外智能评改系统评分有无差异?

### 2.3. 研究工具

本研究以在线评改系统批改网和 Grammarly 为例, 主要采用定量研究的方法。具体研究工具包括, 问卷调查和作文样本分析。

#### 2.3.1. 问卷调查

我们对研究对象进行了英语写作课的课程前与课程后的问卷调查, 问卷共 24 题, 包括选择题、量表题以及问答题。问卷以匿名形式呈现, 内容涉及学生个人学习背景、学生对自己英语写作水平的认知程度、学生对智能评改系统的认可度、对本学期写作课的期望授课方式以及对英语写作课的需求情况等。

#### 2.3.2. 作文样本分析

以 30 名研究对象的四篇英语作文为样本, 比较批改网与 Grammarly 两种在线系统的评分情况。这四篇代表作文的题材为记叙文与说明文, 比如叙述一次经历, 描述一个最喜欢的旅游景点, 从因果的角度讨论人们使用智能手机的上瘾现象, 以及分析在线授课的利与弊。

## 3. 结果与讨论

### 3.1. 英语专业学生对智能评改系统的接受度

课程前的问卷中, 86.21%的研究对象认为智能评改系统对自己的英语习作水平会有提高, 10.34%的研究对象认为不确定, 需要继续使用一段时间再判断, 3.45%的研究对象认为智能评改系统不会对写作有帮助。在授课前, 68.97%的研究对象主动使用过智能评改系统对其他英语作业(非英语作文)进行过校对, 31.03%的研究对象表示之前不太了解, 了解后会使用智能评改系统。经过一学期的追踪, 在课程后的问卷中, 52%的研究对象认为在写作课结束后, 自己会继续使用智能评改系统进行写作评改; 36%的研究对象认为要根据写作的内容以及剩余时间的宽裕度, 可能选择使用智能评改系统; 12%的研究对象则不会继续使用智能评改系统。数据表示在经过一个学期的追踪使用后, 不考虑写作内容以及时间剩余的情况下, 一定会主动在课后使用智能评改系统的研究对象反而由 86.21%降至 53%。在为期一个学期的使用追踪中, 研究对象对智能评改系统的肯定度反而降低了, 究其原因, 笔者将在下文中以学生作文为例详细进行讨论。针对作文评改方面, 研究对象的问卷结果表示, 86.21%的研究对象更希望通过教师的评价提高写作学习效率; 其次分别是 58.62%, 通过智能评改系统评价; 44.83%, 自我评价以及 24.14%, 与其他同学互评。虽然超过一半的研究对象表示可以通过评改系统进行辅助学习, 但教师传统的评改方式依然是目前研究对象最能接受的方式。而与同学互评(生生互评)则相对被选择最少。研究对象对生生互评的信度和效度依然持有保留观点。近年来, 很多研究侧重讨论写作教学的评价体系, 比如在产出导向法的指导下, 很多研究探讨了教师指导下的生生互评是如何提高写作的教学效果等。虽然在这些研究中证实了生生互评会显著提高学生的学习效果, 也提供了一些可以实际操作的教学方法, 但要让学生主动接受

这种评改方式，改变他们以往的观点还有待更多的教师指导、教学观察以及学习追踪。

### 3.2. 英语专业学生自我写作能力的认知状况

课程前与课程后的自我英语写作水平认知情况(满分为 10 分)如图 1 所示。在课程前的问卷中，研究对象期待在英语写作课上有所提高的排名前三的分别为语法、结构和内容。经过一学期的学习，整体来看研究对象在写作内容、结构和语法方面自我认为有所提高，但在词汇方面反而打分稍有下降。这个打分在研究对象比较关心的三个方面均有所提高。写作水平的提高是一个长期积累的过程，在本研究中，研究对象的课程前后的自我认知提高这在一定程度上肯定了本学期写作课程的教师授课方式和方法，也是对本学期课程教师与智能评改系统共同评改作业方式的一种认可。但考虑到问卷中的一些主观问题的设计，问卷采用匿名方式，这使得笔者无法结合研究对象的作业单独分析每位研究对象的认知变化和变化的原因。在之后的研究中，学者可以继续进行深入探讨。

