

Chapter 14

Revisiting Dialogic Feedback in the Age of GenAI: An Investigation Into L2 Writers' Feedback Seeking Interactions



Da Yan, Shaidatul Akma Adi Kasuma, and Mansour Amini

Abstract This study investigates feedback seeking behaviors (FSB) among second language (L2) learners when engaging with GenAI for writing feedback. Utilizing a multiple case study approach, the research recruited two EFL learners with differing language proficiency and technological competence as participants. Data were collected from three writing tasks in an L2 writing course. After data codification using two tailored coding scheme, the researchers performed statistical analysis on the data of on the participants' prompting patterns and interactional sequences. Findings reveal that the learner with higher language and technology competences crafter longer prompts in fewer rounds of interaction, and gradually adapted towards more sophisticated prompting strategies, while the learner with moderate language and technological competences inclined to shorter prompt in more conversational rounds, and kept on using a simpler approach to prompt GenAI. Additionally, learner with higher language and technology competences comparatively engaged in more effective dialogic interaction with GenAI compared to the other learner who exhibited weaker engagement and likelihood to abruptly terminate the dialogic interaction. This paper contributes to the literature by providing insights into the dynamics of FSB in technology-mediated learning environments and suggests directions for future research to further explore the implications of AI-assisted feedback in language education.

Keywords Feedback seeking · Generative artificial intelligence · Dialogic feedback · L2 writing

D. Yan (✉)

School of Foreign Language Studies, Wenzhou Medical University, Wenzhou, China

S. A. A. Kasuma · M. Amini

School of Languages, Literacies & Translation, Universiti Sains Malaysia, Penang, Malaysia

Introduction

In the realm of second language (L2) writing, feedback—particularly written feedback—has emerged as one of the most extensively studied topics (Hyland & Hyland, 2019). Over the last few decades, research has increasingly shifted its focus from examining the impact of feedback on writing quality (Dizon & Gayed, 2021) to exploring how students engage with and utilize feedback (Ranalli, 2021). At the same time, our definition and understanding of feedback has gone through major changes in recent decades. First, there is a growing trend for researchers in the field of instructional feedback to shift the definition of feedback from a transmission of information to a process in which students and teachers actively participated (Carless, 2022). Second, the conversational or dialogic interactions, particularly aiming at knowledge co-construction, during feedback processes have been highlighted as a key element determining feedback success (Er et al., 2021; Steen-Uthheim & Wittek, 2017). Third, the scope of feedback provider have extended from only human agents to technology-, or material-based agents or resources such as rubric, exemplar, and technology-enhanced systems (Lipnevich et al., 2014; Loncar et al., 2023).

However, against the ongoing reconceptualization of feedback as a student-centered and an ecological complex (Chong, 2021; Nieminen et al., 2022), our understanding of how individual learners would interact with technology-based feedback agent, particularly in a dialogic feedback encounter, remain limited. Dialogic feedback is significant for the field of assessment, since it highlights that feedback is a reciprocal process whereby agents provide and receive evaluative comments on their work or performance (Hattie & Timperley, 2007; Winstone et al., 2021). Since the emergency of AI-powered chatbot using generative artificial intelligence (GenAI) or large language models (LLMs), the potential of technology-based agents to fully engage in a meaningful dialogic with human users for the purpose of generating, modifying, and refining evaluative judgment messages have attracted the attentions of researchers and educators in the fields of L2 writing pedagogy and instructional feedback (Durak et al., 2024). To date, there have limited research efforts into the mechanism of L2 writers' strategies to seek feedback from GenAI and the nuances of the human-GenAI dialogic interactions. Feedback seeking has been conceptualized as a lens from which students' agentic participation in and abilities to regulate the whole feedback process could be understood (Leenknecht & Carless, 2023). Therefore, the knowledge gap in how dialogic feedback between L2 writers and GenAI systems should be revisited and understood remained to be addressed. Thus, the present study attempted to delve into the prompting strategies and interactional sequences in the dialogic interactions between L2 writers and GenAI while the former sought feedback from the latter on writing performance. The research would expand our current understanding of both dialogic feedback and human-AI interactions in an educational assessment/feedback context. Particularly, the research findings could deliver key implications for educators and researchers who are interested to promote the advancement of contemporary

feedback and assessment practices. In the subsequent sections, the existing body of scholarship related to the key concepts involved in the study were reviewed in detail.

GenAI's Potential for Feedback on L2 Writing

The application of generative artificial intelligence (GenAI), such as ChatGPT, has significantly transformed the feedback landscape in second language (L2) writing instruction. Early corpus-based automated writing evaluation (AWE) systems, like Pigai, iWrite, and iTEST, played pivotal roles in reducing teacher workload by automating lower-order concerns, such as grammar and spelling corrections (Zhang, 2021). However, these systems were often limited in their ability to provide more comprehensive, context-aware feedback, leading to the rise of AI-based corrective feedback providers, such as Grammarly and QuillBot. These AI tools have been empirically validated as more effective in improving feedback uptake and revision quality, marking a significant improvement over traditional AWE systems (Ngo et al., 2022).

