International Journal of English Literature and Social Sciences
Vol-10, Issue-2; Mar-Apr, 2025

Peer-Reviewed Journal
Journal Home Page Available: https://ijels.com/
Journal DOI: 10.22161/ijels

A Study of Human-AI Interaction Patterns in Artificial Intelligence-Assisted Second Language Writing

Fanxi Shen

School of Foreign Languages, Zhejiang University of Finance and Economics, Hangzhou, China.
ORCID: https://orcid.org/0009-0005-5424-3723

Received: 05 Feb 2025; Received in revised form: 07 Mar 2025; Accepted: 11 Mar 2025; Available online:16 Mar 2025

Abstract— *The aim of this project is to investigate human-computer interaction patterns in AI-assisted second language writing in order to explore students' strategies and patterns when collaborating with generative AI tools, as well as key hidden states in AI-assisted writing. The project utilizes Hidden Markov Models (HMM) and process mining techniques to analyze the potential patterns of student interactions with AI and evaluate the efficacy of different patterns. The findings will reveal the cognitive mechanisms of human-computer interaction in second language writing, expand the theoretical framework of human-computer collaboration, and provide empirical evidence for effectively guiding students to utilize AI to address key questions in second language writing research.*

Keywords— ***Artificial Intelligence; Second Language Writing; Hidden Markov Models; Human-AI Interaction.***

LITERATURE REVIEW

With the rapid development of technology, Artificial Intelligence (AI) has deeply penetrated into the field of education, and the teaching of writing is no exception. The intervention of AI technology has brought new modes and methods to the teaching of writing, and the scope of application has been constantly expanding, from automatic writing assessment systems to advanced language models. This not only provides students with richer learning resources and instant feedback, but also brings convenience and innovation to teachers. However, like any emerging thing, the application of AI in writing teaching faces many challenges, such as the accuracy of the technology, the impact on students' thinking development, and academic integrity. Therefore, it is of great practical significance to study the current situation, advantages, challenges, and response strategies of AI in writing teaching.

1. AI Application in Writing

1.1 Automated Writing Assessment System

Automated Writing Evaluation (AWE) systems are an early and widely used form of artificial intelligence in writing instruction. Huang Jing and He Huaqing (2018) point out in “Research on the Impact of Human-Machine Feedback on Students' Writing Behavior” that Page Essay Grade (PEG), developed by U.S. scholars in the 1960s, was the world's first automatic assessment system for writing, and initially it was only capable of scoring essays, with a relatively single function. With the continuous progress of technology, today's AWE system has become increasingly powerful, and can automatically generate scores while providing students with diagnostic feedback from multiple dimensions, such as grammar, vocabulary, structure and

content. China's Sentencool is a typical example, which has been widely used in English writing teaching in colleges and universities for its advantages of timely marking, recognizing simple errors and providing knowledge of relevant grammar and vocabulary (Gu, Chenghua, and Wang, Li, 2012).

In Huang Jing and He Huaqing's (2018) study, teaching practice was carried out with the help of Sentence Cool Critique Network. Students completed their first drafts and submitted them to the Critique Network, made revisions based on their feedback, followed by peer review, and some students were able to obtain teacher feedback, and finally completed a self-assessment report. The results of the study showed that the human-computer feedback prompted students to increase the opportunity of writing multiple drafts, which is conducive to the realization of the process of writing. An analysis of the number of drafts completed in the students' writing tasks revealed that students completed a total of 417 drafts in the three writing tasks, with 75 drafts in the first draft and 342 drafts in the revised drafts, an average of nearly six drafts per person per writing task. Meanwhile, the quality of students' writing texts improved to varying degrees. In terms of vocabulary spelling and punctuation, students received more machine feedback than manual feedback, which suggests that the AWE system has an advantage in providing feedback at these basic levels. In terms of text quality, both the manual scoring and the machine evaluation showed that the human-machine feedback had a significant positive effect on the quality of students' writing, especially in vocabulary, such as in Writing Task 1, where the academic vocabulary increased significantly in the final draft compared to the first draft, and the difference was significant.

1.2 Intelligent tools based on language models

The application of language modeling represented by ChatGPT in writing instruction has gradually become a hot research topic. Da Yan (2023) conducted a one-week practical study in Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. The study found that students were able to quickly acquire basic skills in writing using ChatGPT and increase their proficiency and ability through collaborative activities. week-long practical study in which undergraduate English majors were exposed to ChatGPT in a L2 writing practicum. The study found that the students were able to quickly acquire the basic skills of writing with ChatGPT, and increased their proficiency and ability through collaborative activities. Students recognized the speed and quality of ChatGPT in generating text that was well structured, had few grammatical errors, and could be produced in different styles, such as academic or casual, depending on the requirements.

