



GPT API-based Chatbots as Adaptive and Facilitative Tutors for L2 English Process Writing

Min-Chang Sung (Gyeongin National University of Education) · Sooyeon Kang (Sinyeon Middle School)



This is an open-access article distributed under the terms of the Creative Commons License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received: February 17, 2025

Revised: March 8, 2025

Accepted: March 17, 2025

Min-Chang Sung (First author)
Associate Professor,
Department of English Education,
Gyeongin National University of Education
Sammak-ro 155, Manan-gu,
Anyang-si Gyeonggi-do, 13910,
Republic of Korea
Tel: +82-31-470-6335
Email: mcsung@ginue.ac.kr

Sooyeon Kang
(Corresponding author)
English Teacher,
Sinyeon Middle School
Moraena-ro 408, Seodaemun-gu,
Seoul, 03717, Republic of Korea
Tel: +87-2-379-4317
Email: englishsk329@naver.com

ABSTRACT

Sung, Min-Chang and Sooyeon Kang. 2025. GPT API-based chatbots as adaptive and facilitative tutors for L2 English process writing. *Korean Journal of English Language and Linguistics* 25, 448-471.

This paper explores the development and application of GPT API-based chatbots as adaptive and facilitative tutors in the context of the EFL process writing by secondary school students. In Study 1, we describe the development of four AI-powered chatbots, each customized for one of the four sequential stages of the L2 writing process, i.e., pre-writing, drafting, revising, and editing. Through prompt fine-tuning using the GPT-4-turbo API and OpenAI Playground, the chatbots were guided to provide stage-specific feedback, prevent unquestioning reliance on AI, and promote L2 writing development. In Study 2, the chatbots were evaluated in a case study with a cohort of six ninth-graders engaged in process writing tasks with the four stage-specific chatbots. We found that the tailor-made chatbots effectively assisted the EFL learners at each stage of the writing process, fostering constructive and creative human-AI collaboration. The learners evaluated the chatbots as helpful and knowledgeable, acknowledging their positive impact on writing quality. Based on these findings, we discuss implications for the design and use of AI chatbots in L2 writing.

KEYWORDS

process writing, GPT-4 turbo API, prompt fine-tuning, writing tutor chatbots, middle school students

1. Introduction

The process-oriented writing approach posits that understanding learners' cognitive processes is central to effective writing (Williams and Burden 1997). This approach values the iterative nature of writing, rather than focusing solely on the final product (Hyland 2003). Studies have demonstrated that process-oriented writing instruction enhances learners' motivation and reduces anxiety toward L2 writing (Hong 2014). It also improves fluency and grammatical accuracy (Butterfield and Jung 2013, Park 2017) and promotes learner autonomy, aiding students in refining writing sub-goals and improving overall writing performance (Nicolás-Conesa et al. 2014). Moreover, process-oriented approaches cultivate a sense of audience, appreciation for feedback, and increased confidence as second language (L2) writers (Kim 2016).

In a process writing class, the teacher acts as an adaptive facilitator who plays multiple roles in contextually relevant ways, such as encouraging students to recognize the fundamental processes of writing, guiding them with timely instruction and feedback through each stage, and supporting the development of individual writing strategies (Hyland 2003, Yang 2018). These roles place substantial cognitive and emotional demands on teachers, often making it challenging to manage the complexities of individual student needs (Lipson et al. 2000).

Therefore, the integration of generative AI in L2 process writing is expected to present transformative opportunities. The dynamic nature of process writing, which involves distinct stages such as brainstorming, organizing, drafting, revising, and editing, demands a variety of specialized skills from L2 writers (De Larios et al. 2002). Generative AI technologies, with their advanced ability to understand task contexts and user intents, can adapt to these different stages and offer stage-relevant assistance (Zou and Huang 2023). Such adaptive AIs may enhance the efficiency of each writing stage and help L2 writers navigate the complexities of process writing while developing their skills.

However, the implementation of AI in L2 writing is fraught with significant challenges. A key concern is learners' tendency to over-rely on AI, expecting it to handle all writing tasks on their behalf (Ningrum 2023). This over-reliance can impede their writing development, reducing them to passive participants in the writing process. Furthermore, ethical issues such as plagiarism and academic dishonesty arise when learners use AI-generated text as their own (Yuan et al. 2024). These challenges necessitate a careful examination of how AI is integrated into the learning environment to ensure it functions as a supportive tool that fosters genuine skill development.

To harness the opportunities of generative AI while strategically addressing its challenges, it is essential to consider technological advancement and instructional design in tandem and create novel environments for enhanced language learning. As one such approach, this paper explores recent innovations in ChatGPT, such as GPT-4-turbo API and OpenAI Playground, across two studies. In Study 1, we use prompt fine-tuning techniques to develop GPT API-based chatbots as adaptive and facilitative L2 writing tutors that offer stage-specific feedback, prevent unquestioning reliance on AI, and promote L2 writing development. In Study 2, we conduct a case study to analyze how L2 learners interact with these chatbots and incorporate feedback into their process writing. We also examine how these chatbots foster human-AI collaboration and promote L2 learners' technological adaptivity. The research questions guiding each study are as follows:

Study 1

1. How can we develop GPT API-based AI chatbots as adaptive and facilitative tutors for L2 English process writing?
2. How can we design the architecture and UI of the chatbots for L2 English learners?

Study 2

1. How do L2 learners interact with the AI writing tutor chatbots during process writing?
2. How does the use of the AI writing tutor chatbots affect L2 learners' writing quality and their perceptions of the chatbots?

2. Use of Generative AI for L2 Writing

The use of generative AI in learning and teaching has demonstrated promising potential for providing pedagogical integration and instructional support (Miao and Holmes 2023, Ruiz-Rojas et al. 2023). This burgeoning trend is particularly apparent in recent advancements in L2 learning and teaching (Law 2024). For example, Yuen and Schlote (2024) explore how AI features within mobile apps enhance language learning by offering personalized and adaptive feedback. Similarly, Kim et al. (2023) emphasize that generative AIs such as ChatGPT have the potential to facilitate supportive learning environments and promote enhanced learning experiences when given appropriate prompts, which points to the importance of prompt fine-tuning. Generative AIs have been found to aid L2 learners and teachers in various pedagogical tasks such as lesson planning, material development, assessment, and personalized feedback (Koraishi 2023, Octavio et al. 2024). These advantages of using generative AIs in diverse contexts of L2 learning can be fully realized only when L2 educators are prepared to integrate AI into their teaching, which underscores the necessity for professional development (Kohnke et al. 2023).

Among a variety of L2 learning contexts, generative AIs have been extensively investigated for L2 writing, probably because of their impressive performance in natural language understanding and generation, i.e., NLU and NLG (Shi and Aryadoust 2024). Many studies have highlighted that generative AIs significantly enhance L2 writers' skills and motivation. For example, Song and Song (2023) assessed the efficacy of ChatGPT in improving academic writing skills and motivation among Chinese college learners of L2 English. The researchers attributed this result to ChatGPT's ability to provide immediate feedback, facilitate idea generation, and enhance overall engagement in the writing process. The implementation of ChatGPT has been also found to have positive effects on L2 writing teachers in various aspects such as time management and self-efficacy (Ghafouri et al. 2024, Steiss et al. 2024).

Noting the efficacy of generative AIs, mostly ChatGPT, in L2 writing pedagogy, a variety of L2 writing activities have been examined such as evaluating ChatGPT's sample summaries of the target texts for revision (Strobl et al. 2024) and utilizing ChatGPT as "a self-learning tool for writing and thinking development" at various stages in process writing (Zou and Huang 2023, p. 1). Such writing activities facilitate the development of L2 writing skills and demonstrate the potential of generative AIs to provide collaborative and interactive learning experiences for L2 writers. For example, Strobl et al. (2024, p. 1) introduced the concept of using ChatGPT as a collaborative agent, namely "a writing buddy" in advanced L2 classes, as both the advanced L2 writers and ChatGPT completed the same summary tasks in the classes.

