



Castledown

 OPEN ACCESS

Technology in Language Teaching & Learning

ISSN 2652-1687

<https://www.castledown.com/journals/tltl/>

Technology in Language Teaching & Learning, 6(3), 1–23 (2024)

<https://doi.org/10.29140/tltl.v6n3.1482>

Comparing Peer, ChatGPT, and Teacher Corrective Feedback in EFL Writing: Students' Perceptions and Preferences



ORIT ZEEVY SOLOVEY 

Tel Aviv University, Israel

solovey@tauex.tau.ac.il

Abstract

The effectiveness of providing Written Corrective Feedback (WCF) in English as a Foreign Language (EFL) classrooms to enhance students' general writing skills has been well-established. With the advent of various feedback modes in modern classrooms, this pilot study aimed to explore EFL students' perceptions and preferences regarding the effectiveness of three distinct feedback modes: peer feedback, artificial intelligence (AI), and teacher feedback. The study also aimed to determine the distribution of feedback across specific writing components, while accounting for revisions made. Fifteen pairs of participants completed a short writing task and received WCF from peers, followed by an AI tool (ChatGPT), and finally from their teacher. Subsequent revisions to their writing were made after each feedback mode. Using a qualitative approach via survey analysis, findings indicate that participants perceived both ChatGPT and teacher feedback as effective, with peer feedback also being generally regarded as effective by most participants. Preferences lean towards teacher WCF and a combination of teacher and ChatGPT WCF. Additionally, the study identified feedback distribution across writing components, with implications discussed for refining EFL classroom feedback practices.

Keywords: Written corrective feedback (WCF); English as a foreign language (EFL); peer feedback; artificial intelligence (AI) feedback

Introduction

The provision of Written Corrective Feedback (WCF) in English as a Foreign Language (EFL) classrooms in schools and universities has been proven effective in improving students' general

Copyright: © 2024 Orit Zeevy Solovey. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. **Data Availability Statement:** All relevant data are within this paper.

writing skills. In the modern classroom, various modes of feedback are now available. Teacher feedback has long been the most traditional and commonly used method of providing WCF. It involves the teacher reviewing students' written work, identifying errors, and providing comments and corrections (Ferris & Roberts, 2001). Peer feedback emerged at a later stage in the field of WCF and involves students providing feedback on their peer's writing, promoting a collaborative learning environment (Storch, 2005). The most recent development in WCF is artificial intelligence (AI) feedback. One AI tool is ChatGPT (Chat Generative Pre-trained Transformer) which has been very popular since its launch in November 2022. This tool interacts with users conversationally by responding to questions and requests, and offers useful features related to writing assessment, including the provision of timely feedback.

The increasing availability of feedback modes in the EFL context, along with constant technological advancements, prompt teachers and researchers to examine students' views, appreciation, and preferences regarding these modes. Some students may prefer the collaborative setting of peer feedback, others may appreciate the teacher's expertise, while some may opt for an AI instant and objective evaluation. Understanding students' views and preferences is important in promoting engagement and motivation, and in laying the groundwork for potentially incorporating student choice in future writing tasks, through adopting a more student-centered approach that caters to individual views and preferences.

To the best of my knowledge, there are no studies that have explored EFL students' perceptions of peer, ChatGPT, and teacher-WCF in combination. Hence, this study seeks to fill this gap and provide insights into how EFL students perceive these feedback modes. This pilot study aims to gauge students' perceptions of the effectiveness of three different feedback modes (peer, ChatGPT, and teacher) and determine which feedback they prefer to receive. An additional aim is to identify the percentage of feedback given to specific components of the writing while also taking into account the students' revisions, in order to gain more insight into the feedback processes.

Literature Review

Collaborative writing, with its inherent practice of immediate feedback exchange among students, substantially enhances writing skills. As conceptualized by Hattie and Timperley (2007), written corrective feedback plays a major role in this process by providing awareness of various aspects of writing performance. Storch's (2005) study comparing texts written by pairs and individual learners highlights the effectiveness of collaborative writing in achieving better task outcomes, grammatical precision, and complexity. Additionally, collaborative writing fosters collective scaffolding, as termed by Donato (1988, 1994), wherein students collectively pool their language knowledge. Empirical evidence suggests that collaborative writing not only promotes consideration of form but also aids in L2 vocabulary acquisition and overall writing proficiency (Garcia Mayo & Imaz Agirre, 2019; Villarreal & Gil-Sarratea, 2019; Fernandez Dobao, 2014; Kim, 2008; Shehadeh, 2011).

In contemporary educational settings, various written corrective feedback methods are available in EFL classrooms: peer, AI, and teacher. The most traditional and commonly used is teacher feedback, which involves the teacher reviewing students' written work, identifying errors, and providing comments and corrections (Ferris & Roberts, 2001). Teacher-led feedback is critical in the revision process for student writers (Sheen et al., 2009). According to Bitchener and Ferris (2012), teachers' WCF, such as edge comments, requests for clarification, and comments on grammatical issues, has been proven to help EFL writers. Teacher feedback is valued for offering expert guidance and instructions tailored to meet individual student needs (Lyster, 1998). It has been suggested that effective teacher feedback

significantly improves students' writing skills, increases motivation, and fosters a positive learning experience (Ferris & Roberts, 2001).

Peer feedback, a more recent advancement in the field, involves students offering feedback on each other's writing, fostering a collaborative learning environment. It has arisen as an innovative method that differs from traditional teachers' WCF (Sheen, 2010b) since teachers traditionally play a dominant role in providing feedback and comments during the writing process (Sheen, 2010a). Peer feedback has been increasingly examined in the literature (Hewett, 2000; Liu & Hansen, 2002) and has demonstrated numerous benefits. Several studies (Connor & Asenavage, 1994; Cho & MacArthur, 2010; Storch, 2005; Topping, 1998) have concluded that peer feedback supports the enhancement of learners' writing capabilities and competencies concerning cognitive, affective, social, and linguistic perspectives. Some researchers (for example, Sato [2013], Sato & Lyster [2012], and Sippel & Jackson [2015]) claim that students' engagement with peer feedback can promote second language learning and can change their role from passive learners to active participants. Peer interactions allow students to engage in active analysis of others' writing, which can lead to a deeper understanding of their own writing weaknesses (Storch, 2005). Peterson and Portier (2014) maintain that peer feedback benefits not only the students, but also the feedback providers. It also helps students to develop critical thinking and language analysis skills (Ferris & Roberts, 2001). As students gain experience using the given criteria to evaluate writing, their self-assessment abilities progress. (Lundstrom & Baker, 2009). Therefore, peer feedback is considered an effective method to improve students' writing skills (Xiao & Lucking, 2008).

Nevertheless, some researchers (for example, Adams et al [2011], Philp et al [2010], and Zhang [1995]) doubt the pedagogic potential of peer feedback due to factors such as the learners' distrust of their own or their peers' linguistic knowledge. Ruegg (2015) found that teacher feedback was more effective in improving grammatical errors than peer feedback. Some studies that compared teacher and peer feedback have shown that student writers prefer teachers' comments over those of their peers, while other studies suggest that peer feedback is beneficial for the construction of the writing process (Yang et al., 2006). It is suggested that EFL learners can effectively take advantage of both methods to improve their writing skills, Tai et al., (2015) compared the effects of the combination of teacher and peer feedback versus using teacher feedback alone on the writing performance of EFL university students within a collaborative online learning system. The results revealed that, in terms of holistic writing skills and the subscales of content, organization, grammar, mechanics, and style, the students in the combined feedback group demonstrated greater improvements than those who received only teacher feedback. Yet, due to the students' limited linguistic knowledge and writing skills, they were able to indicate errors from a reader's perspective, but were less able to supply consistent comprehensive and persuasive feedback. Overall, the students viewed peer review as a demanding task, albeit helpful. Lee (2008) suggests that it is essential to provide proper guidance to peers to ensure accurate error identification and constructive feedback.