Ahmad Azmi AbdelHamid Esmaeil et al. (2023) in Understanding Student Perception Regarding The Use of ChatGPT in Their Argumentative Writing: a Qualitative Inquiry" in which students' perceptions of the use of ChatGPT in their argumentative writing were explored in depth through a qualitative study of 17 students. Students perceived ChatGPT as having significant advantages in providing information and guidance, stimulating writing ideas, and saving time. In terms of information provision, ChatGPT can quickly retrieve relevant content from multiple disciplines and fields, providing students with a wealth of materials and ideas for their writing. It also generates a logical article framework and paragraph content based on the keywords or questions entered by students, helping students to overcome the obstacles in their thinking when writing. In addition, the immediacy of ChatGPT allows students to access help at any time without the constraints of time and space, which greatly improves the convenience of learning.

2. Positive Impact of AI on Writing

2.1 Providing immediate feedback

Automatic Writing Evaluation (AWE) systems and language models can provide students with immediate feedback, which is important for writing instruction. In "A Study of the Effects of a Mixed Human-Computer Feedback Environment on Learning Engagement and Second Language Writing Proficiency," Zhang Ya and Jiang Zhanhao (2022) pointed out that AWE feedback provides frequent, process-oriented diagnostic assessments during the revision process, and that this kind of instantaneous feedback can stimulate students' learning engagement. In their study, students in the experimental group, who received mixed human-computer feedback, showed a significant increase in behavioral and cognitive input compared to the control group, who received only teacher feedback. Specifically, the experimental group had a higher

rate of feedback uptake, and students were able to more actively revise their essays based on the feedback; longer revision time indicated that students paid more attention to the feedback and were willing to spend time to make improvements; and greater awareness of the feedback and ability to use metacognitive strategies, and students were able to better understand the content of the feedback and use metacognitive strategies to monitor and adjust the writing process.

Liu Yingliang, Liu Shengnan, and Yang Jincai (2022) mentioned in "Exploration of Human-Computer Collaborative Teaching and Application in the Perspective of Socio-Cultural Activity Theory -- Taking iWrite Collaborative English Writing Teaching as an Example," that the iWrite English writing teaching and review system can provide timely review and feedback on students' online texts in terms of four dimensions: language, content, chapter structure, and technical specifications. The iWrite English Writing Teaching and Review System, mentioned in iWrite Collaborative English Writing Teaching Example, provides timely review and feedback on online texts submitted by students in four dimensions: language, content, chapter structure, and technical specifications. The system categorizes and counts errors in essays into five categories, such as syntax, lexis, collocations, technical specifications, and other categories, and subdivided into 69 subcategories, and provides concise suggestions for revision. This detailed and immediate feedback helps students identify and correct problems in writing in a timely manner to improve the quality of writing. For example, if students make mistakes in verb tense or singular or plural nouns during the writing process, the system is able to quickly recognize them and give them suggestions for correct usage, so that students can correct their mistakes in time and avoid repeating them in their subsequent writing.

2.2 Enriching Teaching Resources

Artificial intelligence technology brings rich and diverse teaching resources for writing teaching. Jin Liang and Yang Jinsong (2022) pointed out in "Exploring the Transmutation of Higher Vocational Curriculum Structure in the Context of Artificial Intelligence-Enabled Education -- Taking Foreign Language Courses as an Example" that teaching resources present a diversified trend under the support of AI. Teachers can build rich multimodal teaching resources with the help of AI technology, for example, in the senior medical humanities English course, teachers adopt the task-driven method to encourage students to make small videos related to "flu". In the process of production, students not only master medical knowledge, but also express themselves in English, thus improving their English listening, speaking, reading, writing and translation skills. This multimodal presentation of teaching resources integrates text, images, audio, video and other elements to make the learning content more vivid and interesting, which helps to improve students' motivation and participation.

When students in Da Yan's (2023) study used ChatGPT, they found that it provided a wealth of information and examples from multiple disciplines and fields. When students write about scientific and technological topics, ChatGPT can provide the latest scientific research results, relevant cases and opinions from different scholars to help students broaden their ideas and enrich the content of their essays.ChatGPT can also provide different styles of writing examples according to the needs of students, such as formal academic style, easy-to-understand popularization of science and technology style, so that students can learn a variety of ways to express themselves and increase the flexibility and adaptability of writing. and improve the flexibility and adaptability of writing.

2.3 Promoting Personalized Learning

Artificial intelligence helps to personalize writing instruction. According to Liu Yingliang, Liu Shenglan, and Yang Jincai (2022), the resource library of the iWrite system contains a variety of resources for different writing problems, and students can browse and collect the appropriate resources for personalized learning according to their specific problems, such as argumentative essay conceptualization skills, articulation language for high scores in essays, and sentence variety. The system can also analyze students' writing situation and provide teachers with information on students' individual writing strengths and shortcomings, so as to facilitate teachers' precise tutoring. For example, for students who have deficiencies in vocabulary use, teachers can recommend vocabulary learning resources based on the information provided by the system and guide students to conduct targeted practice.

Students in Ahmad Azmi AbdelHamid Esmaeil et al.'s (2023) study perceived that ChatGPT could provide

personalized interactions and guidance based on their needs and help them discover new ideas and perspectives. When students encountered difficulties in writing, ChatGPT was able to provide personalized suggestions and solutions based on the problems and requirements entered by the students. If a student's viewpoint on a certain topic is not clear enough, ChatGPT can guide the student to think about the issue from different perspectives, provide more arguments and methods of argumentation, and help the student to refine his/her viewpoint and improve the depth and breadth of his/her writing.