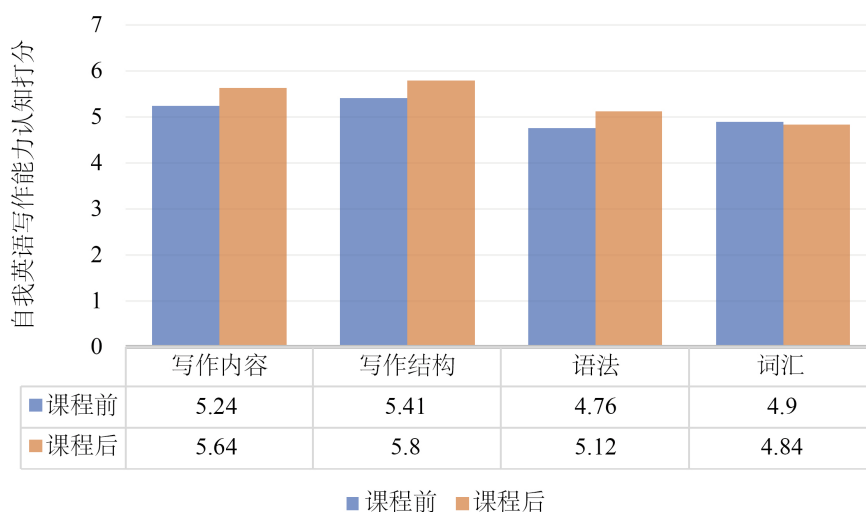
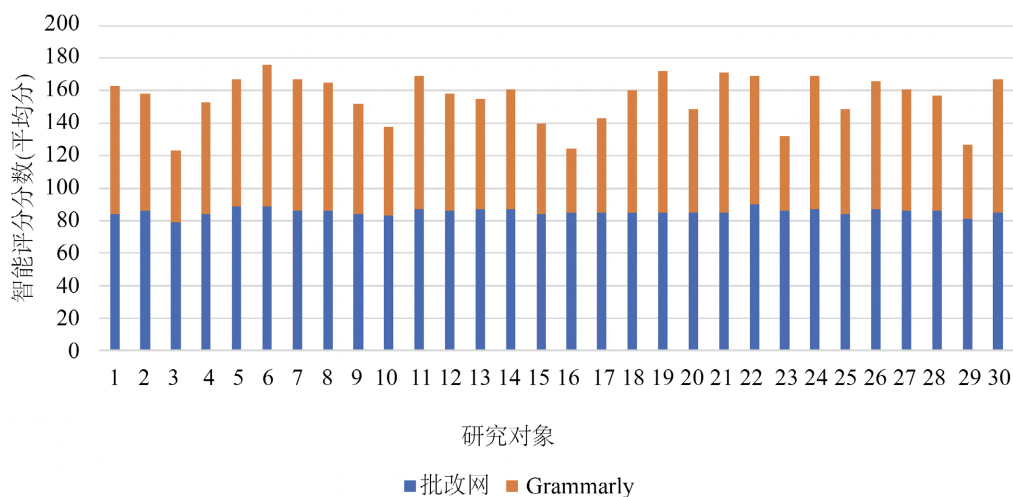


Figure 1. Cognitions of English-major students' self-writing skills for pre- and post-course  
图 1. 课程前与课程后英语专业学生自我写作能力的认知