The most recent development in the field of WCF is AI feedback. Artificial Intelligence has witnessed significant progress in recent years, and the use of AI-based technology is becoming increasingly popular in the field of education. There are multiple ways of applying AI-based tools in student learning environments. Steiss et al., (2024) maintain that AI may function as an automated writing evaluation tool, increasing the amount of feedback students receive and diminishing the burden on teachers to provide frequent feedback to large classes. According to Wu (2024), AI technology has made significant strides in vocabulary acquisition, grammar correction, and discourse generation. AI tools can offer assistance both during and after the writing process. AI technology not only examines grammar and spelling but also offers extensive support in identifying writing problems and suggesting

improvements (Alharbi, 2023). One AI tool is ChatGPT (Chat Generative Pre-trained Transformer), a chatbot that has been very popular since its launch in November 2022 by the United States-based startup OpenAI. ChatGPT interacts with users conversationally, responding to questions and requests in a dialogic manner (OpenAI, 2023). This tool, powered by a large language model, can understand and produce human-like responses across a wide range of topics by computing techniques and a large amount of information to connect ideas and understand prompts in context (Barrot, 2023). Wenzlaff and Spaeth (2022) affirm that ChatGPT is essentially equivalent to humans in composing explanatory answers.

ChatGPT can be a helpful writing tool in the second language classroom by addressing some of the writers' needs and offering useful features related to writing assessment, including the provision of timely feedback. By prompting it to give feedback, ChatGPT identifies the strengths and weaknesses of a text and offers recommendations for improvement. It can also evaluate the clarity, focus, and structure of a text (Barrot, 2023). Several studies have examined the effectiveness of employing ChatGPT as a feedback tool, resulting in diverse findings (Steiss et al., 2024; Wu, 2024). Specifically, a study conducted by Dai et al. (2023) investigated the feasibility of utilizing ChatGPT to offer students constructive feedback aimed at enhancing their learning experience. The findings demonstrated that ChatGPT consistently produces more readable feedback, aiding students in better comprehension and subsequent improvement efforts. However, it was noted that ChatGPT lacks the reliability seen in instructor assessments of student performance. Furthermore, ChatGPT exhibited a notable capability to generate process-focused feedback, which was deemed more effective in shaping students' task strategies than task-focused feedback. The overall conclusion underscores ChatGPT's significant potential in guiding students to enhance their skills and foster the development of learning skills.

Additionally, Su et al. (2023) explored the potential applications of ChatGPT in supporting students across various tasks, including outline preparation, content revision, and proofreading. By providing ChatGPT with an argumentative writing outline and an evaluation rubric, the researchers observed its ability to assess the coherence between claims and sub-claims, evaluate the quality of supporting evidence and rebuttals, and suggest potential counterarguments and claims. This approach has the potential to enhance the structural aspect of the argumentation. In the editing phase, ChatGPT received evaluation rubrics for argumentative writing content. The feedback generated accurately pinpointed the strengths and weaknesses of the writing, offering valuable suggestions for improvement. The researchers acknowledged, however, that due to ChatGPT's limited generative capacity, some suggestions were somewhat vague, with abstract comments, necessitating further elaboration. During the proofreading stage, an evaluation checklist addressing the use of lexical devices, syntactic structures, and textual devices was provided to ChatGPT. While the feedback primarily took an evaluative stance, it lacked sufficient elaboration, limiting its potential for facilitating learning. Furthermore, the consistency of ChatGPT's feedback was noted to be an issue, as it yielded different responses upon each submission of student writing, diminishing the feedback's overall usefulness. Despite these limitations, the researchers highlighted ChatGPT's efficacy as a proofreading aid. When employed to assess individual sentences for grammar, ChatGPT demonstrated the ability to evaluate grammatical accuracy, elucidate the intended meaning of the sentence, and suggest alternative expressions. However, the researchers state that the responsibility rests with students to critically analyze and determine whether to accept the revisions proposed by ChatGPT.

Furthermore, Yoon et al., (2023) evaluated the quality of the feedback produced by ChatGPT as to the coherence and cohesion of essays written by English language learners. They used 50 argumentative essays to be assessed utilizing a rubric. A two-step approach was employed in the

feedback evaluation. First, each sentence in the feedback was categorized into subtypes based on its function (for example, positive reinforcement and problem statement). Then, each sentence in the feedback was evaluated for accuracy and usability in the feedback according to these types. Both the analysis of feedback types and the evaluation of accuracy and usability revealed that most feedback sentences were very abstract, generic, and did not provide concrete suggestions for improvement. The accuracy in detecting repetitive ideas and erroneous use of cohesive devices depended on superficial linguistic features and was often incorrect. They concluded that ChatGPT, without specific training for the feedback generation task, did not offer effective feedback on coherence and cohesion in these students' essays.

Upon receiving feedback of any nature, students stand to benefit from revising their writing. As part of a process-orientated pedagogy, teachers have encouraged their students to write multiple drafts of their papers and have explored various feedback methods to help them improve at each stage of the writing process. (Ferris, 1997). Beason (1993) states that "feedback and revision are valuable pedagogical tools...the research typically indicates that high school and college students improve their drafts upon receiving feedback" (p. 396). WCF aims to teach skills that help students improve their writing proficiency and produce written texts with fewer errors and more clarity (Williams, 2003). The integration of collaborative writing with multiple feedback modes (peer, ChatGPT, and teacher) may offer a rich context for understanding students' perceptions and preferences regarding the different feedback sources.

Methods

The purpose of this pilot study was to gauge students' perceptions regarding the effectiveness of the different feedback modes and to find out which feedback they prefer to receive. An additional aim was to identify the percentage of feedback given to specific components of the writing while considering the revisions performed by students, to gain more insight into the feedback processes. This study was guided by the following three research questions:

- RQ 1: Do students perceive peer, ChatGPT, and teacher WCF as effective?
- RQ 2: What WCF (peer, ChatGPT, teacher) do students prefer to receive?
- RQ 3: What percentage of the WCF is given to specific components of the writing, by peers, ChatGPT, and the teacher?

A university class of 30 freshmen EFL students participated in the study. The students were enrolled in a B2 (upper-intermediate) level course. This is the university's most advanced compulsory English course geared toward improving English reading, writing, and other academic skills to enable students to function competently in the academic and professional world. There were 8 males and 22 females with an age range of 19–26.

The strategy of collaborative writing (Elola & Oskoz, 2010), process approaches, and multiple revisions (Hyland, 2003) were adopted for this study. Students were randomly divided into 15 pairs and were assigned a writing task during which they had to use a checklist to verify that all the writing requirements had been met. Although the assignment was not graded, "The Composition Grading Feedback Sheet" developed by Brown and Bailey (1984) was used in this study. It was adapted to serve as a checklist, as illustrated in Table 8, which consisted of five criteria: organization, content, grammar, mechanics, and quality of expression (see Appendix A). The checklist items align with the writing instruction provided throughout the course, and participants used this same checklist to confirm adherence to writing requirements on two prior occasions preceding the study.

The study's sequencing of feedback modes (peer-ChatGPT-teacher) was strategically devised to optimize feedback effectiveness, considering practical factors such as time constraints, varying familiarity, and the promptness of each mode. The rationale behind this order rests on several key considerations. Firstly, recognizing the inherent time limitations of teacher feedback compared to peer and ChatGPT feedback, the decision was made to schedule teacher feedback last. This ensured students received timely feedback from their peers and ChatGPT before the more time-consuming teacher feedback process. Secondly, the sequence considered students' familiarity with receiving peer feedback versus feedback from an AI tool. By initiating with peer feedback, followed by ChatGPT feedback, the aim was to gradually introduce students to the less familiar mode of automated feedback while leveraging their existing comfort with peer evaluation. Thirdly, prioritizing peer and ChatGPT feedback in the initial rounds allowed for faster turnaround times, enabling students to promptly receive feedback on their written work. This highlights the importance of timely feedback in supporting student learning and improvement. Lastly, while acknowledging the potential for order effects introduced by the sequencing of feedback modes, consistency in evaluation criteria across all feedback modes was ensured. Participants were also briefed on the rationale behind the chosen feedback sequence to maintain transparency and understanding.

