Ziqi, C., Xinhua, Z., Qi, L., & Wei, W. (2026). L2 students' barriers in engaging with form and content-focused AI-generated feedback in revising their compositions. Computer Assisted Language Learning, 39(3), 715–735. https://doi.org/10.1080/09588221.2024.2422478
This is an Accepted Manuscript of an article published by Taylor & Francis in Computer Assisted Language Learning on 06 Nov 2024 (published online), available at: https://doi.org/10.1080/09588221.2024.2422478.

L2 students' barriers in engaging with form and content-focused AI-generated feedback in revising their compositions

Providing corrective feedback to second language (L2) writing constitutes a crucial digital affordance for AI-assisted writing systems. However, L2 writers' revision strategies and obstacles to adopting AI-generated feedback, such as ChatGPT, remain unclear. Forty-five L2 students in a computer science program were tasked with seeking corrective feedback from ChatGPT for their argumentative essays, followed by an analysis of their revisions and rationale for feedback uptake strategies. The findings revealed that approximately 38% of the feedback was either explicitly argued (22%) or ignored (16%). Upon controlling for writing proficiency, participants statistically rejected a significantly higher proportion of feedback at the content level (e.g., evidence) than at the form level (e.g., grammar). Utilizing the Technology Acceptance Model, the reasons for rejecting or ignoring ChatGPT-generated feedback were examined through participants' reflective data, focusing on two perspectives: inconvenience to use and unusefulness. Inconvenient factors included (1) overload feedback, (2) provision of general descriptions instead of specific error highlighting, and (3) repetitive and tedious comments. Themes related to unusefulness encompassed (1) misinterpretation of authors' intentions, (2) lack of clarity and illustrative examples, and (3) extraneous and irrelevant feedback. The implications entail pedagogical strategies to mitigate barriers and foster feedback literacy in AI-assisted educational environment.

Keywords: Generative AI, ChatGPT, AI-generated feedback, uptake, revision strategies

Introduction

AI-generated written feedback or automatic written evaluation systems have been developed to assist teachers' feedback and students' self-feedback practice (Blackie, 2024; Farazouli et al., 2023; Authors, 2023). Their effectiveness in promoting student reflective writing, self-feedback, and self-evaluation practice has recently gained popularity in second language teaching and learning (Barrot, 2020; Li et al., 2023),

exemplified by tools such as Grammarly and Chatbots. One line of inquiry focuses on their contribution to developing second-language writing. At the form level, this includes considerations of lexical complexity and grammar accuracy (Barrot, 2020), while at the content level, it encompasses the quality of argumentation (Guo et al., 2021), relevance of evidence, structure, and organisation (Huang et al., 2021). Two crucial factors underlying these promising outcomes are learners' initiatives in seeking feedback from AI writing systems and subsequent revisions. L2 writers' interactions with automatic written feedback from AI-assisted learning systems and their revision strategies have garnered increasing attention from researchers in language teaching and learning (Zhang & Hyland, 2018). Although Liu et al., (2024) reported negative perceptions of using ChatGPT for language learning in general, it remains less clear what revision strategies the learners use and what challenges second language (L2) writers face when comprehending and incorporating automatically generated feedback. This issue warrants further investigation, especially in light of the advent of ChatGPT, given its remarkable capacity to provide feedback and suggestions at both form and content levels conversationally. Consequently, this study explores L2 writers' utilisation and perceptions of corrective feedback from the AI-assisted writing system (ChatGPT-3.5) across both form and content levels in completing academic writing tasks.

Literature Review

Studies of Technology-assisted feedback

Technology-assisted feedback systems employ various learning tools and platforms, including computers, mobile devices, tablets, and robots, to facilitate corrective feedback among peers or teachers. Regardless of the platform utilised, numerous studies have demonstrated the positive impact of technology-assisted feedback tools on

learners' language acquisition (Bahari, 2022), particularly in terms of (1) increased learner engagement and motivation (Shih, 2011; Saeed & Alharbi, 2023), and (2) cultivation of learners' critical thinking and reflection (Novakovich, 2015). Despite the beneficial effects of integrating technology for feedback purposes in educational settings, barriers persist in teaching practices. Several studies have highlighted that these barriers primarily originate from the teachers' perspective (Abdelhalim, 2023; Li et al., 2019). Firstly, many educators lack competence and confidence in utilising technological resources in their instructional practices. For example, as highlighted by Abdelhalim (2023) in his study involving two Saudi university teachers, teachers' negative attitudes shaped their interaction with technologies and influenced the areas in which they utilised them. Furthermore, through a controlled experiment involving six teachers with varying rates of technology use, the study by Li et al. (2019) suggested that despite teachers possessing sufficient proficiency and confidence in utilising technological resources in their instructional practices, unless they are well-versed in practical pedagogical approaches that enhance learners' learning through the use of technological tools, the integration of educational technology will have minimal impact on corrective feedback practices. Considering the affordances of ChatGPT in language learning, further investigation into its barriers and challenges to its users is warranted.

Studies in AI automatic corrective feedback

An AI-powered learning system offers automatic corrective feedback for written assignments, with numerous tools available for this purpose, such as Grammarly, Pigai, and Criterion. Three primary findings have been reported regarding the benefits of these tools for teaching and learning: (1) improving the accuracy of learners' writing (Barrot, 2020; Barrot, 2021), (2) improving the efficiency of teacher feedback on student writing

(Sanosi, 2022), and (3) promoting greater learner autonomy in learning and using feedback (Moore & MacArthur, 2016; Zhang & Hyland, 2018).