3. Challenges of AI in Writing Teaching

3.1 Accuracy of Feedback

Despite the immediacy of the feedback provided by AI, accuracy and relevance could be improved. Adam Pfau et al. (2023), in Exploring the potential of ChatGPT in assessing L2 writing accuracy for research purposes, investigated the A study of ChatGPT's ability to assess L2 writing accuracy found that ChatGPT frequently underestimates the total number of errors, although it has a high correlation with manual coding in error detection. When analyzing 100 second language compositions at different levels of proficiency, ChatGPT was found to miss errors such as hyphenated sentences, sentence fragments, incorrect or missing prepositions and articles. It is also not stable enough in recognizing the types of errors, and the classification of errors in the same composition may vary at different times.

Huang Jing and He Huaqing (2018) mentioned that most of the feedback from automatic assessment systems such as SentenceCoolCritique.com focuses on the vocabulary section and the more mature grammatical analyses, which tends to only point out the problem and cannot provide correct examples. For some complex grammatical errors or errors in specific contexts, the system cannot accurately identify and correct them. In terms of essay content and chapter structure, the feedback from the automated assessment system is also lacking, and it cannot analyze the logical coherence of the essay and the adequacy of the argument in depth, thus failing to provide students with comprehensive and targeted suggestions for improvement.

3.2 Potential impact on students' thinking skills

Over-reliance on artificial intelligence may have a negative impact on the development of students' thinking skills.Araz Zirar (2023) in Exploring the impact of language models, such as ChatGPT, on student learning and assessment states that If students use language models without critically assessing them and rely only on the information they provide, it may lead to a lack of critical thinking and analytical skills. In the writing process, students who rely too much on the content generated by ChatGPT may lack the ability to think independently and explore issues in depth. When students encounter writing tasks and use ChatGPT-generated articles directly without thinking and analyzing them on their own, they cannot really exercise their thinking skills and find it difficult to improve their understanding and application of knowledge. In the long run, students may develop a habit of dependency and be unable to complete writing tasks independently when faced with situations without the assistance of AI.

Students in Ahmad Azmi AbdelHamid Esmaeil et al.'s (2023) study were also concerned that the use of ChatGPT would cause them to become dependent, which would reduce their learning and critical thinking skills. They argued that excessive use of ChatGPT reduces their opportunities for active thinking and exploration, which is not conducive to the development of independent problem-solving skills and creative thinking, and has a negative impact on long term development.

3.3 Issues of Academic Integrity

The use of artificial intelligence in writing instruction raises serious academic integrity concerns. Da Yan (2023) found that students believe that the use of ChatGPT may undermine educational equity because it helps students generate essays quickly, giving those who use the tool an unfair advantage over the competition. Some students may use ChatGPT to cheat on exams or assignments, a behavior that is difficult to recognize by existing plagiarism detection software. Since the text generated by ChatGPT is somewhat original, it is difficult for existing plagiarism detection systems to determine whether it is plagiarized content or not, which increases the risk of plagiarism among students.

Araz Zirar (2023) also pointed out the increasing use of language models by students to generate assessment assignments and the low reliability of existing detection tools. Current detection tools are often unable to accurately determine whether a text has been generated by a language model or to what extent the content of the language model

has been used. This gives some students an opportunity to take advantage of the situation, and they may obtain high scores through improper means, which undermines the fairness and seriousness of academics. Such behavior is not only unfair to other students, but also not conducive to students' own learning and growth, and cannot truly test students' knowledge mastery and ability level.

3.4 Dilemma of Teacher Role Change

In the process of integrating AI into writing instruction, teachers face the challenge of role transformation. Jin Liang and Yang Jinsong (2022) suggested that in traditional teaching, teachers are the main body of the curriculum, but in the context of AI-empowered education, teachers need to transform into the leaders and instructors in the "teacher-student learning community ', which puts forward new requirements on teachers' teaching concepts and abilities. However, some teachers may have difficulty adapting to this transformation. In the traditional teaching mode, teachers are used to dominating the classroom and imparting knowledge in a single way. In contrast, in an AI-assisted teaching environment, teachers need to learn to utilize new technological tools to guide students to learn independently and interact and collaborate with them more effectively.

Liu Yingliang, Liu Shenglan and Yang Jincai (2022) believe that teachers need to be skilled in the use of human-computer collaborative systems and learn to collaborate with "AI" in teaching, but some teachers currently have insufficient knowledge of and ability to apply related technologies. Some teachers may have a limited understanding of the functions of automatic writing assessment systems and language modeling and do not know how to make full use of these tools to assist teaching. In the process of using them, they may encounter technical problems that cannot be solved or fail to organically integrate the technology with the teaching content, resulting in the inability to fully utilize the advantages of AI in writing teaching.