During the initial round of feedback, each pair was asked to use the checklist to review another pair's writing and provide feedback on areas that required correction. On the checklist, they indicated whether each writing criterion was fulfilled, provided comments on areas requiring correction, and proposed alternative options. Subsequently, each pair had to decide whether they agreed with their peers' comments and make necessary revisions to their work. Next, each pair had to input their revised work into ChatGPT, which was prompted by the checklist, and was instructed to provide feedback as to what needs to be altered. Students received feedback from ChatGPT on the fulfillment of the writing criteria and copied ChatGPT's comments and suggestions into the checklist. Students then assessed this feedback, decided whether to accept it, and proceeded to modify their writing accordingly. Finally, each pair submitted their modified work to the teacher who used the same checklist to identify areas for correction and provide suggestions for improvement. Students had to determine whether they agreed with the teacher's feedback and revise their work. Following these feedback cycles, participants completed a feedback survey (see Appendix B).

The study employed a qualitative approach to explore students' perceptions and preferences regarding different feedback modes and to identify the distribution of feedback across specific components of the writing. To address these objectives, data was gathered using two methods. First, a checklist was employed after each feedback round, necessitating an assessment of writing criteria fulfillment as indicated by peers, ChatGPT, and the teacher. Participants were also directed to copy ChatGPT's comments and suggestions into the checklist. Second, a feedback survey was administered upon the completion of the entire feedback process. The survey addressed relevant aspects of the research objectives, incorporating feedback from multiple sources (peers, ChatGPT, and the teacher). This diversity enhances credibility by providing different perspectives on the effectiveness of various feedback modes, thereby strengthening the findings. Participants evaluated the effectiveness of each mode, confirming the accuracy of their responses and further enhancing credibility. Reflexivity is demonstrated in the rationale behind the feedback modes sequencing, acknowledging potential biases, and study decisions. Additionally, the uniform application of the same checklist by peers, ChatGPT, and the teacher throughout the study reinforces dependability by ensuring transparency and consistency. The detailed explanation of the feedback sequence and the participants' briefing promotes transparency, enabling readers to assess the reliability of the findings. The checklists along with the survey files were uploaded to the course Moodle site by the participants. The researcher later retrieved these files from Moodle and employed content analysis to organize and analyze the data.

Results

The qualitative approach used in this study involved a systematic content analysis of survey data to explore students' perceptions and preferences regarding different modes of written corrective feedback (WCF) – peer, ChatGPT, and teacher feedback. The survey data were divided into three sets, each corresponding to one of the research questions. Five primary codes (A–E) were assigned to organize and analyze the data: A. Types of corrections required, B. Understanding of comments, C. Agreement with comments, D. Perceived helpfulness of comments and reasons for this perception, E. Feedback preference and reasons for this preference.

Categories B–D (Understanding, Agreement, Perceived Helpfulness) addressed the first research question concerning the effectiveness of peer, ChatGPT, and teacher feedback. Understanding the feedback was considered crucial for its effective application. Agreement with feedback reflected students' judgments regarding its validity and relevance, while perceived helpfulness indicated the utility of the provided corrections. Category E (Feedback Preference) was related to the second research question, investigating students' preferences regarding the source of feedback (peer, ChatGPT, or teacher). Category A (Types of Corrections Required) pertained to the third research question, focusing on the distribution of feedback across specific components of writing. It aimed to identify the percentage of WCF directed towards different writing components by peers, ChatGPT, and teachers.

Categories B–D were further subdivided to include responses specific to each feedback mode (peer, ChatGPT, teacher), enabling a detailed analysis of feedback perceptions across different sources. Within each category, all responses to specific questions were gathered to identify different aspects related to feedback. This approach facilitated the identification of prevalent themes or opinions within the data. The frequency of specific types of responses within each category was counted to quantify the prevalence of particular themes or opinions.

Peer Feedback

In the study, a total of 15 pairs provided feedback on various aspects of their peers' work. Respondents were asked about the specific components they believed required correction, and the responses were analyzed. As displayed in Table 1, the most commonly identified areas for improvement were organization and mechanics, with 33.33% of respondents (five pairs) mentioning it, followed by organization alone, which was mentioned by 26.67% of the participants (four pairs). Content and mechanics, as well as content alone, were each mentioned by 13.33% of the respondents (two pairs each). A smaller proportion of participants, 6.67% each, identified organization and grammar, and organization and quality of expression as areas requiring correction (one pair each). Regarding the understanding of comments, all 15 pairs (100%) reported understanding them entirely. When asked whether they agreed with or accepted the comments, the majority, 66.67%, (ten pairs) responded affirmatively. However, 20% of the respondents (three pairs) expressed disagreement with the comments, indicating, for instance, that "It seemed incorrect to us." Two pairs, 13.33% expressed partial agreement, noting, for example, that "We agreed with comments about punctuation, but did not agree that there are not enough explanations/exemplifications of the ideas. We think there are enough." Concerning the perceived helpfulness of the feedback, 73.33% of the respondents (eleven pairs) found it helpful, specifying that "It helped us change the focus of the writing" and "It helped to improve the organization". However, 26.67% (four pairs) did not find the feedback helpful, since, for instance, their peers "...made some comments, but then we read our text several more times and found mistakes that they didn't find".

Table 1 *Summary of Students' Perceptions of Peer Feedback*

Questions:	Responses:
What did your peers think required correction?	33.33% – Organization and mechanics 26.67% – Organization 13.33% – Content and mechanics 13.33% – Content 6.67% – Organization and grammar 6.67% – Organization and quality of expression
Did you understand all the comments?	100% – Yes
Did you agree with/accept the comments? If not, why?	66.67% – Yes 20% – No (“It seemed incorrect to us.”) 13.33% – Partially (“We agreed with comments about punctuation, but did not agree that there are not enough explanations/exemplifications of the ideas. We think there are enough.”)
Did you find the feedback helpful? Why?	73.33% – Yes (“It helped us change the focus of the writing”; “It helped to improve the organization.”) 26.67% – No (“They made some comments, but then we read our text several more times and found mistakes that they didn't find.”)

Note. The study aimed to explore EFL students' perceptions of the effectiveness of peer, ChatGPT, and teacher feedback, and uncover students' preferences. Data analysis results demonstrate that most participants regarded peer feedback as effective. An additional aim was to identify the percentage of the feedback given to specific components of the writing. It was found that peers mostly commented on the organization and mechanics of the writing.

ChatGPT Feedback

As shown in Table 2, ChatGPT provided feedback to 15 pairs suggesting various aspects that required correction. Specifically, 20% of the participants (three pairs) received suggestions related to all categories (organization, content, grammar, mechanics, and quality of expression.) Also, 20% of the participants received feedback related to the content, quality of expression, and grammar. Another 20% of participants were given feedback regarding mechanics and grammar. Additionally, 13.33% of participants (two pairs) received suggestions for organization, mechanics, and grammar, while the same percentage received feedback on organization, content, and grammar. Lastly, 13.33% of participants were provided feedback on content, mechanics, and grammar.

The following instances exemplify the participants' use of ChatGPT and the ensuing feedback. In terms of the writing's organization, students inquired whether a thesis statement was present. Responses varied from affirmations like “Yes, there is a thesis sentence in the provided passage”, to more detailed suggestions when a thesis statement was perceived to be absent. For instance:

Your paragraph does not explicitly state a thesis statement, but the central idea of the paragraph is that To turn the central idea of the paragraph into a thesis statement, you could write something like

In cases where a thesis statement could not be identified, the response indicated:

Based on the information provided, it is unclear what the thesis sentence of your essay is. A thesis sentence is a statement that presents the main idea or argument of the essay and is typically included in the introduction paragraph.