However, the effectiveness of generative AI's feedback compared to human feedback remains a controversial issue. It has been contended that AI-based feedback significantly improves students' writing outcomes, engagement, and feedback literacy. For instance, Rad et al. (2023) showed that the AI application *Wordtune* improved L2 learners' writing outcomes and engagement compared to the control group, highlighting its potential to enhance learning through valuable feedback. Similarly, Shi and Aryadoust (2024) conducted a systematic review of AI-based automated written feedback and concluded that AI-based automated feedback complements human

feedback by providing consistent and detailed support for high-quality writing. Their review showed that AI systems often focused on form-based feedback while also offering meaning-focused feedback, effectively complementing human feedback strategies.

On the other hand, Steiss et al. (2024) provided a more nuanced view by comparing the quality of feedback from ChatGPT and human teachers. Their findings showed that well-trained evaluators provided higher quality feedback than ChatGPT in several key areas such as clear direction, accuracy, and supportive tone. This underscores that using AI-generated feedback instead of human feedback involves a tradeoff between quality and efficiency. Although ChatGPT offers adaptive and useful feedback for L2 writers in a timely manner, it does not yet match the depth and nuance of human feedback in critical areas. Similarly, Zhu et al. (2024) explored AI-generated feedback on English writing tasks and found that AI could provide valuable insights, although human moderation was necessary to ensure the quality and appropriateness of the feedback. Therefore, it is essential that the integration of AI feedback be accompanied by human oversight or intervention to maximize its effectiveness in improving L2 writing skills.

Another controversial aspect of using generative AI in L2 writing is the ethical and practical challenges. One criticism is that the over-reliance on AI tools may hinder L2 students' language development and critical thinking skills (Crossley 2018, Kasneci et al. 2023, Ningrum 2023). There has also been considerable discussion on ethical issues related to plagiarism and academic integrity, as well as practical challenges such as affordability and accessibility of AI tools (e.g., Barrot 2023, Strobl et al. 2024). Considering AI's feedback has developed to the point where it is difficult to distinguish between AI-generated feedback and human feedback, the issues of originality and academic ethics are becoming increasingly complex (Casal and Kessler 2023). Yan (2023) and Yuan et al. (2024) note concerns about academic honesty and equity, urging the development of regulatory policies and pedagogical guidance.

In sum, generative AIs present a promising opportunity to enhance L2 writers' skills and motivation by fostering human-AI interaction and collaboration. However, the usefulness of AI feedback remains controversial when compared to human feedback. Generative AIs also face ethical problems and practical challenges, including the potential hindrance of creative thinking and issues related to academic integrity. To fully tap into the opportunities while addressing these challenges, innovation in both AI technology and L2 writing pedagogy is essential. In this regard, the present study utilizes recent advancements in AI, such as the GPT-4-turbo API and OpenAI Playground, to develop adaptive and facilitative AI chatbots supporting the English process writing of Korean middle school students.

3. Study 1: L2 Writing Tutor Chatbot Development

3.1 Technological Resources

The present study utilized a set of technological resources to design and operate AI chatbots as adaptive and facilitative writing tutors. The first resource is OpenAI's ChatGPT API 4-turbo, a powerful language model designed for generating human-like text based on given prompts. It offers improved efficiency and response accuracy compared to its predecessors, such as GPT API 3.5 or 3.5-turbo, making it suitable for real-time applications. For example, the chatbot for the drafting stage should not provide a complete English sentence or paragraph to prevent L2 writer from over-relying on AI, and, unlike the lower versions of GPT API, the ChatGPT API 4-turbo successfully adheres to this key requirement.

The second technological resource is the OpenAI Playground. It is an interactive platform that allows API users to experiment with various OpenAI models. Among the models, the present study used the *Chat* model which enables API users to design chatbots, providing a user-friendly interface for testing fine-tuned prompts in the chatbot code and observing chatbot responses in real time.

In addition, we used ChatGPT to write the HTML code for the foundational architecture of the chatbots. ChatGPT provided various types of functional code such as framing the chat window, handling user inputs and chatbot responses, and managing conversation history. We also utilized Google Sites, which enables the creation of an accessible and organized platform for L2 writers to interact with the writing tutor chatbots.

3.2 Developmental Procedures

The development of AI writing tutor chatbots involved three main procedures: a) specifying the roles and constraints of the chatbots, b) fine-tuning the chatbot development prompts, and c) creating the web-based user interface of the chatbots.

3.2.1 Specifying roles and constraints

Drawing from earlier research on L2 process writing and the use of AI chatbots in pedagogy (Bahja et al. 2020, De Larios et al. 2002, Guo and Li 2024, Kang and Sung 2024, Kim 2016, Kim et al. 2022), we identified the major challenges that L2 writers experience in each stage of process writing, as well as their potential misuses of AI chatbots. Based on these findings, we defined the types and degrees of assistance that chatbots should offer at different stages of L2 process writing. That is, the chatbots should adapt their assistant roles to different stages of process writing and facilitate L2 writers' agency by scaffolding feedback. For example, the editing chatbot should provide an appropriate amount of editorial support such as providing feedback on grammar and punctuation while it must not rewrite an entire essay on behalf of learners.

3.2.2 Prompt fine-tuning

The chatbot development prompt refers to detailed information that is embedded in a chatbot's code to define its characteristics. This includes specifications such as its role, working context, and response constraints. The development prompt should be distinguished from the user prompt, which is the input provided by users to interact with the chatbot. While the development prompt establishes a chatbot's foundational behavior, the user prompt dynamically influences how the chatbot generates responses in real time.

The development prompt of each chatbot was fine-tuned in an iterative manner until both researchers agreed that the chatbot delivers non-excessive, contextually relevant, and pedagogically valuable feedback to L2 writers.

First, we documented the development prompt of a chatbot based on the previously identified roles and constraints. The prompt was written in Korean since the learners were expected to talk to the chatbot in their L1. Second, we built the chatbot using the developmental prompt at the OpenAI Playground and examined the chatbot responses. Third, when the chatbot did not demonstrate satisfactory performance in fulfilling its intended objectives, the development prompt was revised. When the chatbot showed inconsistency in addressing certain types of request, model conversation pairs of learner request and chatbot response were added to the prompt. These prompt fine-tuning processes were repeated for each of the four AI writing tutor chatbots.

3.2.3 Creating the user interface

We created and merged two types of code for chatbot installation: a) the API code that defines the chatbots' functions was extracted from the final development prompts documented on the OpenAI Playground, and b) the architecture code that designs physical aspects of the chatbots (e.g., height, color) was written by ChatGPT.

The user interface of the chatbots was realized on a website with a homepage and four separate chatbot pages, where the merged code was installed. For example, the API code of the drafting chatbot and the architecture code were merged and installed onto the drafting chatbot page.

3.3 Outcome

3.3.1 Prompt fine-tuning

Four writing tutor chatbots were developed for the four stages of process writing, i.e., pre-writing, drafting, revising, and editing, based on the design principles of L2 learning chatbots (Huang et al. 2021). All the chatbots were assigned the role of an ESL writing tutor helping Korean middle school students with their English writing, but each chatbot, as an adaptive and facilitative tutor, was subject to stage-specific roles and constraints. For this purpose, the development prompts of the four chatbots were fine-tuned, as in Table 1, and included in the API code.

Table 1. Fine-tuning of Chatbot Development Prompts

Chatbot	Roles and Constraints	Development Prompts
Pre-writing	• To focus on the pre-writing stage • To support key tasks during the pre-writing stage • To assist L2 writers in collecting ideas and independent thinking	• Assist students in the pre-writing stage. • Your goal is to increase their motivation, help them gather ideas and decide the purpose and audience of their writing. • Offer a suitable amount of information to encourage independent thinking.
Drafting	• To focus on idea expression and outlining • To prevent L2 writers' over-reliance on the chatbot in English writing • To ensure the importance of student independence	• Help students express their ideas in English and outline their thoughts. • Respond only in Korean and provide just one English word when necessary. • If you are asked to write an English sentence, refuse and provide a key word.
Revising	• To focus on relevant aspects of the revising stage when giving feedback • To deprioritize grammar and spelling, which will be covered in the next stage, i.e., editing	• Read the student's text and give specific feedback on content, organization, and style. • Avoid giving feedback on grammar and spelling whenever possible.
Editing	• To focus on relevant aspects in the editing stage when giving feedback • To promote independent error correction by the students	• Help students correct errors in grammar, spelling, and punctuation in their written work. • Do not rewrite the text for them; instead, give indirect feedback to help students correct the errors themselves.