4. Summary

Comprehensively, it can be seen that the application of artificial intelligence in writing teaching brings many opportunities as well as faces many challenges. It has demonstrated significant advantages in providing immediate feedback, enriching teaching resources, and promoting personalized learning, providing new impetus and ways for the development of writing teaching. However, the issues of accuracy and relevance of feedback, the potential impact on students' thinking skills, academic integrity, and the dilemma of changing teachers' roles need to be addressed in future research and practice. By optimizing the technology, guiding students to use it correctly, strengthening academic integrity education and supervision, as well as improving teachers' technology application ability and other strategies, the role of AI in writing teaching can be better exploited, the deep integration of AI and writing teaching can be realized, the quality and effect of writing teaching can be improved, students' comprehensive writing ability and innovative thinking can be cultivated, and students' future development can be laid with a solid foundation. Future research can further explore how to apply AI more effectively in different teaching scenarios and disciplines, as well as how to establish a more complete evaluation system to comprehensively assess the impact of AI on writing teaching. It can also investigate how to use AI technology to develop students' digital literacy and critical thinking skills, so that they can better adapt to the development needs of the future digital society.

EXPERIMENTAL PROCEDURE

In today's era of deep integration between artificial intelligence and education, it is of great practical significance to explore the dynamics of cooperation between advanced English academic level writers and AI tools. Focusing on this area, this study carefully selected 10 first-year master's degree students as participants, all of whom were enrolled in the Foreign Languages, Literatures and Applied Linguistics program, all of whom had English as their second language, and all of whom had more than ten years of English learning experience. The English master's students were chosen for the study because they are familiar with English academic writing and are able to demonstrate the characteristics of advanced English academic level writing under high standards and complex ideas, which helps to gain insights into the dynamics of this type of writers' work with AI tools.

Prior to participating in the study, all subjects had some

basic knowledge of AI use and had received classroom training in academic writing and AI interaction. The training covered a variety of key aspects, including a detailed explanation of the APA format, which is essential in academic writing, so that students could master how to cite references in a standardized way to avoid academic misconduct, and training in writing AI prompts, so that students could accurately communicate their needs to the AI tool and get a more responsive response.

Prior to the formalization of the experiment, the research team designed and distributed a questionnaire to gather comprehensive information about the participants' backgrounds. The questionnaire covered several key dimensions, including the teacher support and technical support dimension, which focuses on the help and resources that students receive from teachers and technology in their daily learning; the AI literacy dimension, which examines students' understanding, mastery, and application of AI technology; the technology acceptance dimension, which looks at students' acceptance of, attitude toward, and willingness to use the AI tools; and the human-computer interoperability dimension (Zhai, N., & Ma, X, 2022; Shen, Y., & Cui, W, 2024), which focuses on exploring the interaction patterns, collaboration effects, etc. between students and AI tools. Through this questionnaire, the researchers were able to gain a more comprehensive understanding of the basic situation of the participants, providing strong background data support for the subsequent experimental study.

This study utilized an offline experimental design in which an experienced instructor distributed writing tasks to students in class. Students were required to complete a challenging writing task in the same class, i.e., to write an argumentative essay of no less than 500 English words on the topic “Is There a Critical Period for Second Language Acquisition?”. To ensure the quality of the essay, the task has clear and strict requirements on the content and structure of the writing. In terms of content, students are required to support their arguments with well-supported evidence and examples from reliable sources, which requires students to have good literature searching and filtering skills, and to be able to select strong evidence from a wide range of academic resources to support their own views. The structure of the essay must include a clear introduction to capture the reader's attention and introduce the topic of the essay; a body that covers both thesis and counter-argument, demonstrating the student's ability to think deeply about the issue and analyze it; and a strong conclusion that summarizes and sublimates the content of the essay. At the same time, students must cite all the literature correctly in strict accordance with the APA format, which not only reflects the standardization of academic writing, but also helps readers to trace back and check the relevant information. In terms of word count, it is clearly stated that the 500-word requirement does not include a bibliography section, and that the word count should not be less than 450 words or more than 550 words to ensure that the essay can adequately address the point without being too long or too short. In addition, participants were required to complete the writing task in less than 60 minutes, which placed high demands on students' time management and writing efficiency. In order to simulate a real-life writing situation, students were allowed to use any tools they deemed necessary during the experiment, including but not limited to KIMI and Google Scholar, so that they could fully utilize their abilities to authentically demonstrate the process of working with AI tools.

During the course of the experiment, the research team used screen recordings to fully capture the participants' interactions with the writing task and the various tools they chose to use. This recording method was able to record detailed information about the steps students took during the writing process, and the frequency and manner in which they used the tools, providing rich data for subsequent analysis. Prior to any sequence analysis or process mining, the final outcome of the writing task was scored by a professional university lecturer immediately after the data collection phase. The lecturer will make a comprehensive assessment based on the pre-established scoring criteria in terms of content, structure, language expression, literature citation and other aspects of the article to ensure the objectivity and fairness of the scoring. At the end of the experiment, the research team distributed the same questionnaire again in order to compare and analyze the changes in the relevant dimensions of the students before and after the experiment, and further explore the impact of the experimental process on the students.