Table 2 *Summary of Students' Perceptions of ChatGPT Feedback*

Questions:	Responses:
What did ChatGPT suggest correcting?	20% – All categories 20% – Content, quality of expression, and grammar 20% – Mechanics and grammar 13.33% – Organization, mechanics, and grammar 13.33% – Organization, content, and grammar 13.33% – Content, mechanics, and grammar
Did you understand all the comments?	93.33% – Yes 6.67% – Not entirely (“We didn’t understand why the grammar is not correct”)
Did you agree with/ accept the comments? If not, why?	100% – Yes
Did you find the feedback helpful? Why?	100% – Yes (“We saw big differences between our essay compared to the ChatGPT essay. We could see where we were wrong and why”; “ChatGPT corrected the essay very fast and phrased everything better”.)

Note. Data analysis results demonstrate that all participants perceived ChatGPT feedback as effective. It was also revealed that an equal percentage of ChatGPT's comments related to all aspects of writing; content, quality of expression, and grammar; and mechanics and grammar.

Concerning the understanding of the comments, 93.33% (fourteen pairs) indicated that they comprehended the suggestions made by ChatGPT, yet, 6.67% (one pair) reported not entirely understanding the comments, “We didn’t understand why the grammar is not correct”. All participants (100%) reported that they agreed with and accepted the suggestions provided by ChatGPT, and that they thought the feedback was helpful. Participants claimed, “We saw big differences between our essay compared to the ChatGPT essay. We could see where we were wrong and why” and “ChatGPT corrected the essay very fast and phrased everything better”.

Teacher Feedback

As demonstrated in Table 3, analysis of the responses revealed that the most common area for improvement was mechanics and grammar, as reported by 26.67% of the respondents (four pairs). Additionally, 20% of the participants (three pairs) were given feedback related to content, mechanics, and grammar, while another 13.33% (two pairs) reported that the teacher identified organization, content, and grammar as areas in need of corrections. An equal percentage of participants (13.33%) were provided improvement suggestions about organization, content, and quality of expression. Moreover, 13.33% of the participants received feedback on the quality of expression. A smaller proportion of participants, 6.67% each (one pair), was given feedback related to the quality of expression, mechanics, and grammar, and some participants (6.67%) were provided feedback across all categories. All participants (100%) reported that they comprehended and accepted the suggestions provided by the teachers, and they all perceived the feedback as helpful. They thought, for example, that “The teacher corrected the mistakes and gave explanations about every mistake, so we know why it’s a mistake”.

Feedback Preferences

As presented in Table 4, the results indicated that 46.67% of the participants (seven pairs) preferred receiving feedback solely from the teacher. They noted that “The teacher is the expert”, and that they “trust teacher feedback the most and can learn from it”. Additionally, 33.33% of the participants

Table 3 *Summary of Students' Perceptions of Teacher Feedback*

Questions:	Responses:
What did the teacher think required correction?	26.67% – Mechanics and grammar 20% – Content, mechanics, and grammar 13.33% – Organization, content, and grammar 13.33% – Organization, content, and quality of expression 13.33% – Quality of expression 6.67% – Quality of expression, mechanics, and grammar 6.67% – All categories
Did you understand all the comments?	100% – Yes
Did you agree with/ accept the comments? If not, why?	100% – Yes
Did you find the feedback helpful? Why?	100% – Yes (“The teacher corrected the mistakes and gave explanations about every mistake, so we know why it's a mistake”.)

Note. Data analysis results demonstrate that all participants perceived teacher feedback as effective. Additionally, it was observed that the primary emphasis of teacher feedback was on mechanics and grammar, with a subsequent focus on content, mechanics, and grammar in that order.

Table 4 *Summary of Students' Feedback Preferences*

Questions:	Responses:
What feedback or feedback combination do you prefer to receive? Why?	46.67% – Teacher feedback (“The teacher is the expert.”; “We trust teacher feedback the most and can learn from it.”) 33.33% – Teacher and ChatGPT feedback (“Because ChatGPT helped us correct many things in the essay, but then we submitted it to the teacher who noticed more things we need to correct.”; “The teacher is the most reliable source we can get feedback from, but if we want a quick answer when the teacher is not immediately available we can use GPT.”) 6.67% – ChatGPT feedback (“Chat GPT is quick and helpful.”; “GPT instructions and corrections were the clearest.”; “It also gives positive feedback.”) 6.67% – Peer and teacher feedback (“We don't believe in technology, and prefer to get feedback from peers and teachers. We can ask them questions about things we don't understand.”) 6.67% – Peer feedback (“The opinion of peers gives an indication if students like us understand what we wrote, and they know how to give us feedback”.)

Note. The results indicate a predominant preference among participants for teacher feedback, with a subsequent inclination towards a blend of teacher and ChatGPT feedback. Conversely, a lesser preference was observed for peer feedback, a combination of peer and teacher feedback, and feedback solely from ChatGPT.

(five pairs) expressed a preference for receiving feedback from both their teacher and ChatGPT, for the following reasons, “Because ChatGPT helped us correct many things in the essay, but then we submitted it to the teacher who noticed more things we need to correct”, and;

The teacher is the most reliable source we can get feedback from, but if we want a quick answer when the teacher is not immediately available, we can use GPT.

One pair, 6.67%, indicated a preference for receiving feedback only from ChatGPT because it is “... quick and helpful”, its “...instructions and corrections were the clearest”, and “It also gives positive

feedback”. Another pair, 6.67%, preferred a combination of feedback from peers and the teacher. They indicated, “We don’t believe in technology, and prefer to get feedback from peers and teachers. We can ask them questions about things we don’t understand”. Finally, one pair, 6.67%, preferred receiving feedback exclusively from their peers, stating that “The opinion of peers gives an indication if students like us understand what we wrote, and they know how to give us feedback”.

Discussion

This pilot study sought to gauge students’ perceptions regarding the effectiveness of three different feedback modes (peer, ChatGPT, and teacher), and find out which feedback they prefer to receive. It also attempted to identify the percentage of feedback given to specific components of the writing, while considering the revisions performed, to gain more insight into the feedback processes. The first research question addressed whether students perceive peer, ChatGPT, and teacher WCF as effective. According to the findings, all the participants (15 pairs) acknowledged comprehension of their peers’ comments, with the majority (10 pairs), agreeing with and accepting the comments. Yet, 3 pairs expressed disagreement with the comments, while 2 pairs indicated partial agreement. Most respondents (11 pairs) found their peer’s feedback to be helpful, though 4 pairs did not. Regarding the feedback provided by ChatGPT, 14 pairs indicated full comprehension of the comments and suggestions, while only 1 pair reported partial understanding. All participants noted that they agreed with and accepted the suggestions provided by ChatGPT, deeming the feedback helpful. In the context of teacher feedback, all participants affirmed both their understanding of the teacher’s suggestions and their acceptance of them. They all perceived the teacher’s feedback as helpful.

The second research question explored students’ preference for WCF from peers, ChatGPT, and the teacher. The findings reveal that the majority (46.67%) favor teacher-provided WCF, and 33.33% prefer receiving a combination of teacher and ChatGPT WCF. An equal percentage of participants (6.67% each) indicated a preference for peer feedback; a combination of peer and teacher feedback; and ChatGPT feedback, respectively. Several observations were noted upon inspecting the results of the first and second research questions. Firstly, it was observed that while the majority of respondents (11 pairs) perceived their peers’ feedback as beneficial, only 1 pair expressed a preference for getting peer feedback alone, and merely 1 pair favored a combination of peer and teacher feedback. This aligns with Colpitts’ (2016) study that examined student perceptions of receiving teacher corrective feedback, peer corrective feedback, and giving peer corrective feedback. The results showed a considerable preference for receiving teacher feedback. Peer feedback has also been challenged by Lin and Yang (2011), who maintain that the quality of student comments may not be sufficient, and interpersonal relationships may affect learner intentions and the provision of honest comments. Nevertheless, the findings diverge from the conclusions drawn in second language writing research (for example, Hedgcock and Lefkowitz [1992] and Paulus [1999]), which suggest that peer feedback comments have the potential to lead to substantial revisions. Additionally, these studies propose that, in comparison to teacher feedback, revisions informed by peer comments may exhibit improvements in vocabulary, organization, and content.