Automatic corrective feedback can improve learners' writing accuracy by correcting assignment errors. For instance, Barrot's (2020) study examined the impact of Grammarly on L2 learners' writing accuracy. A controlled experiment involving 65 university students from the Philippines demonstrated that students who received feedback tutoring from Grammarly exhibited significantly higher grammatical accuracy in subsequent written tasks than those in the control group. Barrot (2021) attributed this primarily to Grammarly's ability to offer real-time feedback, such as highlighting errors (spelling, grammar, punctuation) and addressing inappropriate language use (intonation, formality, and politeness) within the text. Furthermore, automatic corrective feedback may afford language teachers more time to prepare for classroom instruction. Sanosi (2022) compared the identification of grammatical mistakes by teachers and Grammarly. The findings indicated that the machine is as proficient as experienced teachers in identifying mistakes in students' written work. Consequently, automatic corrective feedback has the potential to streamline the process of identifying and correcting writing errors, thereby enabling teachers to devote more attention to assisting students in developing their writing skills and critical thinking abilities. Additionally, automatic corrective feedback tools promote learner autonomy by facilitating the utilisation of received feedback for further revision. A study by Zhang and Hyland (2018) examined the feedback received by two L2 students from their writing teacher and an automated corrective feedback system. The study concluded that students' engagement with automatic corrective feedback was enhanced by their ownership of their learning process, enabling them to adjust their writing schedules to align with their learning objectives.

Simultaneously, addressing several drawbacks associated with automatic corrective feedback tools is imperative. Firstly, these tools often fail to accurately identify all errors in the text, primarily due to their technical limitations (Lang et al., 2019; Zhang & Hyland, 2018; Liu & Kunnan, 2016). For instance, Liu and Kunnan (2016) discovered in their research that WriteToLearn exhibited inaccuracies in identifying prepositions, word choices, and expressions. Secondly, the feedback provided by these tools tends to be limited in explanation and prone to over-correction, posing challenges for students, particularly those with weaker language skills (Barrot, 2020; Shang, 2019; Yang et al., 2023). For example, Barrot's (2020) study revealed that some students expressed concerns about Grammarly's tendency to over-correct and provide insufficient explanations, leading to frustration among students. Lastly, AlGhamdi (2024) investigated the impact of ChatGPT-generated feedback on the writing skills of 111 first-year computing students, and the students expressed concerns that the feedback provided by ChatGPT was inconsistent. At the same time, Darvishi et al. (2024) revealed that students tend to rely on rather than learn from AI-assisted writing systems. Therefore, the potential overreliance on generative AI-powered Chatbots for feedback generation among second-language learners warrants further scrutiny.

Barriers to Student uptake of corrective Feedback

While corrective feedback practice has positively impacted L2 learners' writing process and performance (Kang & Han, 2015), it would be overly simplistic to assume that it operates seamlessly in a classroom without encountering barriers. Barriers to effective corrective feedback practice have been categorised into factors related to L2 learners and the interactions and communication dynamics between feedback providers. From the learners' perspective, several challenges and difficulties have been identified: (1) a

lack of self-monitoring may lead to diminished motivation in engaging with feedback (Li & Zhang, 2023), (2) limited English proficiency levels may hinder students' ability to comprehend and implement feedback into their revisions (Wang, 2014; Zheng & Yu, 2018), (3) a lack of clarity regarding teachers' expectations and assignment rubrics, and (4) repetitive feedback activities may result in passive student engagement (Wang, 2014).

For example, Li and Zhang (2023) examine the disparities in corrective feedback practices among college students with varying levels of self-monitoring. Interview data indicated that learners with lower self-monitoring encountered difficulties effectively utilising feedback to enhance their learning process, likely due to limited awareness, proficiency in interpreting feedback, and emotional detachment. Additionally, Wang (2014) delves into students' perceptions of peer feedback at a mainland Chinese university, revealing a decline in students' enthusiasm for feedback over time. One plausible explanation for students disregarding feedback from corrective feedback systems is the prevalence of excessive and repetitive feedback activities. Furthermore, the study highlights a lack of understanding regarding teachers' expectations and rubrics for written assignments, leading to feedback that diverges from students' expectations and fosters a negative attitude towards feedback activities.

Secondly, their interactions during the corrective feedback practice may entail two types of challenges: 1) inadequate communication between teachers and students (McMartin-Miller, 2014), and 2) learners' apprehensions regarding interpersonal relationships (Wang, 2014). For instance, in McMartin-Miller's (2014) qualitative research study, interviews were conducted with three American writing lecturers and their students, revealing insufficient communication between teachers and students.

Students often fail to fully comprehend or misconstrue teachers' intentions and their roles in the feedback process. Even within the context of peer feedback, where no social hierarchical relationship exists between feedback providers and recipients, communication may be compromised due to fears of damaging relationships and disrupting class harmony (Wang, 2014).

Since its formal launch in November 2022, ChatGPT has attracted increasing scholarly interest in generative AI. Upon reviewing the existing literature, we found that most empirical studies on applying ChatGPT in academic writing have been conducted in fields such as medicine, physics, and engineering. These studies highlight the robust contextual understanding (Lin, 2023; Ray, 2023) and the ability to learn from human feedback (Ray, 2023) exhibited by ChatGPT. However, its implications for academic writing among second language learners remain ambiguous, particularly regarding the potential for addressing previously mentioned barriers through adapted AI-generated feedback. Furthermore, existing research on ChatGPT has predominantly focused on educators' negative perceptions and concerns about extensive text borrowing, plagiarism, and unethical use (Cotton et al., 2023). The challenges associated with how L2 students learn from and respond to corrective feedback have received scant investigation. Grounded in the Technology Acceptance Model, this study aims to address two research questions:

(1) What are the L2 learners' strategies to engage with ChatGPT-generated corrective feedback, and

(2) What are the barriers for L2 learners to act upon ChatGPT-generated feedback?