Data Coding and Processing

In today's era of globalization, the integration of second language learning and artificial intelligence has become an important research direction in the field of education. This study focuses on the collaborative interaction between L2 learners and AI in the English writing process, aiming to deeply analyze the behavioral characteristics of the learners and the adjustment mechanism of their writing strategies in this process. To this end, the research team carefully collected video data from 11 L2 learners during their collaborative English writing tasks with the AI, which lasted for about 564 minutes and covered all the micro-behaviors of the students during the writing process, ranging from minor operations such as pasting and editing to searching for information. These rich data provide a solid foundation for subsequent in-depth analysis.

In order to conduct a systematic and in-depth analysis of these recorded screen data, the research team chose the ELAN software (Lixin Zhang, 2012) and adopted the method of quantitative content analysis. This method encodes students' writing behaviors in a rigorous chronological order, thus generating a sequence of observable states, which provides precise basic data for the subsequent more complex Hidden Markov Model analysis. Specifically, the analytic approach is divided into three key phases: an open coding phase, an axial coding phase, and a selective coding phase. In the initial open coding phase, the researchers meticulously broke down the data into small units, adding descriptors or codes for each student behavior related to the writing process. These descriptors accurately summarized a specific aspect of the writing process; for example, a student copying textual content from an AI tool might be labeled as “Getting content from AI and pasting it”; modifying and editing the pasted content, labeled “Editing AI-supplied content ”; and looking up information on the Web is labeled as “Finding writing information online.” With such detailed labeling, each behavior is clearly recorded and defined, providing rich raw information for subsequent analysis. As the analysis progressed, the axial coding stage was entered. In this phase, the researchers categorized the numerous codes generated in the previous open coding phase. Similar codes related to acquiring external content such as “acquiring content from AI and pasting” and “acquiring content from other resources on the web and pasting” were grouped together and defined as “external content acquisition behaviors ”; codes related to various editing operations, such as “editing content provided by AI”, “editing and modification of self-created content”, etc., are categorized as “text editing behavior “. Through this categorization, the originally complicated codes become more organized for further analysis. Finally, there is the selective coding stage, in which researchers distill from the categorized codes the themes that express each group of content. These themes are a high level summary of the essence of individual or cluster codes, and can reflect more intuitively the behavioral patterns and characteristics of students in the writing process. After these three stages of meticulous analysis, the research team coded a total of 2,136 actions of 11 students, laying a solid data foundation for subsequent in-depth research.

Hidden Markov modeling, as an extension of the basic Markov chain, is uniquely suited for research in the field of learning sciences (Nguyen, 2024). It provides deep insight into the probability of a sequence consisting of specific random variables or states, and is powerful for analyzing complex dynamic processes. In this study, the researchers used the Python programming language to model sequences of students' writing behaviors based on the HMM model as a way to identify hidden states in the writing process. The central concern of this study is to explore how master's students make adjustments to their own writing strategies in the process of utilizing the AI assistant. In the actual writing process, cognitive activities such as writing plan development and text editing are often hidden and difficult to be directly observed. Yet, these activities have a crucial impact on the final writing outcome and how well students interact with the AI. With Hidden Markov Models (HMM), researchers can effectively model and analyze these implicit cognitive activities that are difficult to observe directly.

The HMM model can capture the potential patterns of students' interactions with the AI during the writing process by analyzing the sequence of students' writing behaviors, revealing the transfer patterns between different states. For example, model analysis may reveal that when students encounter a state of stuck writing ideas, they are more inclined to obtain a large amount of content from the AI and then enter the editing state of such content; while when they

are more satisfied with the writing content, they will be more likely to carry out the refinement of self-created content. Through such an analysis, it is possible to gain insight into students' thinking changes and strategy adjustments in the writing process. In the process of model construction, in order to determine the most appropriate number of model states, the research team comprehensively considered two important indicators, AIC (Akaikei Information Criterion) and BIC (Bayesian Information Criterion). These two indicators can help the researchers choose the model that most accurately reflects the characteristics of students' writing behavior among many possible model structures, thus ensuring the reliability and validity of the research results.

RESULT

In the study of collaborative English writing between second language learners and AI using Hidden Markov Model (HMM), model parameter selection is a key step. We conducted an in-depth analysis of the number of states of the HMM with the help of Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Log Likelihood, the results of which are shown in Fig. 1. The horizontal coordinates in the Figure1 represent the number of HMM states, ranging from 2 to 10, which is chosen based on the pre-experiments and the references to similar studies in the literature, aiming to comprehensively explore the performance of the model under different numbers of states. On the left side of the vertical coordinate is the criterion value, which follows the principle of the lower the better, and the value comprehensively takes into account the degree of model fitting to the data as well as the complexity of the model itself; on the right side of the vertical coordinate is the log-likelihood value, and the higher the value is, the better the model fits to the observed data, i.e., the model is able to better interpret the information in the data.