In the present investigation, students’ reduced inclination toward peer feedback might be attributed to inadequate training in feedback provision. Students in this study were equipped only with a checklist for providing feedback. According to Hansen and Liu (2005), who devised adaptable guidelines for writing instructors in the EFL context, the key to fostering a positive reception of peer feedback lies in meticulous teacher planning and comprehensive student training. They argue that planning and training should be continuous processes, spanning the periods before, during, and after peer response, although the majority should be carried out before peer response commences. Modeling the feedback

process and gradually developing feedback skills throughout the course are essential steps, ensuring that peer feedback motivates students to actively engage in their learning (Ferris & Hedgcock, 2005; Kamimura, 2006).

Secondly, it was noted that while all respondents deemed ChatGPT's feedback helpful, only 1 pair indicated a preference for receiving ChatGPT feedback alone, and 5 pairs expressed a preference for a combination of teacher and ChatGPT feedback. This corresponds with Dai et al.'s (2023) findings that ChatGPT consistently produces readable feedback, aiding students in better comprehension and subsequent improvement efforts, but lacks the reliability seen in instructor assessments of student performance. It could be that students appreciate ChatGPT's immediate feedback, yet sense its incapability of "understanding" the text it generates or the context of the information. Scholars have pointed out that despite their powerful synthesis ability, "chatbots merely learn statistical associations between words...rather than understand their meanings" (van Dis et al., 2023, p. 615). Nevertheless, the results of the current study contrast with some of the findings of Su, Lin, and Lai (2023) that revealed ChatGPT's ability to assess the coherence between claims and sub-claims, evaluate the quality of supporting evidence and rebuttals, and suggest potential counterarguments and claims in an argumentative writing outline.

In the present study, a notable illustration of ChatGPT serving as an effective feedback tool is its provision of explanations, specific corrections, or alternatives to students' work in numerous instances. For example, when a concluding sentence was absent in a piece of writing, ChatGPT provided the following response, as can be seen in Table 5:

Overall, it seems that your paragraph is missing a clear concluding sentence. A concluding sentence is important because it helps to summarize the main points of the paragraph and bring closure to the argument you have made. Here's an example of a concluding sentence that you could use for your paragraph: ... This sentence reiterates the main idea of your paragraph and emphasizes...

Additionally, in cases where vocabulary errors were detected, ChatGPT offered targeted corrections, such as:

For the most part, the vocabulary is correct. However, there are a few minor errors that could be corrected: In the first sentence, "people get confused" could be changed to "people become confused" for consistency with the tense of the rest of

Table 5 ChatGPT's Comment on the Concluding Sentence in One Pair's Writing

Concluding sentence	Overall, it seems that your paragraph is missing a clear concluding sentence. A concluding sentence is important because it helps to summarize the main points of the paragraph and bring closure to the argument you have made. Here's an example of a concluding sentence that you could use for your paragraph: "In conclusion, having almost unlimited options in life can provide us with the freedom to make choices that align with our personal values and goals, enabling us to live happier and more fulfilling lives." This sentence reiterates the main idea of your paragraph and emphasizes the benefits of having almost unlimited options in life.
---------------------	--

Note. During the second round of feedback, each pair input their work into ChatGPT which was prompted by a checklist, and was instructed to provide feedback as to what needs to be altered. Students received feedback from ChatGPT on the fulfillment of writing criteria and copied ChatGPT's comments and suggestions into the checklist.

the text...” or: “In the first sentence, it should be “a” instead of “some” ...In the second sentence, it should be “real” instead of “the real” ..., In the fourth sentence, it should be “their” instead of “there”...and “avoid choosing” instead of “avoid from the choose.

However, this study identified a less effective use of ChatGPT for feedback, demonstrated by the observed inconsistency of responses, as depicted in Table 6 and Table 7. For instance, when different students sought clarification on punctuation correctness, some received responses such as:

As an AI language model, I cannot say for sure that there are no punctuation errors in the essay, as punctuation can be subjective and there are many different styles and conventions that may be used. However, based on my analysis, the essay appears to have good punctuation overall.

Conversely, others were given responses like:

Yes, the punctuation in the essay is generally correct. There are a few minor punctuation errors, such as missing commas and a misspelled word, but they do not significantly affect the clarity of the text.

Moreover, some students received very specific answers like:

“In the first sentence, there should be a comma after “these days” ...In the second main idea sentence, there should be a comma after “study” ...In the third main idea sentence, there should be a comma after “anywhere” ...”

Table 6 ChatGPT’s Comment on the Clarity of Sentences in One Pair’s Writing

All the sentences are clear	As an AI language model, I cannot definitively say that all of the sentences are clear and do not need any improvement, as language is a complex and nuanced subject. However, based on my analysis, the sentences in the essay appear to be well-written and communicate their intended meaning effectively.
-----------------------------	---

Note. Students received feedback from ChatGPT on the fulfillment of writing criteria and copied ChatGPT’s comments and suggestions into the checklist.

Table 7 ChatGPT’s Comment on the Clarity of Sentences in Another Pair’s Writing

All the sentences are clear	The sentences are generally clear, but there are a few areas where the wording could be improved for greater clarity. For example, the second sentence in the paragraph could be revised to say, “Having too many options can be overwhelming and lead to mental exhaustion.” This would clarify the cause-and-effect relationship between having multiple choices and feeling overwhelmed.
-----------------------------	---

Note. Students received feedback from ChatGPT on the fulfillment of writing criteria and copied ChatGPT’s comments and suggestions into the checklist.

Similarly, when students inquired about sentence clarity, responses varied. Some received answers such as:

As an AI language model, I cannot definitively say that all of the sentences are clear and do not need any improvement, as language is a complex and nuanced subject. However, based on my analysis, the sentences in the essay appear to be well-written and communicate their intended meaning effectively.

On the other hand, others were provided with responses like:

The sentences are generally clear; but there are a few areas where the wording could be improved for greater clarity. For example, the second sentence in the paragraph could be revised to say... This would clarify the cause-and-effect relationship between...and..."

Su, Lin, and Lai (2023) also noted a lack of consistency in ChatGPT's feedback, with varied responses emerging upon each submission of student writing. This implies a necessity for multiple inputs into ChatGPT, with subsequent examination of the diverse responses to better understand and address potential issues.

Additionally, it was observed that most students favor teacher-provided WCF. This inclination towards teacher WCF might stem from the traditional educational framework students are accustomed to, wherein instructors traditionally hold the authority to assess and offer feedback on students' work. It may be inherently linked to students' trust in and reliance on feedback from teachers, who are perceived to possess expertise. Such a preference may also be reinforced by the established norms and expectations within the educational environment.

The tendency to favor teacher WCF also aligns with socio-constructivist theories, which emphasize the collaborative nature of teaching and learning and the role of social interactions in knowledge construction (Vygotsky & Cole, 1978; Bruner, 1985; Rogoff, 1990; Lave & Wenger, 1990). According to Vygotsky and Cole's (1978) Zone of Proximal Development, learning occurs within the "zone" where learners engage in collaborative activities with more knowledgeable instructors. With the help of an instructor, students can understand and master knowledge and skills that they would not be able to on their own (Schreiber & Valle, 2013). Once the students master a certain skill they can complete it independently. In this theory, the instructor plays an integral role in students' acquisition of knowledge (Chen, 2015; Schreiber & Valle, 2013).