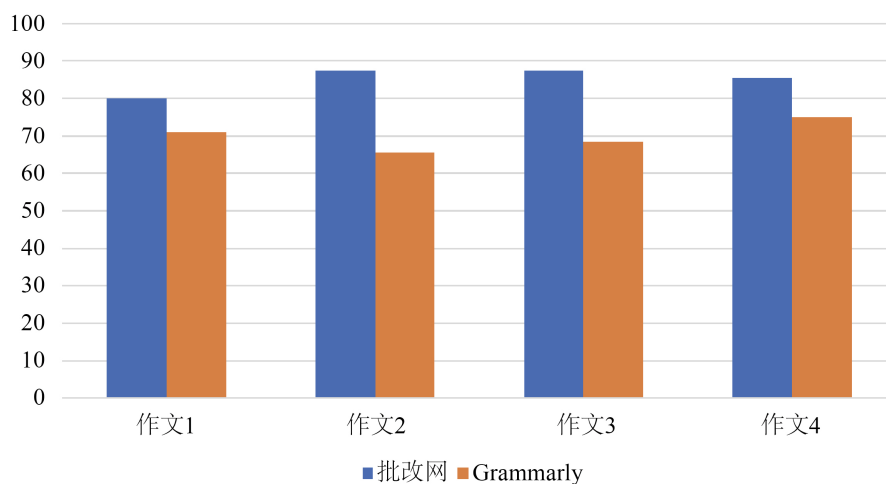
### 3.3. 国内外智能评改系统评分差异

批改网是一个用计算机自动批改英语作文的在线系统。它的原理通过对比学生作文和标准语料库之间的距离，并通过一定的算法将之映射成分数和点评。Grammarly 是通过使用高级机器学习和深度学习在内的各种创新方法，也同样根据不同类型的文章进行校对以实现纠错与打分。本研究以国内智能评改系统批改网和国外智能评改系统 Grammarly 为例，对 30 位研究对象的四篇指定话题英语作文进行了评分对比与分析。图 2 为 30 位研究对象以四篇作文为例的平均分。蓝色表示批改网的平均分，橙色表示 Grammarly 的平均分。蓝色与橙色的长度差异性越大则表示两个批改系统的打分差异越大。以研究对象 3、16、23 和 29 为例，我们可以观察到 Grammarly 的整体平均分明显低于批改网。图 3 为批改网和 Grammarly 两个评改系统对四篇作文的打分平均分。蓝色表示批改网的平均分，橙色表示 Grammarly 的平均分。从图中我们可以直观看到 Grammarly 的评改得分普遍低于批改网的打分。四篇作文以记叙文和说明文为题材，这也表示与打分与作文题材关系并不大。结合整体平均分，我们将具体以研究对象 3 的分析在线授课利弊的作文内容为例，对两个改评系统进行对比。图 4 和图 5 为批改网的具体评改信息。图 4 为针对一篇作文整体的评改信息，包括作文中的各种语法句法警示，分别指出了如标点错误、名词错误、句子



**Figure 2.** The average scores of the 30 participants according to the two intelligent assessment systems, Pigai and Grammarly

**图 2.** 批改网与 Grammarly 两个评改系统对 30 位研究对象的打分平均分



**Figure 3.** The average scores of the four essays according to the two intelligent assessment systems, Pigai and Grammarly

**图 3.** 批改网与 Grammarly 两个评改系统对四篇作文的打分平均分

错误、搭配错误等。图 5 为整体统计后按句点评，其中会详细的以中文为评改语言，提示学生作文中错误之处以及提供推荐表达。这篇作文批改网的打分为 83.5 分。图 6 为同样的一篇作文 Grammarly 评改系统进行的评改。我们可以看到，比如在正确表达方面，有 52 条警示，在表达清晰度方面等 Grammarly 评改系统以进度条的方式进行展现，进度条越靠右边，表示完成度越好。如图 6 所示，这篇作文的整体建议是表达清晰度有些许模糊不够清楚(a bit unclear)。Grammarly 评改系统中会直接在原文画出需要改进之处并提出意见，相比于批改网更直观一些。同样一篇作文，Grammarly 评改系统的打分为 20 分。我们可以清楚的看到两个评改系统对同一篇作文打分的差距非常大。究其原因，很大程度主要是由于两个评改系统的标准语料库与考查方面不同导致的。批改网主要是以修改各种题材的英语作文为导向，语料库收集了很多适合中国学生的写作需要注意的方面，考虑到中国学生容易犯的语法以及表达错误等。针对作文中的每一句话，评价与修改都很详细；而 Grammarly 评改系统则在纠正语法用词标点错误的基础上