As in Table 1, the roles, constraints, and prompts for the writing tutor chatbots are tailored to meet distinct objectives that align with the respective L2 process writing stage they support. For instance, in the pre-writing stage, the chatbot is designed to enhance students' motivation, assist in idea generation, and help determine the

purpose and audience of their writing. Similarly, during the drafting stage, the chatbot aids in expressing ideas and creating outlines while maintaining the emphasis on student independence by refusing to write entire English sentences or paragraphs. In the revising stage, the chatbot focuses on content, organization, and style, offering detailed feedback while minimizing attention to grammar and spelling errors. Lastly, in the editing stage, the chatbot is asked to provide brief feedback to encourage students to identify and correct their own mistakes in grammar, spelling, and punctuation. This structured approach ensures that L2 writers receive targeted support at each stage of process writing.

3.3.2 Code and UI design

The prompts were entered into the chatbot development platform on OpenAI Playground, where we also adjusted the chatbot settings such as the API model (GPT-4-turbo: the most advanced model at the time of development), temperature (0.2: consistent responses for factual relevance), max_tokens (varying from 250 to 600 according to writing stages), and frequency_penalty (0.3: avoiding repetition of the same response for learners' face validity). On OpenAI Playground, the prompts and the chatbot settings could be converted into any of four code types, i.e., Node.js, Python, Curl, and JSON, and we chose the Node.js code considering that the code would be integrated with the HTML-based architecture code for the writing tutor chatbot's physical architecture on a webpage which many students would access simultaneously. Node.js is known for its compatibility with server-side scripting and its efficiency in handling asynchronous I/O operations, which is essential for real-time applications like pedagogical chatbots. Its ability to manage multiple concurrent connections enhances user experience by processing requests swiftly.

After generating the four versions of Node.js API code for the four chatbots and one HTML architecture code for the user interface, we combined each API code and the architecture code into a single code to be entered into a webpage made by Google Sites. The finalized code consists of three major functions, each serving a distinct educational purpose in the development and operation of the chatbots.

The first function pertains to the chatbot design-related commands, which establish the structure and styling of the user interface. These commands include setting the encoding, viewport configuration, and specifying visual styles such as the chat container, messages, and user input fields. These design elements ensure that the chatbot is visually appealing, responsive, and user-friendly, providing a seamless interaction experience for the users.

The second function involves processing the conversation history between chatbots and learners. This function initializes an array to store the conversation history and includes a method to add new messages to the chat. We ensured that a maximum of seven messages is retained in the chatbot history to manage the learner-AI interaction in a rich context.

The third function focuses on handling AI responses and event handling. It enables an asynchronous multiple channel to send user prompts and receive AI-generated replies. It also adds both user and AI messages to the chat and clears the input field after submission. These functionalities collectively ensure that the chatbot interacts dynamically with users, processing inputs and providing intelligent responses in real-time.

We have created a convenient user interface based on the design factors of educational chatbots such as mobility, interactivity, content, performance, and reliability (Bahja et al. 2020). The four chatbots are installed onto four different web pages under a homepage, as in Figure 1, so that L2 writers can easily find the four chatbots using the navigation bar at the top. Each chatbot webpage has a title showing the relevant stage of process writing (e.g., 2. DRAFTING), under which the chatbot is embedded.

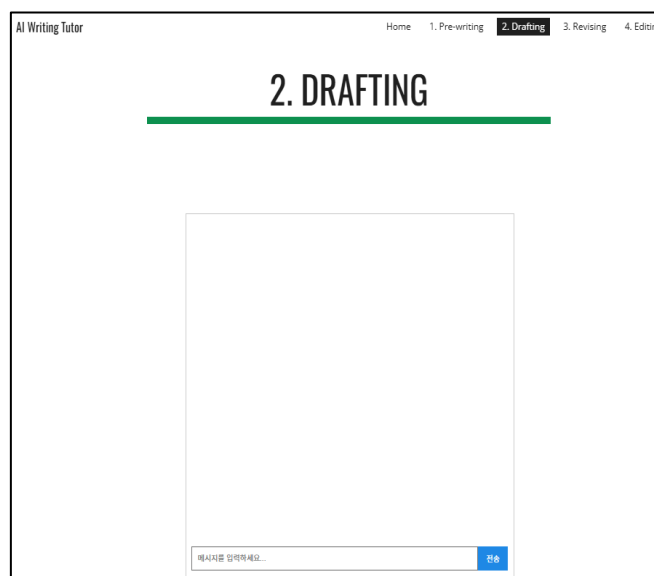


Figure 1. Web User-Interface of Writing Tutor Chatbots

4. Study 2: L2 Process Writing with AI Chatbots

4.1 Participants

A total of six 9th-grade students at B Middle School located in Seoul participated in the study during the first semester of 2024. Five of them were female, and one was male. All the participants had undergone six years of formal English education in Korea, with no prior experience living in an English-speaking country. Based on the results of an online English test provided by Cambridge English (<https://www.cambridgeenglish.org/test-your-english>), their proficiency levels were classified according to the Common European Framework of Reference for Languages (CEFR), with two students at the B1 level and four at the A2 level.

All participants had no prior experience with AI-assisted process writing and uniformly reported difficulties in English writing, particularly with the drafting stage. Therefore, they were all interested in improving their English writing skills using AI technology. They were recruited by one of the researchers of this study, who was working as an English teacher in the middle school. The study was announced to all the 9th graders at the school, and only those who showed interest voluntarily participated. They were informed of the study's purpose, procedures, benefits and risks, and consent was obtained from all participants. Privacy protection measures included the anonymization of personal data, with data management practices following confidentiality guidelines.

4.2 Materials

We developed process writing tasks for Korean middle school students considering contextual relevancy and genre diversity (Hyland 2007). Contextual relevancy was ensured by extracting frequent writing topics from their textbook units, and genre diversity was addressed by incorporating two different genres, i.e., argumentative and expository essays. The argumentative essay focused on the pros and cons of using disposable products, while the expository essay illustrated animals' unique abilities. Both process writing tasks were structured into four stages,

i.e., pre-writing, drafting, revising, and editing.

In order to evaluate the learners' writing process, we developed an evaluation rubric for the four distinct stages of process writing in reference to Ene and Kosobucki (2016) and Liu (2024). Each stage was evaluated based on two assessment criteria, resulting in a total of eight criteria. For the pre-writing stage, the criteria were the relevance and clarity of ideas to the topic (1.1) and the logic and structure of the organization (1.2). In the drafting stage, the criteria included the extent to which the brainstorming results from the pre-writing stage were reflected in the draft (2.1) and the comprehensibility of the draft (2.2). For the revising stage, the criteria were the degree of content revision (3.1) and the degree of organizational and structural revision (3.2). Finally, in the editing stage, the criteria were the completeness and readability of paragraphs (4.1) as well as the level of editing, particularly regarding errors in grammar, spelling, and punctuation (4.2).

We also developed an evaluation survey to assess how the learners perceived the effectiveness of the AI writing tutor chatbots. The majority of the questions were structured using a 5-point Likert scale, in which participants could express their agreement with various statements, ranging from "Strongly disagree (1 point)" to "Strongly agree (5 points)". Referring to Kang and Sung (2024), the Likert scale questions about the effectiveness of the chatbots were organized into five themes: entertainment, convenience, writing effectiveness, interest in writing, and future use. These themes were applied to both the overall chatbot experience and each of the four individual chatbots, resulting in a total of 25 questions. To gain deeper insights, four open-ended questions were also included, asking participants to describe the most helpful feedback, the most interesting aspect of the chatbots, and any difficulties in using the chatbots.

4.3 Procedures

4.3.1 Process writing tasks and chatbot evaluation

Process writing tasks were conducted on two days over a week in a quiet classroom after school. Each meeting took approximately 55 minutes and was scheduled in advance with the participants. During the first meeting, the participants went through the pre-writing and drafting stages, and during the second meeting, they completed the revising and editing stages. Among the six participants, three were provided with an expository writing topic (i.e., Introduce an animal with special talents.), whereas the other three were provided with an argumentative writing topic (i.e., Do you agree with the use of disposable plastics?).