Research design

Research context

Fifty-nine undergraduate students enrolled in an academic writing course within a Computer Science program at a public university in Macao were invited to participate in this study. Their ages ranged from 20 to 21 years old (*SD* = 1.72), with 52 male and seven female students. All participants had completed courses in the first semester and possessed familiarity with the computer operating system, the basics of AI technologies, and English academic writing. Since the classroom instruction language was English, participants likely felt comfortable communicating with the AI system in English. Of the sample, fifty-five were native Cantonese or Mandarin speakers, while two were Portuguese, one English, and one Filipino native speaker. Based on their scores in the English language test of the national university entrance examination in mainland China, students must achieve at least 100 out of 150 to qualify for admission to the school. This score corresponds to the B1 level on the CEFR band scale (Common European Framework of Reference for Languages) (Jin et. al, 2017). Thus, all participants in our study were at least at the B1 level on the CEFR band scale. By the time of data collection, which occurred following the official launch of ChatGPT-3.5 (All subsequent references to ChatGPT are to GPT-3.5) at the end of November 2022, most participants had already utilised the chatbot for academic and entertainment purposes, either then or before.

Data collection

These students were initially contacted to participate in the study through Canvas, a centralised online learning system. Forty-five of them expressed interest in participating. Subsequently, the students were briefed on the data collection procedure.

Initially, they were asked to independently draft a 250-word essay without the assistance of ChatGPT under classroom conditions. The writing prompt was: *To what extent do you agree or disagree that AI technologies may replace human workforces?* This topic was chosen for several reasons: it had been discussed in one of their mandatory subjects during the first semester, Ethics and Professional Issues, therefore, they were familiar with both 1) content, such as key concepts, relevant ideas from supporters and opponents, and 2) linguistic resources, such as terminologies and writing styles.

Secondly, they underwent training to utilise ChatGPT-3.5 to provide corrective feedback on their drafts and revise their essays by the end of April 2023. This training encompassed (1) addressing academic integrity concerns associated with employing ChatGPT for academic assignments, (2) familiarising participants with university guidelines governing the use of ChatGPT for academic purposes, and (3) providing guidelines and examples for writing reflective journals to document their feedback experiences with ChatGPT.

Two prompts were provided to facilitate participants' conversations with ChatGPT: (1) *Please act as an English language teacher and provide corrective feedback on my essay with explanations. My essay is attached below.* and (2) *Please make some suggestions and provide exemplars to improve the quality of my essay based on the attached marking rubrics, which are designed around second language writing, encompassing both form-focused (Grammatical accuracy, Lexical sophistication) and content-focused criteria (Quality of argument structures, Orientations to source evidence)* (Barrot, 2020; Cheong et al., 2022; Authors, 2023; Huang et al., 2021). Participants were encouraged to pose additional follow-up questions to ChatGPT based on its corrective feedback if necessary, such as (1) seeking clarifications on the

feedback and suggestions, (2) requesting more examples to comprehend the feedback, and (3) engaging in arguments with ChatGPT if disagreeing with the feedback. Subsequently, participants were required to submit their responses to the ChatGPT-generated feedback item by item, either in their first language or in English, as they preferred, ensuring that language barriers did not hinder their expression of nuanced perceptions around the feedback's utility. However, subsequent data reflection confirmed that all participants carried out their revision plans in English. As depicted in Table 1, participants were expected to copy and paste ChatGPT-generated feedback into the first column and provide their responses and revisions next to them item by item. Consequently, the collected data in this study encompassed (1) participants' initial drafts of their essays, (2) the corrective feedback and suggestions from ChatGPT, and (3) L2 learners' responses and actual revisions. An example of collected data can be found in Appendix 1.

Table 1. Example of L2 writers' revision strategies on ChatGPT generated feedback

Feedback from ChatGPT	**Your response**
1. Use proper grammar and punctuation - There are several grammatical errors in the essay, including improper use of commas and run-on sentences. Make sure to proofread your work before submitting it.	*I agreed that, and I provide made revision accordingly. I change "get in touch with people" to "easily connect with people". "had convenient our daily lives" to "has made our daily lives more convenient". "book for position" to "make reservations".*
2. Use appropriate language: Choose your words carefully and use appropriate language for your audience and purpose. Avoid using overly technical terms or jargon that your readers may not understand. Also, make sure to vary your sentence structure to keep your writing interesting.	*I don't agree that, think my words are all very easy. I don't use any jargon in my essay. The evidence I chosen were closed to our lives.*
3. Clarify and elaborate on your points - Some of your arguments could be more	*I agree with that, but I can't do so. Because that will over the word count.*

specific and detailed. For example, you could provide examples of jobs that have been replaced by machines or describe how addiction to virtual technology affects young people's lives.	

Data analysis

To address the first research question (RQ1), "What are the L2 learners' strategies to engage with ChatGPT-generated feedback?" participants' responses to ChatGPT-generated feedback were examined. Drawing on previous research in corrective feedback and automatic written feedback, two frameworks were utilised to facilitate the coding process: levels of revisions and the taxonomy of revision strategies (refer to Table 2). Regarding the level of revisions, Fan and Xu (2020) distinguished between two types of feedback: form-focused revision and content-focused revision. Form-focused feedback, typically provided directly by teachers or peers, targeted various error types at the linguistic level, encompassing sentence structures, word choices, and verb forms. Intriguingly, they reported that content-focused feedback primarily occurred through negotiation. This implies that feedback providers prefer oral discussions and debates with recipients regarding content-related issues, such as evidence and examples. Concerning revision strategies, three categories were identified based on previous studies in corrective feedback, focusing on the rate of uptake or acceptance of received feedback (Wei et al., 2022): (1) accept and revise accordingly, (2) ignore or refrain from responding to the received corrective feedback (Zhang, 2020), and (3) demonstrate disagreement and confusion by engaging in arguments with the feedback providers, particularly in cases of non-error corrective feedback (Yang et al., 2023).