Figure 4. Overall assessment information of Pigai  
图 4. 批改网的整体评改信息

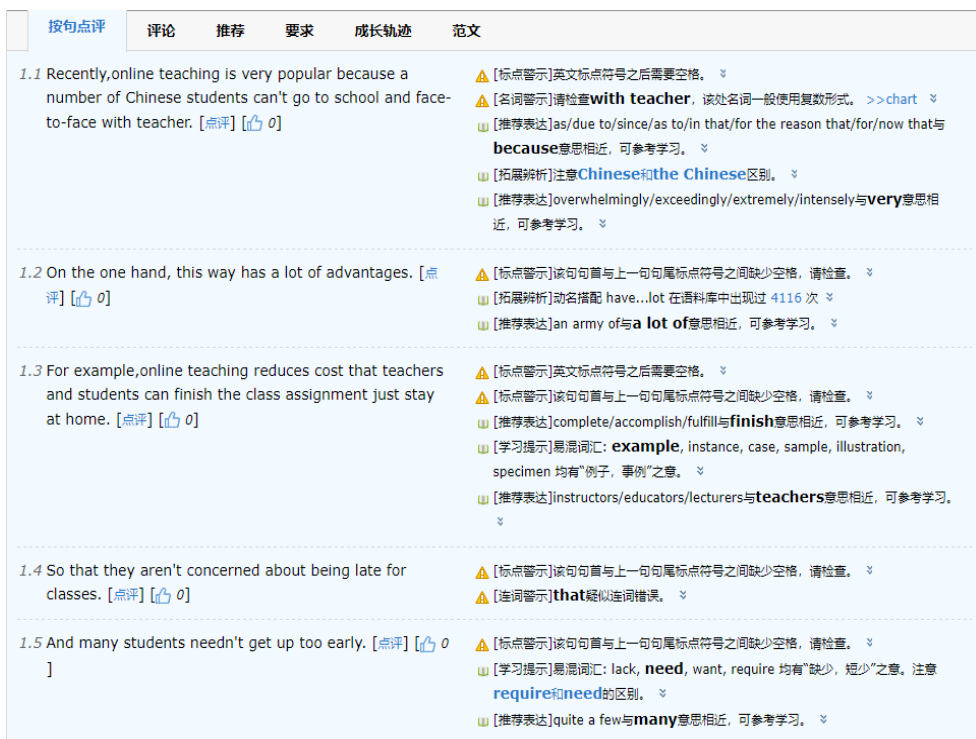


Figure 5. Detailed assessment information of Pigai  
图 5. 批改网的详细评改信息

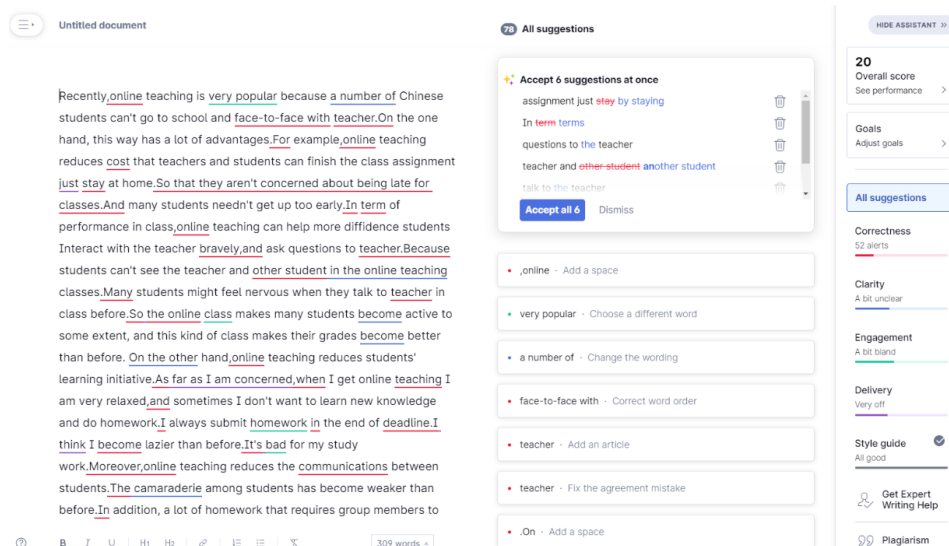


Figure 6. Overall assessment information of Grammarly  
图 6. Grammarly 评改系统的整体评改信息

同时考虑英语写作的语气与风格。它的使用人群更广一些，语料库的范围及类型也更多。但在长句与被动句的修改方面，两个评改系统都有待进一步提升准确性。这也是导致研究对象在学期末追踪反馈时表示一定会主动在课后使用智能评改系统比例下降的原因之一。

#### 4. 教学建议与启示

首先，不可否认的是智能评改系统一定程度辅助解决了教师批改费时低效的现象，缓解了教师的批改压力以及工作强度，并为教师提供了打分参考。评改系统中提供的修改参考也在某种程度上弥补了教师批改的局限性，智能评改系统可以指出一些教师人工批改中忽略的写作问题。但智能评改系统也会出现错误评改，一些正确的表达被系统标记为不够准确，而且评改系统修正的内容不能全盘接受，需要学生自己或在教师的指导下进行再次复查。其次，在教学过程中授课教师应该正确引导学生对智能评改系统的认识。智能评改系统的打分固然是一个重要的参考标准，但是最终的课程得分并不会以智能评改网的打分为准，具体的反馈和建议对学生应该是更重要的。让学生知道如何去修改，怎样的表达更为准确及地道的传递信息才是写作课最终达到的目的之一。根据本研究的调查结果，通过一学期对研究对象的学习追踪，学生更希望授课教师提供作文的评改而不是主要依靠智能评改系统或者同学之间互相进行写作评价，这也为之后的英语写作评改模式的研究提供了一定的背景基础。智能评改系统的使用利弊分明，授课教师应该充分利用其优势，避开系统的弊端，教师可以在之后的英语写作教学中考虑利用智能评改与教师人工评改相结合的方式进行作文点评，在未来的研究中，研究者也可以着重进行这方面的写作评改体系探讨。

#### 基金项目

本文系第十一批中国外语教育基金项目“人工智能辅助下‘产出导向法’在英语口语及写作教学中的设计与实施研究”(ZGWYJYJJ11A057)的成果。

#### 参考文献

- [1] 霍俊燕. 课程思政视阈下大学英语写作网络直播课教学探索与实践[J]. 海外英语, 2021(24): 145-146.

- 
- [2] 徐雪林. 课程思政在高校英语专业写作课程中的实践研究——以四川某院校英语学院“英语基础写作(3)”课程为例[J]. 英语教师, 2022, 22(22): 109-112.
- [3] 文秋芳. “产出导向法”与对外汉语教学[J]. 世界汉语教学, 2018, 32(3): 387-400.  
<https://doi.org/10.13724/j.cnki.ctiw.2018.03.008>
- [4] 宋皓. 大学英语写作网络智能评分与教师人工评分的对比研究[J]. 教育现代化, 2016, 3(13): 72-73.  
<https://doi.org/10.16541/j.cnki.2095-8420.2016.13.035>