Once seated, each participant was provided with worksheets containing clear directions for their tasks for each stage and a tablet PC with a keyboard to access the AI writing tutor chatbots. The worksheet provided the topic, explanations of each stage, and directions for tasks in each stage (for further details, refer to the Appendix). The teacher highlighted that the participants should complete the tasks with assistance from the corresponding AI chatbots. To help them engage effectively, a brief pre-task training session was provided on how to initiate questions and sustain conversations with the chatbots. Additionally, the participants were encouraged to think critically throughout their interactions with the chatbots.

After any questions regarding the procedures and directions were resolved, they began to chat with the AI writing tutor chatbots and complete the tasks on the worksheet in a sequential manner. They were allowed to chat either in Korean or English, but they all ended up conversing with the chatbot in Korean. At the end of every stage, the teacher collected the conversation logs along with their task outputs for event-focused discourse analysis (Wortham and Reyes 2020).

For the first two stages (i.e., pre-writing and drafting) on Day 1, they used the worksheet and the AI writing

tutor chatbots. When the drafting stage was completed, every participant had the first draft of their writing on the worksheet. For the revising and editing stages on Day 2, they started to use Google Docs for typing, revising, and editing their writings based on the feedback from and interactions with the chatbots.

At the end of the second meeting, the learners assessed the AI writing tutor with the evaluation survey, which took about 10 minutes.

4.3.2 Writing process evaluation

A total of four raters evaluated learners' writing process based on the evaluation rubric with the eight criteria (see Section 4.2). Two of the raters were the authors of this study while another was a professor specializing in English education, and the other was an English teacher with a master's degree in English education, possessing over a decade of teaching experience. Through email, the learners' writings, the worksheets they used, and the evaluation rubric were provided to the raters. Each rater was instructed to independently score the participants' products over the four writing stages from 0 (=Inadequate) to 4 (=Excellent) and provide justifications for their scores. The score 4 means the learner exceeds expectations with outstanding quality, while the score 0 means the learner fails to meet basic requirements.

After collecting the first-round scores and comments, a statistical adjustment using Z-scores was performed to ensure fairness and comparability in scoring. This procedure was essential as the absence of a rigorous pre-training phase, despite the provision of a detailed evaluation rubric, may have increased variability in scoring patterns due to raters' individual perspectives and prior experiences. To address this, we calculated Z-scores based on each rater's scores. When the inter-rater discrepancies for a particular student's scores on a single criterion differed significantly, by more than one standard deviation, from the average Z-score, the raters were requested to review and modify their scores and justifications at their discretion. The final dataset was established based on the consensus reached.

4.4 Data Analysis

To offer a deeper insight into learner-chatbot interaction, a case-study approach was chosen. A close reflection on the particularities of a case allows access to "educational reality" (Van Lier 2005, p. 195). Considering there has been a lack of attention to what actually happens during human and AI interaction among learners in educational context, analysis of the particularization will "inform pedagogical development in the use of AI and the design of AI for education" (Wang et al. 2023, p. 3). The analysis of learner-chatbot interactions aimed to identify both commonalities and particularities in learners' use of the chatbot at each stage of the process writing approach, as well as the chatbot's role. To achieve this, the two researchers conducted a thorough, iterative review of the chatbot-student logs, systematically extracting patterns and distinctions.

To examine the impact of AI writing tutor chatbots on L2 learners' writing quality, the mean rating scores across eight criteria (i.e., 1.1-4.2) were calculated for the six learners' process writing performance. Referring to the scoring rationale provided by the raters, the two researchers conducted an in-depth review of the students' written texts to reveal the distinctive characteristics of each case. We also examined the evaluation survey results to understand how the learners reacted to chatbot use at different stages of writing. For this, the mean and standard deviation of their responses on a five-point Likert scale were calculated, and the responses to open-ended questions were scrutinized to identify case-specific patterns.

4.5 Findings

4.5.1 Interaction between L2 learners and chatbots

The interaction between the chatbot and learners at the pre-writing stage is characterized by learners starting with general questions and gradually narrowing down to more specific ones. In response, the chatbot provides relevant information while also encouraging learners to engage more deeply by prompting them to clarify their opinions, refine their questions, and incorporate relevant details into their writing.

Excerpt 1 illustrates a typical interaction in expository writing, where participants first asked broad questions about animals with special talents, selected one animal based on the chatbot's suggestions, and then requested more detailed information. As the learners and chatbots interacted in Korean, the excerpts are presented in Korean-English bilingual notation, where the English translation is provided in italic with major patterns highlighted in bold.

Excerpt 1. Interaction at Pre-Writing [Jia*, expository]

Jia: 특별한 재능을 가진 동물들을 알려줘.

Tell me about animals with special talents.

AI: 동물들 중에는 각기 다른 특별한 재능을 가진 종들이 많아요. 예를 들어, ...

여러분이 생각하는 동물 중 어떤 동물의 재능에 대해 더 알아보고 싶나요?

Many animals have unique abilities! For example, ...

Is there a specific animal whose talent you'd like to learn more about?

Jia: 개미의 특별한 재능을 알려줘.

Tell me about the special talents of ants.

AI: 개미는 놀라운 조직력과 팀워크로 유명해요. ... 어떤 주제로 글을 쓰고 싶은지 생각해 보세요.

*Ants are known for their incredible teamwork and organization. ... **Think about what topic you want to write about.***

Jia: 개미의 특별한 재능을 10줄로 말해줘.

Tell me about the special talents of ants in 10 lines.

AI: ... 개미는 또한 효율적인 음식 수집 시스템을 가지고 있어서, 발견한 음식을 다른 개미에게 신호를

보내 공유합니다. 이러한 재능에 대해 글을 쓸 때, 어떤 점이 가장 인상 깊었나요?

*... Ants also have an efficient food collection system, using signals to share discoveries with other ants. **When writing about their abilities, which aspect do you find most interesting?***

Jia: 개미가 개미 친구들에게 신호를 보내는 방법을 알려줘.

How do ants send signals to their friends?

Similarly, the participants composing argumentative essays on disposable plastics initially explored both advantages and disadvantages before deciding on a stance and seeking supporting evidence. This is well illustrated in Ian's case in Excerpt 2.

*All participants' names are pseudonyms.

Excerpt 2. Interaction at Pre-Writing [Ian, argumentative]

Ian: 여러가지 의견들이 있을텐데 먼저 일회용 플라스틱 사용에 대해서 찬성하는 입장에 대해 알려줘.

There may be various opinions, but first, tell me about the arguments for using disposable plastics.

AI: 일회용 플라스틱 사용을 찬성하는 입장에서는...

Those in favor of using disposable plastics ...

Ian: 그럼 이제 플라스틱 사용에 대해 금지하는 입장을 말해줘.

Then, please state the position on banning plastic use.

AI: 플라스틱 사용을 금지하는 입장에서는 ... 이러한 문제점들을 어떻게 해결할 수 있을까요?

The position advocating for banning plastic ... How can we address these issues?

Ian: 근데 이런식으로 금지를 하면 우리 생활에 편리함과 경제적인 이점에서 손해를 보지 않을까?

But if we ban it like this, won't we lose convenience and economic benefits in our daily lives?

AI: 플라스틱 금지로 인한 단점을 고려할 때, 대체재 개발과 적응 과정이 중요합니다. ...

이러한 변화에 대해 어떻게 생각하나요?

When considering the drawbacks of banning plastic, the development of alternatives and the adaptation process are crucial. ... What do you think about these changes?

Ian: 그럼 너의 의견은 혹시 금지하는 쪽이니 아니면 찬성하는 쪽이니?

So, are you in favor of banning or not?

AI: 저는 여러분이 글을 쓰기 전에 다양한 관점을 고려해볼 것을 권장합니다. ...

어떤 점이 가장 중요하다고 생각하나요?

I recommend considering various perspectives before forming an opinion. ...

What do you think is the most important aspect?

Ian: 난 그래도 환경을 먼저 생각해야 된다고 생각해.

I still think we should prioritize the environment.

At the drafting stage, the interaction between the chatbot and learners is characterized by participants inquiring about the translation of Korean words or phrases into English and occasionally seeking additional information to complement their writing, with the chatbot supporting the drafting process by providing relevant information based on their requests.