Table 2. Levels and taxonomy of revision strategies

Students' revision strategies	Content-focused	Form-focused

Accept		
Reject - Argue		
Reject - Ignore		

To address research question RQ2, "What are the barriers for L2 learners to act upon ChatGPT generated feedback?" participants' negative evaluations and justifications for not acting upon received ChatGPT feedback were identified and examined. Negative expressions such as rejection, disagreement, and confusion regarding the feedback were identified. Their justifications were analysed using the Technology Acceptance Model (TAM), which proposes easiness and usefulness as variables predicting new users' intentions for continuous use. The first author and corresponding author initially coded responses from the first 15 participants and collaboratively developed the framework. After resolving all disagreements, they coded responses from the remaining 30 participants, achieving an agreement rate of 87%.

To answer the second research question, our research focused on assessing and understanding learners' negative attitudes towards using AI-generated feedback. We strategically adapted the conventional TAM variables to align with our narrowed focus on uncovering challenges and barriers. Rather than employing the original framework's emphasis on perceived usefulness and perceived ease of use, we shifted our attention to their negative counterparts. Specifically, we focused on perceived unusefulness, a deviation from perceived usefulness, and perceived uneasiness/inconvenience, a deviation from perceived ease of use. This modification allowed us to capture and interpret the specific apprehensions and obstacles learners perceive more accurately. Under the newly conceptualised dimension of *uneasiness/inconvenience*, as shown in Table 3, we identified three critical sub-categories reflective of learner experiences: (a) overload feedback, (b) provision of general descriptions instead of specific error

highlighting, and (c) repetitive and tedious comments. Similarly, within *unusefulness*, our study outlined sub-categories including (a) misinterpretation of authors' intentions, (b) lack of clarity and illustrative examples, and (c) extraneous and irrelevant feedback.

Table 3. Coding framework on the barriers of acting upon ChatGPT-generated feedback

Categories	**Sub-categories**	**Explanations**
Uneasiness/inconvenience to use	Overload feedback	ChatGPT provided more than two suggestions in one piece of feedback. L2 learners responded in a selective way and ignored some suggestions
	Provision of general descriptions instead of specific error highlighting	ChatGPT described one type of mistake. L2 learners did not revise accordingly and asked for the locations of these mistakes.
	Repetitive and tedious comments	ChatGPT highlighted or corrected the same type of mistakes. L2 learners felt overwhelmed by the large number of mistakes and refused to revise all of them
Unusefulness	Misinterpretation of authors' intentions	ChatGPT revised essay, which changed the L2 learners' original ideas
	Lack of clarity and illustrative examples	ChatGPT did not explain the suggestions clearly enough
	Extraneous and irrelevant feedback	ChatGPT pointed out mistakes which did not exist in the essay or were irrelevant to the essay.

Results

Quantitative phase

Overall, 62% of the generative AI-generated feedback focused on content, encompassing the quality of argument structures and orientations to source evidence. Meanwhile, 38% related to form-related mistakes and errors, such as lexical resource selection and grammatical issues. Irrespective of the feedback focus, students accepted

the majority (62.2%) of feedback provided by ChatGPT. Conversely, approximately 38% of feedback was rejected in two ways: (1) "participants argued with the Chatbot, defending their decisions and refusing to make changes" (22%), and (2) "participants ignored feedback from the Chatbot without further communication or revisions" (16%). To provide further specificity, the percentages of AI Chatbot feedback on form-focused issues that were "accepted," "argued," and "ignored" were 68.1%, 15.6%, and 16.3%, respectively. In contrast, the corresponding percentages of feedback on content were 58.6%, 25.7%, and 15.8%. This indicates that participants demonstrated a higher level of disagreement with content-focused feedback.

Table 4 provides an overview of L2 learners' strategies for engaging with feedback on Form or Content. Notably, more students (62%) focused on content than form-focused feedback (38%). This suggests that students prefer generative AI feedback on grammar, spelling, and text structure errors over non-error feedback such as content clarity and logical coherence.

Table 4. Descriptive data and proportion of the students' Revision Strategies

Focus of Feedback	Student's Revision Strategies	Total number
Form (38%)	Accept (68.1%)	92
	Argue (15.6%)	21
	Ignore (16.3%)	22
Content (62%)	Accept (58.6%)	130
	Argue (25.7%)	57
	Ignore (15.8%)	35
Sum		357

An ANCOVA test was conducted to examine whether there was a statistically significant difference among learners' revision strategies after controlling for their pre-experiment academic writing proficiency. Assumptions for conducting the ANCOVA

test were assessed. The Kolmogorov–Smirnov test indicated no violations of normality. Levene's test results suggested equality of variances across samples. No significant interaction between the covariate and the dependent variable was observed, indicating no violation of the assumption of homogenous regression slopes (F (1,87) = 3.77, p = .06; F (1,87) = 3.74, p = .06; F (1,87) = 1.64, p = .20).

Table 5. ANCOVA of Students' Revision Strategies to Feedback

Students' Revision Strategies	Form	Content	*SS*	*MSE*	*F (1,87)*	*Sig.*	η^2
	Percentage in each group (%)						
Accept	68.1	58.6	0.68	0.23	5.39	.002**	.163
Argue	15.6	25.7	0.69	0.23	12.14	.000***	.302
Ignore	16.3	15.8	0.10	0.03	3.51	.019*	.116

***p < .001., **p < .01., *p < .05.

According to the ANCOVA test results presented in Table 5, the focus of the feedback, whether form-focused or content-focused, significantly influences students' acceptance of AI Chatbot-generated feedback. Specifically, the acceptance level of generative AI feedback on Form-focused was higher compared to Content-focused feedback (F (1,87) = 5.39, p = 0.002, partial η2 = 0.163), indicating that L2 students found the AI-generated form-focused corrections and suggestions to be more accurate or in line with their expectations than the content-focused feedback. Moreover, students are more inclined to argue generative AI feedback on writing Content than Form (F (1,87) = 12.14, p = 0.000, partial η2 = 0.302). Participants tend to be more sceptical regarding feedback related to content. Lastly, students are more likely to ignore corrective feedback related to Form than feedback associated with Content (F (1,87) = 3.51, p = 0.019, partial η2 = 0.116).