The third research question addressed the distribution of WCF across specific aspects of writing provided by peers, ChatGPT, and the teacher. The results indicate that within peer feedback, the areas for improvement were primarily organization and mechanics, with organization alone being next. Subsequently, attention was directed toward content and mechanics alongside content alone. Following these, feedback addressed organization and grammar, as well as organization and quality of expression. Hence, it was reasonable to anticipate that participants' revisions would primarily address the organization of their writing, with some attention to content and mechanics, while placing relatively less emphasis on grammar and quality of expression. Lack of sufficient linguistic knowledge among reviewers may contribute to the inaccurate identification of linguistic errors. This concern is evident in criticisms challenging the validity of feedback associated with peer review in EFL classrooms, especially when reviewers have lower language proficiency (Covill, 2010; Guenette, 2007).

After the initial round of revisions and inputting the text into ChatGPT for feedback, a smaller number of pairs received suggestions to work on organization. This corresponds with the prior revisions

made by the participants. The areas marked for improvement in ChatGPT feedback included various recommendations. Some pairs were advised to pay attention to all aspects of their writing, while others were directed to focus on content, quality of expression, and grammar. Certain pairs were guided to address issues related to mechanics and grammar. Fewer pairs received guidance to address either organization, mechanics, and grammar; organization, content, and grammar; or content, mechanics, and grammar. Notably, ChatGPT placed less emphasis on the necessity for enhancing the quality of expression, while it precisely identified mechanics, content, and notably, grammar as areas warranting attention. Given the inherently abstract nature of expression quality, this observation suggests a potential requirement for refining the prompts employed in assessing expression quality. As noted by Atlas (2023), language models like ChatGPT may deviate from the intended direction if prompts or questions are not carefully worded. Thus, it is crucial to formulate clear and detailed prompts, enabling the model to precisely comprehend the desired task. Another important principle is to encourage ChatGPT to elaborate on its responses by providing feedback and requesting further information.

Following the second round of revisions and the submission of the revised work for evaluation by the teacher, the feedback provided by the teacher highlighted specific areas requiring improvement. The primary focus centered on mechanics and grammar, followed by content, mechanics, and grammar. Additionally, the teacher addressed the following as requiring corrections: organization, content, and grammar; organization, content, and the quality of expression; and quality of expression alone. In fewer instances, pairs received feedback concerning the quality of expression, mechanics, and grammar; as well as feedback spanning all writing categories. Much like ChatGPT's assessment, the teacher's evaluation emphasized the need to address mechanics, content, and, notably, grammar. Furthermore, it also identified the quality of expression as an area that warranted attention and improvement.

Despite technological advances and the potential of ChatGPT to function as a written corrective feedback tool, it is worth noting that even after students received feedback from the tool and made subsequent revisions, the teacher still identified errors in their writing. As some participants noted, "GPT helped us a lot with making the writing good but the teacher paid attention to things that we and GPT didn't see." Teachers have the expertise, experience, and real-world context that contribute to the identification of writing errors and the provision of appropriate feedback. According to Atlas (2023), although ChatGPT and other language models prove beneficial in proofreading and editing student writing, they should not be viewed as replacements for human editing and feedback. While these tools can help to identify errors and issues, they may not always be able to understand the context and intent of the writing. Consequently, it is important to employ ChatGPT and other language models in conjunction with human editing and feedback to guarantee the most accurate and effective results.

Implications for Practice and Future Research

The study's findings regarding the perceived value of ChatGPT's feedback among participants have implications for the EFL classroom. Drawing from autonomous learning theory, as explained by Little (1991), which emphasizes the importance of learners taking ownership of their learning process and actively seeking feedback to enhance their knowledge, it becomes evident that integrating peer and AI-based feedback can foster learner autonomy. These forms of feedback offer immediate and accessible support for self-directed learning. However, effective autonomous learning necessitates learners' ability to critically evaluate and integrate feedback into their learning strategies, a process that may benefit from instructors' guidance and scaffolding, as discussed by Benson (2013). Therefore, recognizing the potential of ChatGPT's feedback, teachers could consider its incorporation into the language learning process to further empower students in their autonomous learning efforts. This integration would necessitate training students to formulate clear and detailed prompts, ensuring the

model precisely comprehends the task at hand. Moreover, encouraging students to prompt ChatGPT to elaborate on its responses can enhance the depth of feedback. Additionally, as proposed by Atlas (2023), teachers can train a ChatGPT model using a dataset of student essays previously graded and corrected by the teacher. This trained model can then analyze new essays, flagging errors for the teacher's review and subsequent feedback to the student. Yet, it is important to consider the potential inconsistency in responses from ChatGPT. This underscores the necessity for multiple inputs into ChatGPT and subsequent examination of diverse responses to better understand and address potential issues.

Rather than banning the use of AI tools in the classroom, educators can harness these tools to facilitate student learning, provided appropriate guidance is given (Otsuki, 2020; Hellmich & Vinall, 2021; Carvalho et al, 2022). Teachers often face the challenges of large class sizes and time constraints which may limit the level of individual attention that students receive (Lee, 2008). In this context, AI feedback can complement teacher feedback by providing immediate support, enabling students to work on their writing independently and making timely corrections. As suggested by Barrot (2023), ChatGPT can be useful as an L2 writing assistant and should be adopted as a supplementary tool for essay composition. According to Barrot (2023), students can be encouraged to write their original outputs first and then refine them using ChatGPT. Teachers can emphasize the value of the writing process, and leverage ChatGPT's editing capabilities to teach students correct language forms and style. Students can instruct ChatGPT to edit their work, review the changes made by ChatGPT, and decide whether to accept or reject them based on their assessment.

While integrating ChatGPT into EFL classrooms holds promise, it also presents practical challenges and requires careful consideration of educational institutions' readiness to implement such technology effectively. Educational institutions need to ensure they possess the essential technical infrastructure to incorporate AI-based tools such as ChatGPT into their classrooms. This involves securing reliable internet connectivity, providing access to computers, and ensuring compatibility with the software platform. Teachers need training to effectively integrate ChatGPT into their instructional practices. This training should involve not only the technical aspects of using the tool but also strategies for integrating AI feedback with traditional teaching methods and addressing potential challenges. Integration of ChatGPT should be aligned with the pedagogical goals and objectives of the language learning curriculum. Teachers should consider how AI-based feedback complements instructional practices and supports students' language learning goals.

Students also require instruction in the proper use of AI tools. Warschauer et al. (2023) introduce an AI literacy framework aimed at seamlessly integrating AI tools into L2 writing instruction. The framework suggests several key steps. Firstly, students must grasp the fundamental functionalities, strengths, weaknesses, and biases of AI writing tools. Secondly, they should develop proficiency in accessing and navigating these tools to address specific communication tasks, such as drafting papers or gathering background information. Thirdly, students need to expertly prompt AI to generate content aligned with their writing objectives. Fourthly, they should verify the accuracy of AI-generated content throughout the writing and revision stages. Lastly, students must learn to ethically and effectively incorporate AI-generated texts into their writing, acknowledging and citing their use of AI in the authoring process. They claim that these elements are essential to building second language students' understanding of how to use AI to support their learning.

Several limitations of this pilot study must be noted. Firstly, the sequence of feedback provision—peer-ChatGPT-teacher—employed in this study could have influenced the results. It is plausible that altering the feedback order might have produced different outcomes. Secondly, the small number of participants means that generalization of the findings is difficult to make. Thirdly, the quality and

effectiveness of feedback may differ across teachers, peers, and AI systems, introducing potential variability in students' perceptions and preferences. Fourthly, students' perceptions and attitudes toward feedback can be influenced by various factors, including their prior experiences, language proficiency, and individual preferences. These factors were not addressed in the current study. Lastly, this study only captured students' perceptions at a particular point in time. It did not account for potential shifts in attitudes or preferences over time, especially in light of the novel use of ChatGPT as a feedback tool.