However, it seems that the particular expectations of what the drafting chatbot would provide varied among learners. Ian and Lael engaged with the chatbot as a resource for evaluating their own English sentences in terms of grammatical accuracy and semantic appropriateness, whereas Jia and Mina approached it as a personal assistant capable of offering support across a wide range of inquiries. For example, as in Excerpt 3, Lael focused on syntactic-semantic accuracy of her English compositions, requesting the chatbot to translate her own English sentences into Korean and check for any errors.

Excerpt 3. Interaction at Drafting [Lael, argumentative]

Lael: You can use plastic when you go camping or do some outdoor activities를 번역해줘

"You can use plastic when you go camping or do some outdoor activities." Translate this.

AI: "캠핑이나 야외 활동을 할 때 플라스틱을 사용할 수 있습니다."

"You can use plastic when camping or doing outdoor activities."

Lael: 문장에서 어색한 부분은 없어?

Are there any awkward parts in the sentence?

AI: 네, 문장이 자연스럽고 명확합니다.

No, the sentence is natural and clear.

On the other hand, Jia attempted various types of request such as translating her English sentences, checking grammatical accuracy, and asking for words that matched her proficiency level, as in Excerpt 4. This case exemplifies the use of the chatbot as an adaptive personal assistant.

Excerpt 4. Interaction at Drafting [Jia, expository]

Jia: ants have speacl talent 한국어로 번역해줘 (author note: “speacl” is the learner’s typo.)

Translate “ants have speacl talent” into Korean.

AI: “ants have special talent”을 한국어로 번역하면 ...

If you translate “ants have special talent” into Korean, ...

Jia: 협동 영어로

How do you say ‘hyeopdong’ in English?

AI: “cooperation”이라는 단어를 사용할 수 있습니다.

You can use the word “cooperation.”

Jia: 쉬운 단어로

In simpler words?

AI: “teamwork”라는 단어를 사용할 수 있습니다.

You can use the word “teamwork.”

Jia: what teamwork lets know가 맞는 문장인지 알려줘

Is “what teamwork lets know” a correct sentence?

Meanwhile, Yena appeared to perceive the drafting chatbot primarily as a dictionary, while Sora, despite acknowledging the restriction on translation, still attempted to utilize it as a translation tool. For example, in Excerpt 5, Sora attempted to translate entire sentences twice, suggesting a degree of dependency and an inclination to use the chatbot as a translation tool.

Excerpt 5. Interaction at Drafting [Sora, argumentative]

Sora: 친환경 대체제 보다 훨씬 효율적이다

Much more efficient than eco-friendly alternatives

AI: 이 문장을 영어로 표현하고 싶으시다면, 중요한 단어를 알려드릴게요: “efficient”.

이 단어를 사용해서 문장을 구성해보세요. ...

If you want to express this sentence in English, I’ll give you an important word: efficient.

Try to construct a sentence using this word. ...

Sora: 성급하게 플라스틱 사용을 금지하면 많은 문제를 야기할 수 있다.

Banning plastic use too hastily could cause many problems.

AI: “Prematurely”

In the revising stage, learners sought feedback on content and organization, and the chatbot offered positive feedback, suggestions and examples. In response, most students asked follow-up questions to gather more specific information and find appropriate English expressions for revision. Excerpt 6 depicts typical interactions among the learners while using the revising chatbot.

Excerpt 6. Interaction at Revising [Mina, expository]

Mina: (초안 제시) 글의 구성과 내용에 대한 피드백을 해줘

(Draft Presentation) Please provide feedback on my writing in terms of content and organization.

AI: 글의 주제인 개미의 팀워크에 대해 잘 설명하고 있습니다. 그러나 몇가지 개선할 점이 있습니다. ... 각 역할에 대한 설명을 조금 더 확장하여 개미들의 팀워크가 어떻게 효과적으로 가능한지 구체적인 예시를 추가하는 것이 좋겠습니다. ... 마지막 문장에서 “Now you can feel that ants are smart and excellent.” 대신, 개미들의 특별한 사회구조와 팀워크가 왜 놀라운지를 요약하여 강조하는 결론을 제시하는 것이 좋겠습니다..

The essay explains the topic of ant teamwork well. However, there are a few areas for improvement. ... Expanding the descriptions of each role and adding specific examples of how the teamwork among ants functions effectively would be beneficial. ... In the final sentence, instead of saying “Now you can feel that ants are smart and excellent,” it would be better to present a concluding statement that summarizes and emphasizes why the ants' unique social structure and teamwork are remarkable. ...

Mina: 일개미가 어떻게 협력해 음식을 찾고, 집을 짓는지 구체적으로 설명해줘.

Explain in detail how worker ants cooperate to find food and build their anthill.

AI: 일개미들은 매우 조직적이고 효율적인 방법으로 협력하여 음식을 찾고 집을 짓습니다. ...

Worker ants cooperate in a very organized and efficient manner to find food and build a house. ...

Mina: 이동하다 영어로

“idonghada” in English

AI: “move”

The interaction between the chatbot and students at the editing stage is characterized by students submitting their revised essays and requesting corrections. In response, the chatbot's prominent role was to make suggestions regarding capitalization, punctuation, grammaticality, and naturalness. Four students ended the interaction after receiving feedback from the chatbot, while one participant (Sora) asked if there was additional feedback or if it was acceptable not to correct certain phrases she believed clearly conveyed the intended meaning. The remaining student (Yena) inquired about the meaning of a word or sentence provided by the chatbot, which shows her desire to verify whether the sentence edited by the chatbot conveyed the same meaning she intended (See Excerpt 7).

Excerpt 7. Interaction at Editing [Yena, expository]

AI: “This is the first things of dolphin’s special abilities”에서 “thing” 대신에 “aspect”를 사용하고 ...

In “This is the first thing of dolphin's special abilities,” you can replace “thing” with “aspect” ...

Yena: aspect 뜻이 뭐야?

What does “aspect” mean?

AI: “aspect”는 어떤 사물이나 상황의 특정 부분이나 측면을 의미하는 단어입니다. 예를 들어, ...

“Aspect” refers to a particular part or facet of an object or situation. For example, ...

Yena: This is the first aspect of dolphin’s special abilities 어떻게 해석해

How should I interpret “This is the first aspect of dolphin's special abilities”?

In summary, at the pre-writing stage, the learners interacted with chatbots to gather information and refine their topic ideas. The chatbot primarily provided relevant information and prompted the learners to think critically. During drafting, learners frequently sought English expressions and occasionally requested additional information, with the chatbot offering appropriate expressions and information; notably, this stage revealed the greatest number of particularities among the learners. In the revising stage, the learners asked for comprehensive feedback on content and organization by submitting their entire drafts, and the chatbot provided detailed feedback and suggestions for improvement. At the editing stage, the learners focused on correcting grammar and improving the naturalness of their writing, with the chatbot making grammatical and stylistic recommendations.

4.5.2 Effects of chatbot use on writing quality

Figure 2 shows the mean rating scores in the eight criteria (i.e., 1.1~4.2) of the six learners' performance in process writing. It was revealed that, during the pre-writing stage, all learners produced relevant and clearly organized brainstormed ideas in Korean, directly related to their writing topics, except for one (Jia) who showed a minor problem in organizing ideas. At the drafting stage, with chatbot assistance, learners generally succeeded in integrating their brainstormed ideas into their initial drafts in English, although one learner (Sora) exhibited only partial reflection. Regarding the comprehensibility of the first drafts, four learners produced clear and understandable texts with minor issues. However, two learners' (Jia and Sora) writings were somewhat difficult to follow due to some unclear sentences.