Qualitative phase

The students' revision strategies in response to AI-generated feedback were utilised to substantiate and elucidate the quantitative findings, addressing the second research question. The researchers analysed the interview transcripts to identify pertinent themes that could inform suitable study inquiries. Tables 6 and 7 outline the barriers encountered by L2 learners in acting upon ChatGPT-generated feedback guided by the Technology Acceptance Model.

Three themes were identified to capture students' inconvenient barriers when receiving feedback.

Table 6. Sub-themes under the category of Inconvenience

	Sub-themes	Percentage
Inconvenience	Overload feedback	63.04%
	Repetitive and tedious comments	28.26%
	Provision of general descriptions instead of specific error highlighting	8.70%

Concerning information overload in feedback, ChatGPT presents L2 learners with extensive feedback containing numerous suggestions. However, learners tended to respond to or disregard specific suggestions selectively. The abundance of suggestions can make it challenging for learners to discern which recommendations are most crucial or relevant to their learning goals. Consequently, they may inadvertently overlook valuable insights or fail to allocate their attention appropriately. For example:

> M23: *Its summary is too long to be suitable for its summary. I think it will be better if it gives me a short one.*

The repetitive and tedious comments at the form level pertain to ChatGPT consistently highlighting or correcting the same errors made by L2 learners. This

continual inundation of errors overwhelmed the learners, leading them to decline to revise each mistake. The inconvenience stems from ChatGPT's feedback becoming monotonous and predictable, as it repeatedly emphasises the same errors. This repetition may have rendered the learning process tedious and demotivating for L2 learners. For example:

> M15: *I also want to express myself in simpler language, the Chatbot keeps reminding me of that, but my vocabulary is small and I can only express myself as much as possible.*

ChatGPT provided a generic explanation of errors, offering a general description rather than highlighting specific mistakes. Consequently, L2 learners refrained from making necessary revisions and sought clarification on the specific instances of these errors. This inadvertently hinders learners' ability to recognise and rectify their linguistic shortcomings, as ChatGPT does not pinpoint the exact locations of errors. For example:

> F8: *ChatGPT didn't give me the exact sentences, like which sentence is unclear, just told me to revise all the paragraph, so I will not make any change on this point.*
> M16: *I don't agree with that, it doesn't point out the place I didn't statement clearly. I don't understand where I didn't state clear.*

Table 7. Sub-themes of Uselessness Guided by the Technology Acceptance Model

	Sub-themes	**Percentage**
Uselessness	Misinterpretation of authors' intentions	67.02%
	Lack of clarity and illustrative examples	22.34%
	Extraneous and irrelevant feedback	10.64%

Misinterpreting authors' intentions and ideas refers to instances where ChatGPT may revise the essay, altering the L2 learners' original ideas. When L2 learners

compose an essay, they invest significant effort in expressing their ideas and conveying their intended messages. However, ChatGPT, an AI language model, may not fully grasp the nuances and subtleties of the learners' writing. Consequently, it may misinterpret the authors' intentions and make revisions that diverge from the original meaning. For example:

> F5: I t*hink ChatGPT's comment is wrong, because I just want to speak from a neutral point of view to show some good and bad things coming out. Therefore, I will not make any change on this point.*

Second, lack of clarity and illustrative examples means ChatGPT provided insufficient explanations for its suggestions, making it challenging for L2 learners to comprehend and apply them thoroughly. Without adequate examples, learners may find connecting the suggestions to real-life writing situations challenging, hindering their ability to apply the knowledge effectively. For example:

> M30: *I do not understand these comments, I don't know which verb to replace.*

Extraneous and irrelevant feedback entails instances where ChatGPT identifies mistakes that do not exist in the essay or are irrelevant to the task. This discrepancy may stem from the model's inability to comprehend the intended meaning or the specific task requirements. Extraneous feedback can be frustrating for L2 learners who depend on AI language models for assistance, resulting in confusion and wasted time. For example:

> F8: *I think ChatGPT's comment has some wrong, because my first sentence of the paragraph is "On the other hand, technology also has some disadvantages." And ChatGPT give me the same sentence.*

Discussions

Regarding the first research question, this study reveals that L2 writers are more

receptive to form-focused level corrective feedback generated by ChatGPT, such as lexical resource selection and grammatical issues. This finding aligns with previous research conducted by Yang and Liu (2017), which emphasises the ability of AI corrective feedback systems to locate grammatical and syntactic errors accurately, whose direct suggestions help students to self-correct. Furthermore, collaborative feedback efforts between AI systems and teachers have been proposed as effective by Zhang and Hyland (2018), as such feedback aligns with established grading criteria and integrates L2 writers' prior feedback experiences. Additionally, providing examples, highlighting errors, and offering rewritten versions of problematic sentences have enhanced students' writing proficiency, particularly in reducing error rates related to grammar, vocabulary, spelling, and conjunctions (Song, 2019).