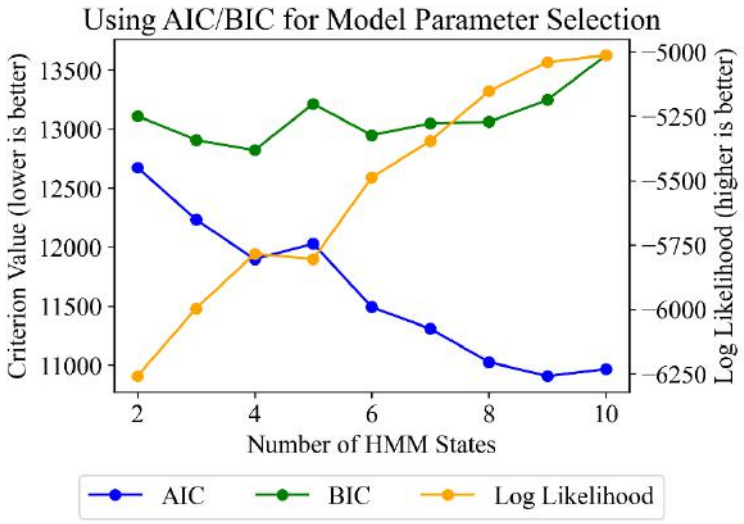

Fig1

Observing the trend of the curves in the figure, we can find that the blue AIC curve shows a continuous decrease with the increase of the number of HMM states. When the number of states is 2, the AIC value is relatively high, around 12600; when the number of states gradually increases to 10, the AIC value decreases to around 11000, which indicates that with the increase of the number of states, the model is able to capture more complex patterns and potential laws in the data, and thus fit the data better. However, this decreasing trend also hints at a potential risk of overfitting, as an overly complex model may overlearn the noise in the training data, resulting in poor generalization on new data.

The green BIC curve, on the other hand, is characterized by a decreasing and then increasing trend. The BIC reaches its lowest value around 12800 when the number of states is around 4. After that, as the number of states continues to increase, the BIC value starts to rise, and BIC evaluates the model not only by considering the model's goodness-of-fit, but also by imposing a stricter

penalty on the model's complexity. This means that BIC prefers models that explain the data better and are relatively concise in order to avoid overfitting problems due to model complexity.

The orange log-likelihood curve climbs from a low value at 2 states to a relatively high value at 10 states, from around -5100 to -5000. This intuitively reflects that as the number of states increases, the model's ability to fit the data increases, and it is able to better match the observed data distribution. However, as mentioned earlier, increasing the number of states simply in pursuit of high log-likelihood values may sacrifice the generalization ability of the model.

In the analysis of the hidden states of the HMM model (Figure 2), we finally identified four representative hidden states (State1 - State4), and the main writing behaviors of students and their probabilities of occurrence in each state are shown in Figure 2. In order to present this information clearly, each state is presented as a circular graph, which details the writing behaviors with higher probability (> 0.05) and their percentage in that state, and for behaviors with probability less than 0.05, they are uniformly labeled as "Actions with probability < 0.05".

State1 can be summarized as Read and Review and all of them belong to non-human interaction behaviors of students. In State1 (green circle), the behavior "ReadEssay" (reading an essay) dominates with a high percentage of 39.0%. This indicates that in this state, students spent a great deal of time and effort reviewing the essay they had already written, trying to thoroughly evaluate the essay in terms of its overall structure, logical coherence, grammatical accuracy, and relevance of the content, among other aspects. This is closely followed by "EditContent" with a percentage of 14.4%, which indicates that once students find problems in the process of reading an article, they will promptly revise and adjust the content in order to improve the quality of the article. In addition, "Delete Previously Pasted Generated Content" (Delete Previously Pasted Generated Content) accounted for 9.8%, "Read Generated Reference " (Read Generated Reference) accounted for 9.1%, and behaviors such as "Read Generated Reference" also reflect to some extent the students' prudence in screening and integrating AI-generated content as well as references, and that they screen the information they obtain, removing the parts that do not meet the requirements and retaining the useful information.