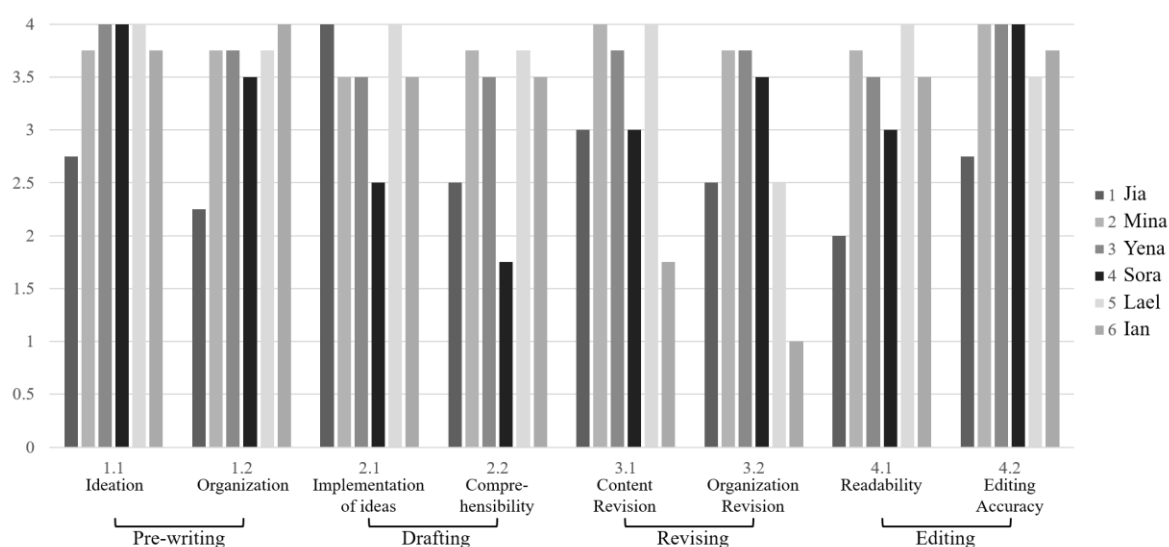


Figure 2. Results of Writing Process Evaluation

During the revising stage, significant improvements were observed in the content of the learners' writings. With the exception of one learner (Ian) whose writing content was overall well-developed from the outset, all students incorporated additional details by consulting with the chatbot, resulting in more comprehensive and persuasive texts. Excerpt 8, for example, shows how Lael revised her text based on chatbot's suggestions. She attempted to include background information explaining her disagreement with the ban on single-use plastics in the introduction and provided the economic advantages of plastic compared to other materials.

Excerpt 8. Lael's Revision Process [Argumentative]

1) Chatbot's Suggestions at Revision Stage

In the introduction, you clearly present an argument against banning the use of plastic, *but providing a bit more background information on why this issue is important would help the reader better understand the context of your writing.* (...) Second paragraph: While you mention the cost-effectiveness of plastic production, *adding specific cost data or comparative examples would make your argument more convincing.*

2) Revised Version: Added Parts Underlined

Some people say that we should stop using plastic. But, I don't agree with it. If we stop using we can face many problems. From now on, I'm going to tell you why I don't agree with stopping using plastic. (...) Second, production of plastic is simple and economical. If we assume that we will make same size of bottles, plastic bottles will be the cheapest bottles among bottles that are made of glass, iron or paper. Because raw material prices of plastic is low. It shows that plastic can be mass-produced. So It's affordable. (...)

However, in terms of organizational revision, the results varied among the learners. While half of the learners (Mina, Yena, Sora) showed moderate to substantial improvements, such as clearer paragraph organization or more explicit introductions and conclusions based on the chatbot's suggestions, the remaining three learners (Jia, Lael, Ian) showed only minor improvements in this area. Two of them (Lael, Ian) demonstrated clear organization from the outset, suggesting that these learners did not perceive a need to improve their writing's structure. The remaining learner (Jia) struggled with organizing their writing from the beginning, which may explain why she had difficulties in revising her draft. Adjusting the organization may have posed a considerable challenge for this learner.

Excerpt 9 depicts how Mina revised the organization of her writing based on the chatbot's suggestion. To achieve a clearer structure, she divided the introduction, body, and conclusion into separate paragraphs, decided to pose a question before explaining the ants' social structure, and incorporated a summary sentence in the conclusion based on feedback from the chatbot.

Excerpt 9. Mina's Revision Process [Expository]

1) Chatbot's Suggestions at Revision Stage

There is a need to clarify the structure of the writing. For example, providing an explanation of the ants' social structure immediately after the question 'So how do the ants do these things?' would aid readers in comprehending the information more easily. (...) Instead of the final sentence 'Now you can feel that ants are smart and excellent,' it would be better to provide a concluding statement that summarizes and emphasizes why the ants' unique social structure and teamwork are remarkable.

2) Revised Version: Added Parts Underlined

Ants are very cooperative animals. They find their food efficiently and build their house. So how do the ants do these things?

The answer is from their social structure. The social structure includes the queen ant, worker ants, and male ants. The queen ant handles the group's reproduction. And male ants reproduce with the queen ant. Worker ants find food, build the whole house. (...)

As you read, every ant does their own work. And this makes teamwork! Now you can feel that ants are smart and excellent.

At the editing stage, except for one learner (Jia), all the learners' writings were rated as overall complete and easy to read. Particularly, their writings showed very few or no errors in terms of grammar, spelling and punctuation. In case of Jia, although her writing has been improved compared to the drafts in terms of readability, the consistency and smoothness between sentences and the overall flow remained irregular. Her writing also still contained several errors.

4.5.3 Effects of chatbot feedback on learner perceptions

Table 2 shows that all the participants acknowledged that the AI writing tutor chatbot was fun, easy, and convenient. They also indicated that they became more interested in English writing due to the experience of writing with the aid of the chatbot. They perceived it as very useful for their English writing and showed a strong desire to use it for their future writing.

In terms of learners' perception of the chatbot for each stage of writing, it turned out that the use of the chatbot for the pre-writing, revising, and editing stages was particularly fun, easy, and convenient. The participants also demonstrated that the use of the chatbot for each stage contributed to their increased interest in each stage and expressed a desire to use it in the future. Although none of them were dissatisfied with the use of the chatbot for the drafting stage, their perception of its usefulness and its contribution to their interest in the drafting stage of writing were less strong than in other stages. This result appears to stem from the fact that, as revealed by open-ended survey responses regarding difficulties with the AI writing tutor, four students (Jia, Mina, Sora and Yena) expressed a need for additional support during the drafting stage because the chatbot provided only individual words rather than complete sentences in English. Learners responded to this constraint differently: while some primarily used the drafting chatbot to obtain word-level Korean-to-English translations, others relied on it to review and refine their sentence composition.

Table 2. Learners Responses after Using Chatbots Across Different Writing Stages

Stage	Theme: Mean (SD)				
	Entertainment	Convenience	Contribution to Writing	Interest in Writing	Future Use
Overall	4.5 (0.5)	4.7 (0.5)	4.8 (0.4)	4.3 (0.7)	4.8 (0.4)
Pre-writing	4.3 (0.5)	4.0 (0.8)	4.5 (0.5)	4.3 (0.5)	4.7 (0.5)
Drafting	4.2 (0.4)	4.0 (0.6)	4.3 (0.7)	3.8 (0.9)	4.5 (0.8)
Revising	4.8 (0.4)	4.7 (0.5)	4.8 (0.4)	4.5 (0.5)	5.0 (0.0)
Editing	5.0 (0.0)	4.8 (0.4)	5.0 (0.0)	4.8 (0.4)	5.0 (0.0)

Except for one student (Lael) who noted the chatbot's helpfulness during the drafting stage, specifically highlighting its provision of unfamiliar words, students perceived chatbot feedback during the revising and editing stages to be the most helpful for their writing. They appreciated how the chatbot provided feedback on content and organization, as well as on grammatical errors and unnatural language.

The most interesting aspect of using the AI writing tutor chatbot, according to the students, was its ability to provide the information they sought. Additionally, one participant (Lael) found the chatbot's provision of individual words during the drafting stage particularly interesting. She found it interesting that the drafting chatbot did not provide a whole sentence, as this approach fostered autonomy and helped improve English proficiency. Another participant (Yena) noted that the chatbot's editing capabilities were novel and intriguing.

5. Discussion

We have demonstrated how generative AI can be effectively developed and integrated into L2 English process writing, offering both theoretical insights and practical applications. Now we want to discuss the major findings with respect to the three research questions.

Study 1 examined how we developed GPT API-based AI chatbots as adaptive and facilitative tutors for L2 process writing. The major goal of this research was to develop useful L2 process writing tutor chatbots with the GPT API, and the goal was achieved through a multi-layered approach encompassing strategies, tools, and processes. This achievement highlights the significant potential of generative AI in L2 process-oriented writing instruction (Law 2024).