Additionally, statistical analysis has revealed L2 writers' significantly distinct revision strategies concerning form-focused and content-focused feedback: the uptake rates for form-focused feedback are considerably higher than those for content-focused feedback. This discrepancy can be attributed to L2 writers' perceptions of AI feedback (Mayordomo et al., 2022). Given their proficiency in the subject matter of the written assignments—Artificial Intelligence and Robots—these participants from the Computer Science program expressed dissatisfaction with feedback about examples, references, and arguments in their writing. Furthermore, this finding is consistent with research by Yang et al. (2023), indicating that students often overlook non-error feedback when refining their written assignments. Despite its linguistic significance, such feedback lacks contextual information and examples, akin to the content-based feedback examined in our study. Embracing this type of feedback can pose challenges for students.

This study addresses the second research question by examining the challenges or barriers encountered by L2 writers in utilising Generative AI technology to aid their essay writing, focusing on two perspectives: inconvenience and perceived usefulness. Specifically, inconvenient feedback from generative AI technology encompasses (1) overload feedback from ChatGPT, (2) repetitive and tedious comments on a form level from ChatGPT, and (3) ChatGPT providing general descriptions rather than highlighting specific mistakes. These findings corroborate previous research on AI-assisted writing, such as Barrot (2020) highlighted challenges related to generating accurate and comprehensive feedback using AI technology, including overcorrection, cognitive overload, and limited metalinguistic explanation. The limitations of ChatGPT in providing satisfactory responses may be attributed to existing technical issues (Lang et al., 2019; Liu & Kunnan, 2016), which hinder AI's understanding of the complexity of language and nuances of writing. Additionally, the low level of feedback literacy among L2 writers can impede their ability to effectively use and interpret automated feedback from AI-assisted writing systems (Zheng & Yu, 2018). This lack of understanding and familiarity with AI-generated feedback may hinder L2 writers' engagement with and benefit from the feedback provided by these systems, potentially resulting in feedback being ignored.

Moreover, this study delineates three specific challenges associated with the perceived uselessness of feedback provided by generative AI technology. These challenges encompass ChatGPT's misinterpretation of authors' intentions and ideas, the lack of clarity and illustrative examples in ChatGPT-generated feedback, and the occurrence of extraneous and irrelevant feedback from ChatGPT.

Firstly, a misalignment often arises between the author's perspective and the feedback provided by ChatGPT, particularly regarding content-focused feedback. This

observation aligns with the findings of Yang et al. (2023), who noted that students exhibit less inclination to utilize feedback incorporating linguistic resources (e.g., suggestions for collocations) than feedback solely correcting errors. ChatGPT's form-focused feedback emphasises grammar and spelling corrections, typically more manageable for students to address and, consequently, more readily accepted than feedback addressing content. Moreover, confronting content-focused feedback from AI-generated sources may necessitate a higher level of AI feedback literacy among students, effectively reconciling contradictory feedback. Feedback from various sources may encompass distinct domains (e.g., form, content), perspectives (AI, teachers, peers), and processes (revision, embellishment), requiring students to acquire reliable knowledge from diverse sources to evaluate AI-generated feedback effectively. With this ability, students could make well-informed decisions when integrating AI recommendations into their scholarly compositions. Previous research on feedback literacy consistently underscores the positive correlation between students' feedback literacy levels and their proficiency in integrating feedback from multiple sources into their revised assignments (Xu & Zhang, 2022; Zhang & Mao, 2023). Consequently, considering students' varying feedback literacy levels, one critical advantage of employing generative AI technologies in feedback practice lies in their capacity to address form-level errors during revision processes.

Secondly, this study emphasises the limited effectiveness of clarity and illustrative examples provided by the AI system. Similar findings were reported by Yang et al. (2023), who emphasised the importance of offering examples and contextual information in feedback provided by AI-based writing systems, as it enhances learner engagement in the learning process. In our study, the probable reason for this finding may stem from the inadequate historical data and communication accessible to

ChatGPT. This constrains its understanding of students' writing characteristics and experiences, impeding its capacity to furnish tailored micro-level feedback with clear examples.

Lastly, extraneous and irrelevant feedback from ChatGPT, such as errors not relevant to the essay, was observed. This finding resonates with Barrot's (2021) research, which cautioned against the potential pitfalls and ethical considerations surrounding AI-generated content. Barrot (2021) highlighted how AI auto-correction systems often substitute uncommon or highly context-dependent words with incorrect ones. This underscores the growing importance of equipping students with the skills to identify and critique erroneous information generated by AI chatbots. Enhancing AI systems' ethical judgment and discernment capabilities, alongside refining training data and algorithms, becomes imperative to address these issues and ensure the delivery of accurate and reliable feedback to L2 writers.

Implications and Limitations

The implications of this study encompass at least two dimensions. From a teacher's perspective, this study emphasises the significance of comprehending the receptiveness of second language learners to various feedback types when seeking feedback from AI-based writing systems. For instance, Generative AI can be effectively utilised by teachers to provide feedback on form-focused errors in students' writing. However, it is crucial for teachers to critically evaluate the feedback from Generative AI before using it for content-focused feedback. Also, teachers should acknowledge these barriers to students receiving AI-generated feedback and devise strategies to integrate AI technology into their feedback practices effectively. They can assist students in utilising AI systems for feedback effectively and navigate potential limitations and challenges. From the perspective of L2 writers, the findings of this study underscore the necessity

for a comprehensive understanding of the limitations and benefits of AI-generated feedback systems. L2 writers should know the potential challenges when seeking feedback from such systems. They should critically evaluate AI-generated feedback with rubrics, teacher feedback, and their learning needs and objectives (Nicol & Kushwah, 2023; Yan & Carless, 2021), particularly the content-focused corrective feedback generated from AI. Furthermore, L2 writers can utilise the findings of this study to standardise their prompts when interacting with artificial intelligence and receive more constructive feedback in their writing processes (Espejel et al., 2023; Henrickson & Meroño-Peñuela, 2023), prompt design (e.g., few-shots, Chain of Thought prompting, etc) helps to effectively design and get better results on different tasks with LLMs. Overall, the implications of this study provide valuable insights for L2 writers and teachers, offering guidance on the effective integration and utilisation of AI technology for feedback in L2 writing processes.