Given that much is still unknown about the use of ChatGPT as a WCF tool in the EFL context, future research should investigate its long-term effects. Furthermore, there is a need for longitudinal studies focusing on feedback combination strategies. Such studies can provide insights into catering feedback approaches to diverse learner needs in EFL contexts. Another avenue of future study could be an exploration of the factors influencing students' preferences for specific feedback methods. Uncovering these factors could help refine feedback practices and identify areas warranting improvement in feedback processes.

Conclusion

This qualitative pilot study attempted to explore EFL students' perceptions of the effectiveness of peer, artificial intelligence (AI), and teacher feedback, and to uncover students' preferences. The participants completed a short writing task which was given written corrective feedback by peers, an AI tool (ChatGPT), and the teacher. Each feedback mode was followed by students' revision of their writing. The data analysis results demonstrated that participants prefer to receive teacher WCF, and a combination of teacher and ChatGPT WCF. An additional aim was to identify the percentage of the feedback given to specific components of the writing. The distribution of WCF across specific aspects of the writing has been discussed. The implication for the EFL classroom is that teachers can integrate ChatGPT into the language learning process, provided appropriate guidance is given. Future research should investigate the long-term effects of using ChatGPT as a WCF tool in the EFL context, study feedback combination strategies, and explore the factors influencing students' preferences for specific feedback methods.

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. Funding for this research was covered by the author(s) of the article.

References

- Adams, R., Nuevo, A., & Egi, T. (2011). Explicit and implicit feedback, modified output, and SLA: Does explicit and implicit feedback promote learning and learner-learner interactions? *Modern Language Journal*, 95(1), 42–63. <https://doi.org/10.1111/j.1540-4781.2011.01242.x>
- Alharbi, W. (2023). AI in the Foreign Language Classroom: A Pedagogical Overview of Automated Writing Assistance Tools. *Education Research International*. <https://doi.org/10.1155/2023/4253331>
- Atlas, S. (2023). ChatGPT for higher education and professional development: A guide to conversational AI. *College of Business Faculty Publications*. https://digitalcommons.uri.edu/cba_facpubs/548
- Barrot, J.S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, 100745. <https://doi.org/10.1016/j.asw.2023.100745>
- Beason, L. (1993). Feedback and revision in writing across the curriculum classes. *Research in The Teaching of English*, 27, 395–421.
- Benson, P. (2013). *Teaching and researching: Autonomy in language learning*. Routledge.

- Bitchener, J., & Ferris, D. (2012). *Written corrective feedback in second language acquisition and writing*. Routledge.
- Brown, J. D., & Bailey, K. M. (1984). A categorical instrument for scoring second language writing skills. *Language Learning*, 34(4), 21–38. <https://doi.org/10.1111/j.1467-1770.1984.tb00350.x>
- Bruner, J. (1985). Vygotsky: A historical and conceptual perspective. *Culture, Communication, and Cognition: Vygotskian perspectives*, 21, 34.
- Carvalho, L., Martinez-Maldonado, R., Tsai, Y.S., Markauskaite, L., & De Laat, M. (2022). How can we design for learning in an AI world? *Computers and Education: Artificial Intelligence*, 3, Article ID 100053. <https://doi.org/10.1016/j.caeai.2022.100053>
- Chen, B. (2015). Social constructivism of language and meaning. *Croatian Journal of Philosophy*, 15(43), 87–113.
- Cho, K., & MacArthur, C. (2010). Student revision with peer and expert reviewing. *Learning and Instruction*, 20(4), 328–338.
- Colpitts, B. D. (2016). Japanese students' perceptions of peer corrective feedback in an EFL classroom. *Acta Humanistica et Scientifica Humanities Series*, 49, 345–358.
- Connor, U., & Asenavage, K. (1994). Peer response groups in ESL writing classes: How much impact on revision? *Journal of Second Language Writing*, 3(3), 257–276. [https://doi.org/10.1016/1060-3743\(94\)90019-1](https://doi.org/10.1016/1060-3743(94)90019-1)
- Covill, A. E. (2010). Comparing peer review and self-review as ways to improve college students' writing. *Journal of Literacy Research*, 42(2), 199–226. <https://doi.org/10.1080/10862961003796207>.
- Dai, W., Lin, J., Jin, H., Li, T., Tsai, Y. S., Gašević, D., & Chen, G. (2023). Can large language Models provide feedback to students? A case study on ChatGPT. In *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)* (pp. 323–325), Orem, UT, USA. doi: 10.1109/ICALT58122.2023.00100.
- Donato, R. (1988). *Beyond group: A psycholinguistic rationale for collective activity in second-language learning*. University of Delaware.
- Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf, & G. Appel (Eds.), *Vygotskian approaches to second language research* (pp. 33–56). Ablex.
- Elola, I., & Oskoz, A. (2010). Collaborative writing: Fostering foreign language and writing conventions development. *Language Learning & Technology*, 14(3), 51–71.
- Fernandez Dobao, A. (2014). Attention to form in collaborative writing tasks: Comparing pair and small group interaction. *Canadian Modern Language Review*, 70(2), 158–187. <https://doi.org/10.3138/cmlr.1768>
- Ferris, D.R. (1997). The Influence of Teacher Commentary on Student Revision. *TESOL Quarterly*, 31(2), 315–339, <https://doi.org/10.2307/3588049>.
- Ferris, D., & Hedgcock, J. (2005). *Teaching ESL composition: Purpose, process, and practice*. Lawrence Erlbaum.
- Ferris, D. R., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, 10(3), 161–184, [https://doi.org/10.1016/S1060-3743\(01\)00039-X](https://doi.org/10.1016/S1060-3743(01)00039-X).
- Garcia Mayo, M.P., & Imaz Agirre, A. (2019). Task modality and pair formation method: Their impact on patterns of interaction and LREs among EFL primary school children. *System*, 80, 165–175. <https://doi.org/10.1016/j.system.2018.11.011>
- Guenette, D. (2007). Is feedback pedagogically correct?: Research design issues in studies of feedback on writing. *Journal of Second Language Writing*, 16(1), 40–53. <https://doi.org/10.1016/j.jslw.2007.01.001>
- Hansen, J. G., & Liu, J. (2005). Guiding principles for effective peer response. *ELT Journal*, 59(1), 31–38. <https://doi.org/10.1093/elt/cci004>

- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research* 77(1), 81–112, <https://doi.org/10.3102/003465430298487>.
- Hedgcock, J., & Lefkowitz, N. (1992). Collaborative oral/aural revision in foreign language writing instruction. *Journal of Second Language Writing*, 1(3), 255–276. [https://doi.org/10.1016/1060-3743\(92\)90006-B](https://doi.org/10.1016/1060-3743(92)90006-B)
- Hellmich, E., & Vinall, K. (2021). FL instructor beliefs about machine translation: Ecological insights to guide research and practice. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, 11(4), 1–18.
- Hewett, B. L. (2000). Characteristics of interactive oral and computer-mediated peer group talk and its influence on revision. *Computers and Composition*, 17(3), 265–288. [https://doi.org/10.1016/S8755-4615\(00\)00035-9](https://doi.org/10.1016/S8755-4615(00)00035-9)
- Hyland, K. (2003). *Second language writing*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511667251>.
- Kamimura, T. (2006). Effects of peer feedback on EFL student writers at different levels of English proficiency: A Japanese context. *TESL Canada Journal*, 23(2), 12–39.
- Kim, Y. (2008). The contribution of collaborative and individual tasks to the acquisition of L2 vocabulary. *The Modern Language Journal*, 92(1), 114–130. <https://doi.org/10.1111/j.1540-4781.2008.00690.x>
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
- Lee, I. (2008). Understanding teachers' written feedback practices in Hong Kong secondary classrooms. *Journal of Second Language Writing*, 17(2), 69–85. <https://doi.org/10.1016/j.jslw.2007.10.001>
- Lin, W. C., & Yang, S. C. (2011). Exploring students' perceptions of integrating Wiki technology and peer feedback into English writing courses. *English Teaching: Practice and Critique*, 10(2), 88–103.
- Little, D. (1991). *Learner Autonomy I: Definitions, issues and problems*. Dublin: Authentik.
- Liu, J., & Hansen, J. G. (2002). *Peer response in second language classroom*. University of Michigan Press.
- Lundstrom, K., & Baker, W. (2009). To give is better than to receive: The benefits of peer review to the reviewer's own writing. *Journal of Second Language Writing*, 18(1), 30–43.
- Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20(1), 51–81. <https://doi.org/10.1017/S027226319800103X>
- OpenAI. (2023). *ChatGPT: Optimizing language models for dialogue*. <https://openai.com/blog/chatgpt/>
- Otsuki, G.J. (2020). OK computer: To prevent students cheating with AI text-generators, we should bring them into the classroom. *The Conversation (blog)*. <https://theconversation.com/ok-computer-to-prevent-students-cheating-with-ai-text-generators-we-should-bring-them-into-the-classroom-129905>
- Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing*, 8(3), 265–289. [https://doi.org/10.1016/S1060-3743\(99\)80117-9](https://doi.org/10.1016/S1060-3743(99)80117-9)
- Peterson, S. S., & Portier, C. (2014). Grade one peer and teacher feedback on student writing. *Education* 3–13, 42(3), 237–257, DOI: 10.1080/03004279.2012.670256.
- Philp, J., Walter, S., & Basturkmen, H. (2010). Peer interaction in the foreign language classroom: What factors foster a focus on form? *Language Awareness*, 19(4), 261–279. <https://doi.org/10.1080/09658416.2010.516831>
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. Oxford University Press.
- Ruegg, R. (2015). The relative effects of peer and teacher feedback on improvement in EFL students' writing ability. *Linguistics and Education*, 29, 73–82.