Our main strategy was to make four different chatbots, which shared the same identity as an ESL writing tutor helping Korean middle school students but were assigned to different stages of process writing (cf. Zou and Huang 2023). This strategy was feasible because OpenAI Playground allows the customization of chatbots' roles and constraints via prompt fine-tuning. The chatbot development prompts were fine-tuned to promote the major functions of each process-writing stage and to prevent possible drawbacks such as learners' over-reliance and chatbots' information overload (Barrot 2023, Strobl et al. 2024).

In addition, the design of the chatbot architectures and web interfaces appeared to meet the two criteria for usefulness of language learning chatbots (Huang et al. 2021). The first criterion is *usability*, which demands that chatbots for language learning should be easy to use and personalized. Our writing tutor chatbots appear to meet this criterion since a) the user interface is simple and intuitive, and b) every learner-chatbot interaction is unique as the chatbots' responses are generated in real time. The second criterion is *utility*, which demands that the chatbots perform various pedagogical roles and offer interpersonal and cohesive communication. This criterion is also satisfied since a) different chatbots are to focus on different stages of process writing, and b) the learner-chatbot interaction occurred in contextually cohesive ways.

Study 2 concerns how L2 learners interact with AI chatbots during process writing. Our findings indicate that both the learners and the chatbots are oriented to the major task of each step when they interact with each other, an important principle for L2 learning chatbots (Kim et al. 2022). For example, at the pre-writing stage, they appeared to have varying orientations according to their target genres. In the context of expository writing, the chatbot catalogued a variety of facts, then the learners chose an animal and asked for more facts. On the other hand, in the context of argumentative writing, the learners became more explorative of different stances, and the chatbot was more provocative, encouraging the learners to present their own opinions and decide how they would incorporate the information into their writing. Therefore, it is argued that the chatbots mediate between stage-specific writing tasks and L2 learners (Li 2018).

Another noteworthy finding is that the learners came up with various collaboration strategies to address tasks at hands. For example, as noted earlier, the drafting chatbot was designed not to provide complete English sentences, and the learners had to write their own sentences. Then, some learners were uncertain whether their sentences express what they wanted to express. To address this uncertainty, they asked the chatbots to translate their English sentences into Korean or evaluate the grammaticality of their own sentences, and assessed the accuracy and appropriateness of their language based on the chatbots' responses. Provided that the learners had not interacted with any writing tutor chatbots before, their use of various strategies has significant implications for L2 writing pedagogy. First, when L2 learners are provided with intelligent tools, they promptly understand what they can and cannot do with the tools. Second, L2 learners creatively collaborate with intelligent tools to address specific challenges in L2 writing.

Finally, we have investigated the effects of the AI chatbots on L2 learners' writing quality and perceptions. As shown in Figure 2, the learners generally succeeded in performing the major tasks in the drafting and editing stages as the chatbots provided learner-friendly feedback for brainstorming and error correction. However, variations were observed during the revision stage, particularly in organizational revision. Two learners (Ian, Lael) successfully structured their writing from the outset with support from the pre-writing chatbot and made minimal changes in organization at later stages. On the other hand, three other learners (Mina, Yena, Sora) were able to improve their writing structure with guidance from the revising chatbot, while one learner (Jia) continued to experience difficulties. These results suggest that the accessibility of the GPT API chatbots' feedback for process writing may vary among L2 learners. In particular, revising organizational aspects requires skills such as identifying inadequacies, restructuring for the target audience, and ensuring text cohesion and coherence, which pose different degrees of challenges for L2 learners. Therefore, chatbot feedback related to organization may necessitate human support or instructional intervention, such as the examination of revision cases (cf. Zhu et al. 2024).

We have also found that the L2 learners, after the process writing task, perceived the AI chatbots as knowledgeable and helpful in English writing and wanted to use the chatbots in the future. In particular, they gave the highest scores to the revising and editing chatbots, reporting that these two chatbots' customized feedback on content, organization, grammar, and naturalness was the most helpful for their writing. In contrast, the learners' evaluation of the drafting chatbot was less positive than the other chatbots. Given that drafting was reported as the most challenging stage, participants may have expected more substantial support during this phase, such as full-sentence generation. This unmet expectation could have contributed to their less positive evaluation. Although half of the students reported that they felt difficulties in writing English sentences at the drafting stage, the rater evaluation of their drafts was higher than that of the other drafts (mean scores 3.46 vs. 3.17). This may imply that the learners' perceived difficulties caused by the AI chatbot's limited assistance facilitate their cognitive processes and thus contributes the overall writing quality.

6. Conclusion

The present study investigated the development and implementation of GPT API-based AI chatbots as adaptive and facilitative tutors for L2 English process writing among Korean middle school students. The major findings revealed that these chatbots could effectively assist students at various stages of the writing process, including pre-writing, drafting, revising, and editing. The stage-specific feedback provided by the chatbots helped students generate ideas, structure their writing, and refine their drafts with improved clarity and coherence.

The significance of this study lies in its innovative approach to integrating AI technology in L2 writing instruction for adaptive and facilitative scaffolding. Our chatbots encouraged independent thinking and limited excessive feedback to address potential drawbacks, such as overreliance on the chatbot and information overload for the learner. This implies that the future design of AI-powered pedagogical chatbots should consider both potential benefits and drawbacks.

Moreover, this study fills a gap in the research by focusing on a less-studied population, i.e., secondary EFL students, thereby providing valuable insights into the applicability of AI in diverse educational contexts. For example, the chatbots in this study were used alongside various materials such as teacher demonstrations, worksheets, and cloud-based apps, suggesting that the integration of AI technology into the classroom should be purpose-driven and supported by accessible resources.

However, the study also identified several limitations, suggesting avenues for future research. First, the writing tutor chatbots were developed using a single API, GPT-4-turbo; therefore, it remains uncertain which type of API best supports L2 process writing. Further research is needed to explore the potential of recent LLM-based APIs, such as GPT-4o and DeepSeek, for L2 process writing. Second, the small sample size of only six students in Study 2 restricted the generalizability of the findings. In particular, there may be self-selection bias as the participants may have had a greater interest in using technology for learning, which could have influenced their engagement with and perceptions of the chatbots. To mitigate this bias, future studies should seek a more representative sample, including students with varying levels of technological familiarity and interest. Third, the participants' feedback on using the chatbots was quite positive, which may be attributed to their lack of prior experience incorporating chatbots into English writing. Therefore, future research should adopt a long-term approach to examine whether these positive perceptions persist over time and thus are not bound to the novelty effect. Finally, this study evaluated the impact of chatbots through a single writing task. Longitudinal research involving multiple writing tasks and a thorough analysis of students' writing proficiency development over time is necessary to ascertain the long-term benefits and challenges of using AI chatbots as writing tutors. By addressing these limitations, future research will be better positioned to determine the efficacy and adaptability of AI tools in L2 writing pedagogy.

References

- Bahja, M., R. Hammad and G. Butt. 2020. A user-centric framework for educational chatbots design and development. In *HCI International 2020-Late Breaking Papers: Multimodality and Intelligence: 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19-24, 2020, Proceedings 22*, 32-43. Springer International Publishing.
- Barrot, J. S. 2023. Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing* 57, 100745.
- Butterfield, A. L. and C. K. Jung. 2013. A hybrid approach to second language writing. *Journal of the Korean English Education Society* 12(1), 123-139.
- Casal, J. E. and M. Kessler. 2023. Can linguists distinguish between ChatGPT/AI and human writing?: A study of research ethics and academic publishing. *Research Methods in Applied Linguistics* 2(3), 100068.
- Crossley, S. A. 2018. Technological disruption in foreign language teaching: The rise of simultaneous machine translation. *Language Teaching* 51(4), 541-552.
- De Larios, J. R., L. Murphy and J. Marín. 2002. A critical examination of L2 writing process research. In S. Ransdell and S. Barbier, eds., *New Directions for Research in L2 Writing*, 11-47. Springer.
- Ene, E. and V. Kosobucki. 2016. Rubrics and corrective feedback in ESL writing: A longitudinal case study of an L2 writer. *Assessing Writing* 30, 3-20.
- Ghafouri, M., J. Hassaskhah and A. Mahdavi-Zafarghandi. 2024. From virtual assistant to writing mentor: Exploring the impact of a ChatGPT-based writing instruction protocol on EFL teachers' self-efficacy and learners' writing skill. *Language Teaching Research* Online First, 1-23.
- Guo, K. and D. Li. 2024. Understanding EFL students' use of self-made AI chatbots as personalized writing assistance tools: A mixed methods study. *System* 124, 103362.
- Hong, K. 2014. How to assess L2 English writings from L2 English teachers' perspective. *The Journal of Mirae English Language and Literature* 19(4), 365-390.
- Huang, W., K. F. Hew and L. K. Fryer. 2021. Chatbots for language learning—Are they really useful? A systematic