Bearing in mind the implications of our study, there are still some limitations that we should be aware of. Firstly, our results indicate that L2 students reported that the ChatGPT-generated overload feedback was one of the reasons why they refused to receive feedback. In reflection, we could have evaluated and improved the effectiveness of ChatGPT prompts through a pilot study, and this limitation could be addressed in future studies. Secondly, our study did not collect participant dialogues with ChatGPT, focusing on feedback evaluation rather than interaction. This presents a limitation in that it omits analysis of potential struggles with the AI's language level for learners at low English levels. Further research could be conducted to explore this gap in greater detail. Additionally, a primary limitation of our study is the use of convenience sampling. This non-probabilistic sampling method may result in a sample that is not representative of the broader population, thus affecting the generalizability of our

findings. Future research should consider using probabilistic sampling methods to enhance representativeness and reliability. Lastly, our study is limited to the responses of L2 students to generative AI feedback on their academic writing. Therefore, the results may not be generalizable to other learners, such as L1 language learners and K12 students.

Reference

Abdelhalim, S. M. (2023). Online writing instruction: Understanding university EFL instructors' beliefs and practices as a complex system. *Language Teaching Research*, Advanced online publication. https://doi.org/10.1177/13621688221150855

AlGhamdi, R. (2024). Exploring the impact of ChatGPT-generated feedback on technical writing skills of computing students: A blinded study. *Education and Information Technologies*. https://doi.org/10.1007/s10639-024-12594-2

Bahari, A. (2022). Affordances and challenges of technology-assisted language learning for motivation: A systematic review. *Interactive Learning Environments*, *31*(9), 5853–5873. https://doi.org/10.1080/10494820.2021.2021246

Barrot, J. S. (2020). Integrating Technology into ESL/EFL Writing through Grammarly. *RELC Journal*, *53*(3), 764–768. https://doi.org/10.1177/0033688220966632

Barrot, J. S. (2021). Using automated written corrective feedback in the writing classrooms: effects on L2 writing accuracy. *Computer Assisted Language Learning*, *36*(4), 584–607. https://doi.org/10.1080/09588221.2021.1936071

Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 1–13. https://doi.org/10.1080/02602938.2024.2335321

Cheong, C. M., Luo, N., Zhu, X., Lu, Q., & Wei, W. (2022). Self-assessment complements peer assessment for undergraduate students in an academic writing task. *Assessment & Evaluation in Higher Education*, *48*(1), 135–148. https://doi.org/10.1080/02602938.2022.2069225

Cotton, D., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1–12. https://doi.org/10.1080/14703297.2023.2190148

Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers and Education/Computers & Education*, *210*, 104967. https://doi.org/10.1016/j.compedu.2023.104967

Espejel, J. L., Ettifouri, E. H., Alassan, M. S. Y., Chouham, E. M., & Dahhane, W. (2023). GPT-3.5, GPT-4, or BARD? Evaluating LLMs reasoning ability in zero-shot setting and performance boosting through prompts. *Natural Language Processing Journal*, *5*, 100032. https://doi.org/10.1016/j.nlp.2023.100032

Fan, Y., & Xu, J. (2020). Exploring student engagement with peer feedback on L2 writing. *Journal of Second Language Writing*, 50, 100775. https://doi.org/10.1016/j.jslw.2020.100775

Farazouli, A., Pargman, T. C., Laksov, K. B., & McGrath, C. (2023). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment & Evaluation in Higher Education*, 1–13. https://doi.org/10.1080/02602938.2023.2241676

Guo, Q., Feng, R., & Hua, Y. (2021). How effectively can EFL students use automated written corrective feedback (AWCF) in research writing? *Computer Assisted Language Learning*, *35*(9), 2312–2331. https://doi.org/10.1080/09588221.2021.1879161

Henrickson, L., & Meroño-Peñuela, A. (2023). Prompting meaning: a hermeneutic approach to optimising prompt engineering with ChatGPT. *AI & Society*. https://doi.org/10.1007/s00146-023-01752-8

Huang, W., Hew, K. F., & Fryer, L. K. (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. https://doi.org/10.1111/jcal.12610

Jin, Y., Wu, Z., Alderson, C., & Song, W. (2017). Developing the China Standards of English: challenges at macropolitical and micropolitical levels. *Language Testing in Asia, 7*(1). https://doi.org/10.1186/s40468-017-0032-5

Kang, E. Y., & Han, Z. (2015). The Efficacy of written corrective feedback in improving L2 Written Accuracy: A Meta-Analysis. *The Modern Language Journal*, *99*(1), 1–18. https://doi.org/10.1111/modl.12189

Lang, F., Li, S., & Zhang, S. (2019). Research on Reliability and Validity of Mobile Networks-Based Automated Writing Evaluation. *International Journal of Mobile Computing and Multimedia Communications*, *10*(1), 18–31. https://doi.org/10.4018/ijmcmc.2019010102

Li, G., Sun, Z., & Jee, Y. (2019). The more technology the better? A comparison of teacher-student interaction in high and low technology use elementary EFL classrooms in China. *System*, *84*, 24–40. https://doi.org/10.1016/j.system.2019.05.003

Li, Y., & Zhang, L. J. (2023). Self-regulation and student engagement with feedback: The case of Chinese EFL student writers. *Journal of English for Academic Purposes*, *63*, 101226. https://doi.org/10.1016/j.jeap.2023.101226

Li, Y., Sha, L., Yan, L., Lin, J., Raković, M., Galbraith, K., Lyons, K., Gašević, D., & Chen, G. (2023). Can large language models write reflectively. *Computers & Education: Artificial Intelligence*, *4*, 100140. https://doi.org/10.1016/j.caeai.2023.100140