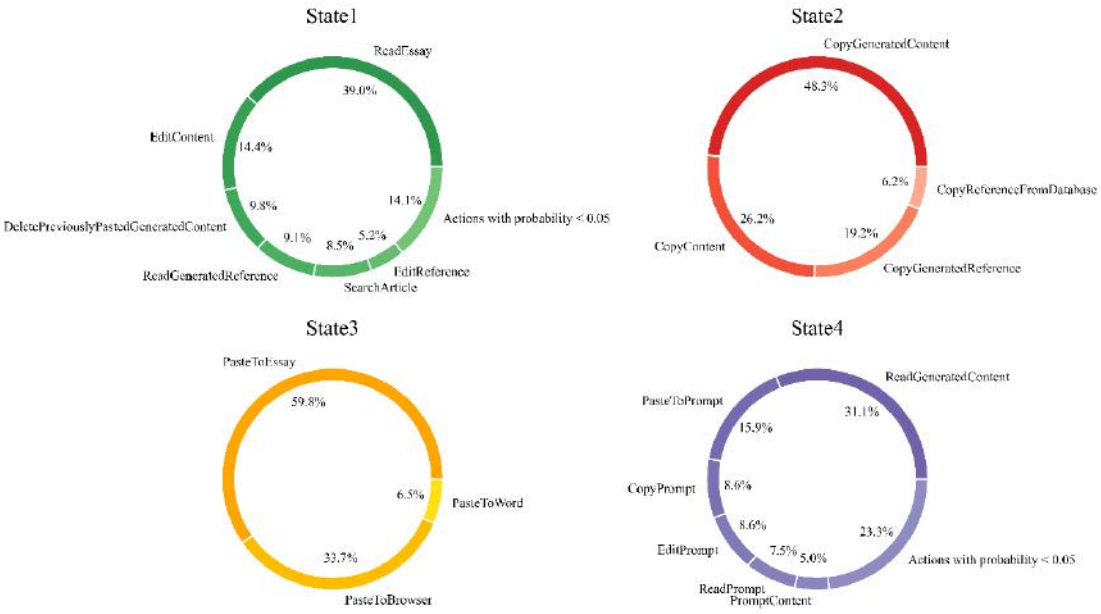

Fig.2

State2 can be summarized as copy. state2 (red circle) is notably characterized by "CopyGeneratedContent", which accounts for 48.3%, almost half of all behaviors in this state. This clearly indicates that in this state, students highly rely on AI-generated content to enrich their writing. "CopyContent" accounted for 26.2% of all behaviors, further indicating that students not only obtain content from AI, but also may copy relevant information from other

sources to fulfill their writing needs. This behavioral pattern may appear in the early stages of writing, when students face blocked ideas, lack of relevant materials, or lack of familiarity with the writing topic, they will use AI to quickly obtain a large amount of information to provide a basis and inspiration for writing.

State 3 can be summarized as Paste. state3 (orange circle) has "PasteToEssay" (Paste to Essay) as the main behavior, accounting for 59.8%. This means that in this state, students mainly worked on integrating various content previously acquired from AI or other sources into their essays. "PasteToWord" (Paste to Document) at 33.7% and 'PasteToBrowser' (Paste to Browser) at 6.5%, on the other hand, reflect that students are transferring content between different platforms and tools and organizing process. They may first paste the acquired content into a temporary document or browser for preliminary screening and editing, and then paste the finalized content into the article to construct a complete framework and content system of the article.

State4 can be summarized as prompt-related. since prompt involves interaction with the AI, the behaviors are all related to AI interactions. in State4 (purple circle), the percentage of "ReadGeneratedContent" is 31.1%, indicating that students in this state mainly read and understand AI-generated content in depth. "PasteToPrompt" accounts for 15.9%, which may mean that in the process of reading AI-generated content, students will, according to their own needs and understanding of the content, paste the relevant information into the prompt box to further interact with the AI in order to obtain content that better meets their expectations. In addition, behaviors such as "CopyPrompt" (8.6%) and "EditPrompt" (7.5%) also show that students are constantly adjusting the prompts for interaction with the AI, trying to optimize the prompts to guide the AI to generate more content. Optimizing prompts to guide the AI to generate more accurate and useful content to better meet their writing needs.

DISCUSSION

In terms of model parameter selection, AIC and BIC provide us with two different perspectives for weighing model complexity and goodness-of-fit.AIC focuses more on the model's ability to fit the data, and it encourages us to select complex models that more fully describe the data. In some cases, when our goal is to fit the available data as accurately as possible, such as in some data mining and exploratory analyses, AIC may lead us to choose models with a higher number of states. However, this choice may entail the risk of overfitting, making the model much less predictive on new data.

In contrast, BIC is more stringent on the complexity of the model while considering the model's goodness-of-fit. This makes BIC more inclined to choose models that are relatively simple and can explain the data better, which helps to improve the generalization ability of the model. In practice, BIC may be a more appropriate choice if we are more concerned about the performance of the model on new data in the future, such as in predictive modeling and classification tasks. Therefore, in this study, we need to consider both the AIC and BIC criteria together, combine the specific purpose of the study and the characteristics of the data, and find an optimal balance between the model's fitting effect and complexity, and determine the number of HMM states that is most suitable for the data as 4.

From the results of the hidden state analysis of the HMM model, the different states reflect that students present diverse behavioral patterns in the process of collaborative writing with the AI, which are interrelated and together constitute a dynamic writing process.The reading and editing behaviors embodied in the non-HMM behaviors of State1 students may be the students' initial self-assessment and optimization of the article, which is an indispensable part of the writing process. process, which is an integral part of the writing process and helps to improve the quality and logic of the essay.The behavior of extensive copying of AI-generated content in State2 manifests that copying is a major feature of the students' interaction with the AI. In teaching, teachers need to guide students to treat AI-generated content correctly, not only to learn to learn from it, but also to make in-depth thinking and innovation on the basis of it.The Paste in State3 shows that students are transforming the acquired information into a part of their own essays, which requires students to have a certain degree of information filtering and organizing ability. Educators can help students improve their skills in this area through relevant lessons and training, so that they can make better

use of AI and other resources to complete writing tasks.The adjusted behavior of interacting with AI in State4 shows that students are constantly exploring how to better collaborate with AI in order to obtain content that better meets their needs. This also suggests that future AI-assisted writing tools need to further optimize the interaction interface and features to better meet students' individual needs and improve the efficiency and effectiveness of students' collaboration with AI.