- review of chatbot-supported language learning. *Journal of Computer Assisted Learning* 38(1), 237-257.
- Hyland, K. 2003. *Second Language Writing*. Cambridge University Press.
- Hyland, K. 2007. Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing* 16(3), 148-164.
- Kang, S. and M. Sung. 2024. EFL students' self-directed learning of conversation skills with AI chatbots. *Language Learning & Technology* 28(1), 1-19.
- Kasneci, E., K. Sessler, S. Küchemann, M. Bannert, D. Dementieva, F. Fischer, U. Gasser, G. Groh, S. Günnemann, E. Hüllermeier, S. Krusche, G. Kutyniok, T. Michaeli, C. Nerdel, J. Pfeffer, O. Poquet, M. Sailer, A. Schmidt, T. Seidel, ... and G. Kasneci. 2023. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences* 103, 102274.
- Kim, H. 2016. Changes in EFL students' perceptions and approaches to writing in a process-oriented writing class. *Modern English Education* 17(1), 49-67.
- Kim, H., H. Yang, D. Shin and J. Lee. 2022. Design principles and architecture of a second language learning chatbot. *Language Learning & Technology* 26(1), 1-18.
- Kim, S., J. Shim and J. Shim. 2023. A study on the utilization of OpenAI ChatGPT as a second language learning tool. *Journal of Multimedia Information System* 10(1), 79-88.
- Kohnke, L., B. L. Moorhouse and D. Zou. 2023. Exploring generative artificial intelligence preparedness among university language instructors: A case study. *Computers and Education: Artificial Intelligence* 5, 100156.
- Koraishi, O. 2023. Teaching English in the age of AI: Embracing ChatGPT to optimize EFL materials and assessment. *Language Education & Technology* 3(1), 55-72.
- Law, L. 2024. Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open* 100174.
- Li, M. 2018. Computer-mediated collaborative writing in L2 contexts: An analysis of empirical research. *Computer Assisted Language Learning* 31(8), 882-904.
- Lipson, M. Y., J. Mosenthal, P. Daniels and H. Woodside-Jiron. 2000. Process writing in the classrooms of eleven fifth-grade teachers with different orientations to teaching and learning. *The Elementary School Journal* 101(2), 209-231.
- Liu, X. 2024. Formative writing assessment: An EFL teacher's beliefs and practices. *Changing English* 31(2), 200-210.
- Miao, F. and W. Holmes. 2023. *Guidance for Generative AI in Education and Research*. UNESCO Publishing.
- Nicolás-Conesa, F., J. R. de Larios and Y. Coyle. 2014. Development of EFL students' mental models of writing and their effects on performance. *Journal of Second Language Writing* 24(1), 1-19.
- Ningrum, S. 2023. ChatGPT's impact: The AI revolution in EFL writing. *Borneo Engineering & Advanced Multidisciplinary International Journal* 2(SI), 32-37.
- Octavio, M. M., M. V. G. Argüello and J. T. Pujolà. 2024. ChatGPT as an AI L2 teaching support: A case study of an EFL teacher. *Technology in Language Teaching & Learning* 6(1), 1-25.
- Park, J. 2017. Benefits of the process approach in L2 writing: From the prospective teacher's perspectives. *The Journal of Linguistic Science* 82, 117-136.
- Rad, H. S., R. Alipour and A. Jafarpour. 2023. Using artificial intelligence to foster students' writing feedback literacy, engagement, and outcome: A case of Wordtune application. *Interactive Learning Environments* 32(9), 5020-5040.
- Ruiz-Rojas, L., P. Acosta-Vargas, J. De-Moreta-Llovet and M. Gonzalez-Rodriguez. 2023. Empowering education with generative artificial intelligence tools: Approach with an instructional design matrix. *Sustainability*

- 15(15), 11524.
- Shi, H. and V. Aryadoust. 2024. A systematic review of AI-based automated written feedback research. *ReCALL* 36(2), 187-209.
- Song, C. and Y. Song. 2023. Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology* 14, 1260843.
- Steiss, J., T. Tate, S. Graham, J. Cruz, M. Hebert, J. Wang, Y. Moon, W. Tseng, M. Warschauer and C. B. Olson. 2024. Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction* 91, 101894.
- Strobl, C., I. Menke-Bazhutkina, N. Abel and M. Michel. 2024. Adopting ChatGPT as a writing buddy in the advanced L2 writing class. *Technology in Language Teaching & Learning* 6(1), 1-19.
- Van Lier, L. 2005. Case study. In E. Hinkel, ed., *Handbook of Research in Second Language Teaching and Learning*, 195-208. Erlbaum.
- Wang, X., Q. Liu, H. Pang, S. C. Tan, J. Lei, M. P. Wallace and L. Li. 2023. What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education* 194, 104703.
- Williams, M. and R. L. Burden. 1997. *Psychology for Language Teachers: A Social Constructivist Approach*. Cambridge University Press.
- Wortham, S. and A. Reyes. 2020. *Discourse Analysis Beyond the Speech Event*. Routledge.
- Yan, D. 2023. Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies* 28(11), 13943-13967.
- Yang, H. 2018. Influence of learner training on students' process writing in automated writing evaluation-supported class. *Multimedia-Assisted Language Learning* 21(3), 88-114.
- Yuan, Y., H. Li and A. Sawaengdist. 2024. The impact of ChatGPT on learners in English academic writing: Opportunities and challenges in education. *Language Learning in Higher Education* 14(1), 41-56.
- Yuen, C. L. and N. Schlote. 2024. Learner experiences of mobile apps and artificial intelligence to support additional language learning in education. *Journal of Educational Technology Systems* 52(4), 506-525.
- Zhu, Y., L. Zhang, R. Zong and H. Sun. 2024. Exploring AI-generated feedback on English writing: A case study of ChatGPT. *US-China Foreign Language* 22(3), 144-153.
- Zou, M. and L. Huang. 2023. The impact of ChatGPT on L2 writing and expected responses: Voice from doctoral students. *Education and Information Technologies* 29(11), 13201-13219.

Examples in: English, Korean
 Applicable Languages: English
 Applicable Level: Secondary

Appendix. Worksheets for Argumentative Essay Writing (Originally in Korean)

English Writing Class with AI Writing Tutor

Topic: Do you support banning single-use plastics?

⇒ Decide your stance and write an argumentative essay in English

1. Pre-Writing

- **Explanations:** After discussing with the AI Writing Tutor and reviewing both the pros and cons, decide your stance and outline what content you will include in your essay.
- **Steps:** Discuss with the #1 AI Tutor. => Submit the conversation log to the survey.
=> Complete the following 1) and 2).

- 1) My stance: _____
- 2) Summary of brainstorming results

Summary of Brainstorming Results

2. Drafting

- Explanations: Based on the brainstorming results, discuss with the AI Writing Tutor and write the first draft.
- Steps: Discuss with the #2 AI Tutor while writing the essay. => Submit the conversation log to the survey.

[illegible]

3. Revising

- Explanations: Send your first draft to the AI Writing Tutor, receive feedback on the structure and content of your writing, then revise your essay incorporating the evaluation results.
- Steps: Submit the essay to the #3 AI Tutor for evaluation and discuss with the AI Tutor.
 - => Submit the conversation log to the survey.
 - => Summarize the evaluation results and revise your essay.
- Evaluation results by the AI Tutor

Strengths	Areas for Improvement

- Type on Google Docs incorporating the evaluation results.

4. Editing

- Explanations: Send the revised draft to the AI Writing Tutor for correction and receive feedback. Based on the corrections, create the final version.
- Steps: Submit the essay to the #4 AI Tutor for evaluation and engage in discussion.
 - => Submit the conversation log to the survey => Write the final draft.