Lin, P. M. (2023). ChatGPT: Friend or foe (to corpus linguists)? *Applied Corpus Linguistics*, *3*(3), 100065. https://doi.org/10.1016/j.acorp.2023.100065

Liu, G., Darvin, R., & Ma, C. (2024). Exploring AI-mediated informal digital learning of English (AI-IDLE): a mixed-method investigation of Chinese EFL learners' AI adoption and experiences. *Computer Assisted Language Learning*, Advanced online publication. https://doi.org/10.1080/09588221.2024.2310288

Liu, S., & Kunnan, A. J. (2016). Investigating the application of Automated writing evaluation to Chinese undergraduate English majors: A case study of WriteToLearn. *The CALICO Journal*, *33*(1), 71–91. https://doi.org/10.1558/cj.v33i1.26380

Mayordomo, R. M., Espasa, A., Guasch, T., & Martínez-Melo, M. (2022). Perception of online feedback and its impact on cognitive and emotional engagement with feedback. *Education and Information Technologies*, *27*(6), 7947–7971. https://doi.org/10.1007/s10639-022-10948-2

McMartin-Miller, C. (2014). How much feedback is enough?: Instructor practices and student attitudes toward error treatment in second language writing. *Assessing Writing*, *19*, 24–35. https://doi.org/10.1016/j.asw.2013.11.003

Moore, N. S., & MacArthur, C. A. (2016). Student use of automated essay evaluation technology during revision. *Journal of Writing Research*, 8(1), 149–175. https://doi.org/10.17239/jowr-2016.08.01.05

Nicol, D., & Kushwah, L. (2023). Shifting feedback agency to students by having them write their own feedback comments. *Assessment & Evaluation in Higher Education*, 1–21. https://doi.org/10.1080/02602938.2023.2265080

Novakovich, J. (2015). Fostering critical thinking and reflection through blog-mediated peer feedback. *Journal of Computer Assisted Learning*, *32*(1), 16–30. https://doi.org/10.1111/jcal.12114

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, *3*, 121–154. https://doi.org/10.1016/j.iotcps.2023.04.003

Saeed, M. A., & Alharbi, M. A. (2023). Towards fostering Saudi EFL learners' collaborative engagement and feedback literacy in writing. *Assessing Writing*, *56*, 100721. https://doi.org/10.1016/j.asw.2023.100721

Sanosi, A. B. (2022). To err is human: Comparing human and automated corrective feedbak. *Information Technologies and Learning Tools*, *90*(4), 149–161. https://doi.org/10.33407/itlt.v90i4.4980

Shang, H. (2019). Exploring online peer feedback and automated corrective feedback on EFL writing performance. *Interactive Learning Environments*, *30*(1), 4–16. https://doi.org/10.1080/10494820.2019.1629601

Shih, R. (2011). Can Web 2.0 technology assist college students in learning English writing? Integrating Facebook and peer assessment with blended learning. *Australasian Journal of Educational Technology*, *27*(5). https://doi.org/10.14742/ajet.934

Song, Z. (2019). Investigating Chinese EFL college students' writing through the Web-Automatic Writing Evaluation Program. *English Language and Literature Studies*, *9*(3), 20. https://doi.org/10.5539/ells.v9n3p20

Wang, W. (2014). Students' perceptions of rubric-referenced peer feedback on EFL writing: A longitudinal inquiry. *Assessing Writing*, *19*, 80–96. https://doi.org/10.1016/j.asw.2013.11.008

Wei, W., Cheong, C. M., Zhu, X., & Lu, Q. (2022). Comparing self-reflection and peer feedback practices in an academic writing task: a student self-efficacy perspective. *Teaching in Higher Education*, *29*(4), 896–912. https://doi.org/10.1080/13562517.2022.2042242

Xu, L., & Zhang, T. (2022). Engaging with multiple sources of feedback in academic writing: postgraduate students' perspectives. *Assessment & Evaluation in Higher Education*, *48*(7), 995–1008. https://doi.org/10.1080/02602938.2022.2161089

Yan, Z., & Carless, D. (2021). Self-assessment is about more than self: the enabling role of feedback literacy. *Assessment & Evaluation in Higher Education*, *47*(7), 1116–1128. https://doi.org/10.1080/02602938.2021.2001431

Yang, H., Gao, C., & Shen, H. (2023). Learner interaction with, and response to, AI-programmed automated writing evaluation feedback in EFL writing: An exploratory study. *Education and Information Technologies*, *29*(4), 3837–3858. https://doi.org/10.1007/s10639-023-11991-3

Yang, L., & Liu, Y. (2017). Comparison between Peer Feedback and Automated Feedback in College English Writing: A Case Study. *Open Journal of Modern Linguistics/Open Journal of Modern Linguistics*, *07*(04), 197–215. https://doi.org/10.4236/ojml.2017.74015

Zhang, T., & Mao, Z. (2023). Exploring the development of student feedback literacy in the second language writing classroom. *Assessing Writing*, *55*, 100697. https://doi.org/10.1016/j.asw.2023.100697

Zhang, Z. V. (2020). Engaging with automated writing evaluation (AWE) feedback on L2 writing: Student perceptions and revisions. *Assessing Writing*, *43*, 100439. https://doi.org/10.1016/j.asw.2019.100439

Zhang, Z., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*, *36*, 90–102. https://doi.org/10.1016/j.asw.2018.02.004

Zheng, Y., & Yu, S. (2018). Student engagement with teacher written corrective feedback in EFL writing: A case study of Chinese lower-proficiency students. *Assessing Writing*, *37*, 13–24. https://doi.org/10.1016/j.asw.2018.03.001