- Sato, M. (2013). Beliefs about peer interaction and peer corrective feedback: Efficacy of classroom intervention. *Modern Language Journal*, 97(3), 611–633. <https://doi.org/10.1111/j.1540-4781.2013.12035.x>
- Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language Acquisition*, 34(4), 591–626. <https://doi.org/10.1017/S0272263112000356>
- Schreiber, L. M., & Valle, B. E. (2013). Social constructivist teaching strategies in the small group classroom. *Small Group Research*, 44(4), 395–411.
- Sheen, Y. (2010a). Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition*, 32(2), 203–234. <https://doi.org/10.1017/S0272263109990507>
- Sheen, Y. (2010b). Introduction: The role of oral and written corrective feedback in SLA. *Studies in Second Language Acquisition*, 32(2), 169–179. doi:10.1017/S0272263109990489
- Sheen, Y., Wright, D., & Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System*, 37(4), 556–569. <https://doi.org/10.1016/j.system.2009.09.002>
- Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing*, 20(4), 286–305. <https://doi.org/10.1016/j.jslw.2011.05.010>
- Sippel, L., & Jackson, C. N. (2015). Teacher vs. peer oral corrective feedback in the German language classroom. *Foreign Language Annals*, 48(4), 688–705. <https://doi.org/10.1111/flan.12164>
- Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. B. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction*, 91, 101894.
- Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing*, 14(3), 153–173. <https://doi.org/10.1016/j.jslw.2005.05.002>
- Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57, 100752. <https://doi.org/10.1016/j.asw.2023.100752>
- Tai, H.C., Lin, W.C., & Yang, S.C. (2015). Exploring the effects of peer review and teachers' corrective feedback on EFL students' online writing performance. *Journal of Educational Computing Research*, 53(2), 284–309. <https://doi.org/10.1177/0735633115597490>
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249–276.
- van Dis, E. A., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: Five priorities for research. *Nature*, 614(7947), 224–226. <https://doi.org/10.1038/d41586-023-00288-7>
- Villarrreal, I., & Gil-Sarratea, N. (2019). The effect of collaborative writing in an EFL secondary setting. *Language Teaching Research*. Epub ahead of print 15 February 2019. DOI: 10.1177/1362168819829017.
- Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard University Press.
- Warschauer, M., Tseng, W., Soobin, Y., Webster, T., Jacob, S., Du, Q., & Tate, T. (2023). The affordances and contradictions of AI-generated text for writers of English as a second or foreign language. *Journal of Second Language Writing*, 62, 101071. <https://doi.org/10.1016/j.jslw.2023.101071>
- Wenzlaff, K., & Spaeth, S. (2022). Smarter than Humans? Validating how OpenAI's ChatGPT model explains Crowdfunding, Alternative Finance and Community Finance. *Validating how OpenAI's ChatGPT model explains Crowdfunding, Alternative Finance and Community Finance*. (December 22, 2022). <https://ssrn.com/abstract=4302443>
- Williams, J. G. (2003). Providing feedback on ESL students' written assignments. *The Internet TESL Journal*, 9, 1–5.

- Wu, Y. (2024). Study on the impact of utilizing ChatGPT and other AI tools for feedback in EAP writing classrooms on the discursive writing performance of English major students. *Transactions on Social Science, Education and Humanities Research*, 4, 143–150.
- Xiao, Y., & Lucking, R. (2008). The impact of two types of peer assessment on students' performance and satisfaction within a Wiki environment. *Internet and Higher Education*, 11(3–4), 186–193. <https://doi.org/10.1016/j.iheduc.2008.06.005>
- Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15(3), 179–200. <https://doi.org/10.1016/j.jslw.2006.09.004>
- Yoon, S. Y., Miszogladd, E., & Pierce, L. R. (2023). Evaluation of ChatGPT Feedback on ELL Writers' Coherence and Cohesion. *arXiv preprint arXiv:2310.06505*. <https://doi.org/10.48550/arXiv.2310.06505>
- Zhang, S. (1995). Reexamining the affective advantage of peer feedback in the ESL writing class. *Journal of second language writing*, 4(3), 209–222. [https://doi.org/10.1016/1060-3743\(95\)90010-1](https://doi.org/10.1016/1060-3743(95)90010-1)

Appendix A

Table 8 *Feedback Checklist.*

Criteria	✓ if criterion fulfilled	Comments: What needs to be corrected/improved?
Organization		
Thesis statement		
Three ideas		
Each idea is followed by two additional sentences		
Concluding sentence		
Content		
The thesis indicates the topic, and includes a claim about the topic		
The three ideas support the thesis statement		
Each idea is followed by two additional sentences that explain/exemplify the idea		
The concluding sentence restates the thesis and sums up the ideas/gives a suggestion/a personal statement about the topic		
Grammar		
The sentences are grammatically correct		
Mechanics		
The punctuation is correct		
The spelling is correct		
The choice of words is accurate		
Quality of expression		
All the sentences are clear		
All the sentences are logically connected		

Note. A feedback checklist is used in each review round to ensure the fulfillment of specific writing criteria and to incorporate comments from the evaluators. The checklist is adapted from the criteria developed by Brown and Bailey (1984), "The Composition Grading Feedback Sheet".

Appendix B

Feedback Survey

1. a) What did your peers think required correction? *Circle your answers.*
Organization Content Grammar Mechanics Quality of expression
b) Did you understand all the comments?
c) Did you agree with/accept the comments?
 If not, why?
d) Did you find the feedback helpful?
 Why?
2. a) What did ChatGPT suggest correcting? *Circle your answers.*
Organization Content Grammar Mechanics Quality of expression
b) Did you understand all the comments?
c) Did you agree with/accept the comments?
 If not, why?
d) Did you find the feedback helpful?
 Why?
3. a) What did the teacher think required correction? *Circle your answers.*
Organization Content Grammar Mechanics Quality of expression
b) Did you understand all the comments?
c) Did you agree with/accept the comments?
 If not, why?
d) Did you find the feedback helpful?
 Why?
4. a) What feedback or feedback combination do you prefer to receive? *Check (✓) your preferred feedback mode, or specify your preferred combination.*

 ___ Peer feedback
 ___ ChatGPT feedback
 ___ Teacher feedback
 ___ A combination of: _____