In addition, these hidden states do not exist in isolation from each other, but influence and transform each other. Students may switch between different states during the writing process according to their needs, problems encountered, and control of writing progress. This dynamic behavioral pattern provides rich information for us to deeply understand the collaboration mechanism between students and AI, as well as valuable references for further optimizing AI-assisted writing tools and teaching strategies. For example, we can design a more intelligent prompting system based on the behavioral characteristics of students in different states, provide corresponding suggestions and guidance when students are in a specific state, help students better complete the writing task, and promote the improvement of students' writing ability at the same time.

CONCLUSION

This study focuses on collaborative English writing between second language learners and AI, and analyzes it using Hidden Markov Model (HMM). In the selection of model parameters, with the help of Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and log-likelihood value, it is found that AIC decreases continuously with the increase of the number of states in HMM, which enhances the fitting ability but has the risk of overfitting; BIC decreases and then rises, which tends to be a succinct and well-fitting model; log-likelihood value rises with the number of states, and the overpursuit sacrifices the ability of generalization, and in combination with the HMM model, it can be identified as four states.

In the HMM model hidden state analysis, four representative states were identified. state1 focuses on reading and reviewing articles, which are non-human-computer interaction behaviors that reflect students' self-assessment and optimization; state2 features copying of content and is highly dependent on the AI to obtain writing materials; state3 focuses on pasting and integrating information into the article; and state4 focuses on prompt-related behaviors around interactions with the AI State3 focuses on pasting and integrating information into the article; State4 focuses on prompt-related behaviors around interaction with the AI, exploring better ways of collaboration.

The different states reflect diverse and dynamically related writing behavior patterns. This not only provides a basis for understanding the collaboration mechanism between students and AI, but also provides a reference for optimizing AI-assisted writing tools and teaching strategies, such as designing an intelligent prompting system to help students improve their writing skills. Future research can further explore how to promote the efficient collaboration between students and AI based on these states to improve the effect of second language writing teaching.

ACKNOWLEDGEMENT

This paper is supported by the research project [24XJKT069] at Zhejiang University of Finance and Economics. A Study of Human-AI Interaction Patterns in Artificial Intelligence-Assisted Second Language Writing

REFERENCES

[1] Adam, P., Charlene, P., & Yiran, X. (2023). Exploring the potential of ChatGPT in assessing L2 writing accuracy for research purposes. Research Methods in Applied Linguistics, 2(100083). https://doi.org/10.1016/j.rmal.2023.100083

[2] Ahmad Azmi AbdelHamid Esmaeil, Dg Norizah Ag Kiflee Dzulkifli, Ismail Maakip, Ovelyn Ozy Matanluk, & Sanen Marshall. (2023). Understanding Student Perception Regarding The Use of ChatGPT in Their Argumentative Writing: A Qualitative Inquiry. Jurnal Komunikasi: Malaysian Journal of Communication, 39(4), 150 - 165. https://doi.org/10.17576/JKMJC-2023-3904-08

[3] Araz, Z. (2023). Exploring the impact of language models, such as ChatGPT, on student learning and assessment. Review of Education, 11, e3433. https://doi.org/10.1002/rev3.3433

[4] Carlson, M., Pack, A., & Escalante, J. (2023). Utilizing OpenAI's GPT-4 for written feedback. TESOL Journal, 00, e759. https://doi.org/10.1002/tesj.759

[5] Da, Y. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. Education and Information Technologies, 28, 13943 - 13967. https://doi.org/10.1007/s10639-023-11742-4

[6] Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024) Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education*, 49(5), 847-864.

[7] Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. - P. T. (2023) Ethical Principles for Artificial Intelligence in Education: A Delphi Study. Sustainability, 15(10), 8171.

[8] Shen, Y., & Cui, W. (2024). Perceived support and AI literacy: the mediating role of psychological needs satisfaction. Frontiers in Psychology, 2024, 15, 1415248.

[9] Zhai, N., & Ma, X. (2022). Automated writing evaluation (AWE) feedback: A systematic investigation of college students' acceptance. Computer Assisted Language Learning, 35(9), 2817-2842.

[10] 顾成华，王丽.(2012).基于句酷批改网的大学英语写作教学实证研究.扬州大学学报(高教研究版)(04),92-96.doi:10.19411/j.cnki.1007-8606.2012.04.021.

[11] 黄静，何华清. (2018). 人机反馈对学生写作行为的影响研究. 外语电化教学, (1), 19 - 24.

[12] 金靓，杨劲松. (2022). 人工智能赋能教育背景下高职课程结构嬗变探究 —— 以外语类课程为例. 职教论坛, (5), 65 - 70.

[13] 李书影. (2018). 智能评价环境下学习者自主写作行为与成效研究. 外语电化教学, (3), 47 - 52.

[14] 刘应亮，刘胜蓝，杨进才. (2022). 社会文化活动理论视域下人机协同教学及应用探索 —— 以 iWrite 协同英语写作教学为例. 中国电化教育, (11), 108 - 116.

[15] 张亚，姜占好. (2022). 人机混合反馈环境对学习投入和二语写作水平的影响研究. 外语界, (4), 40 - 48.