Since its introduction, GenAI has garnered substantial attention for its potential as a powerful feedback provider in L2 writing. Researchers have investigated various aspects of its functionality, revealing promising results. For example, ChatGPT has demonstrated an ability to offer accurate grammatical and syntactical corrections through its vast repository of pre-trained language data (Fang et al., 2023). Additionally, GenAI's conversational nature allows learners to engage in multi-turn interactions, enabling them to iteratively refine their prompts and feedback, thus improving the quality of revisions over time (White et al., 2023). This dynamic interaction fosters a more personalized and adaptive feedback experience, enhancing learners' engagement and motivation (Yan, 2023).

Moreover, GenAI's ability to generate feedback across various linguistic constructs—beyond mere grammatical correctness—further distinguishes it from earlier systems (Mizumoto & Eguchi, 2023). Learners can use ChatGPT to seek feedback on higher-order concerns, such as content development, coherence, and organization, making it a versatile tool in L2 writing pedagogy. Importantly, ChatGPT is not just a one-time feedback provider but an interactive partner in the revision process, encouraging learners to iteratively improve their writing by engaging in a dialog with the AI (Bang et al., 2023).

Despite its advantages, the integration of GenAI into L2 classrooms also presents challenges. The quality of feedback provided by the AI is heavily influenced by the user's ability to craft effective prompts, and issues such as AI-generated hallucinations—where incorrect or irrelevant information is provided—remain a concern (Tonmoy et al., 2024). Additionally, effective use of GenAI requires a higher level of AI literacy among students, necessitating appropriate scaffolding and support from teachers (Yan, 2023).

Overall, GenAI tools like ChatGPT offer unprecedented opportunities to enhance feedback practices in L2 writing. By enabling more interactive, personalized, and

adaptive feedback processes, they hold the potential to significantly improve learners' writing outcomes while alleviating some of the burdens traditionally placed on teachers and peers. However, careful consideration of the challenges associated with AI feedback tools is necessary to fully realize their potential in educational contexts.

Dialogic Feedback

In recent years, the understanding of feedback in the context of education has gone through significant change from perceiving it as a transmission of information, particularly following an one-directional model from teacher to student, to a process featuring students' active involvement (Carless, 2022). Feedback is now increasingly viewed through a socio-constructivist lens, where learning is a co-constructed process between teachers and students, with a focus on dialogue, participation, and mutual engagement (Evans, 2013). This shift highlights the importance of dialogic feedback, which moves beyond mere information transfer to become an interactive exchange that promotes deeper student reflection and understanding (Steen-Utheim & Wittek, 2017). Dialogic feedback focuses on building relationships of trust and creating opportunities for students to negotiate meanings and clarify expectations, enhancing their capacity to act on feedback (Carless, 2012). Studies have shown that this form of feedback encourages emotional and relational support, which is crucial for maintaining student engagement, as well as fostering environments where students feel safe to express themselves (Hounsell, 2007). This relational aspect of feedback, where emotions and understanding play a key role, marks a clear departure from earlier feedback models that often overlooked the student's emotional experience and sense-making process (Steen-Utheim & Wittek, 2017).

Existing literature on dialogic feedback has been conducted in both the context of teacher-fronted and peer-based feedback. For example, Steen-Utheim & Wittek (2017) examined the dialogic feedback interactions between a teacher and his students in a Norwegian higher education context. The longitudinal and ethnographic study revealed that the dialogic teacher-student interactions provided emotional and relational support, maintained the sustainability of the dialogues, enabled students to better express themselves, and contributed to individual growth. Similarly, the interventional study conducted by Zhao et al. (2024) demonstrated that students exposed to the dialogic peer feedback interventional conditions outperformed those who followed a non-dialogic model of peer feedback in attitude, writing performances, and more constructive revisions. Moreover, Zhao et al. (2024) proposed a three-step model explaining the mechanism of dialogic feedback in a peer-based setting. The three-step feedback model involves a dialogic process aimed at enhancing the effectiveness of peer feedback in writing. The first phase, feedback, involves peer reviewers providing written comments on a peer's work. While essential, this step alone can be insufficient, as students may not fully understand or know how to apply the feedback. In the second phase, feedback-on-feedback, the feedback

receiver (writer) responds to the initial feedback by asking for clarification, negotiating meaning, or seeking more specific advice, fostering deeper engagement. The final phase, re-feedback, allows the original feedback provider to respond to these queries, clarifying or refining their feedback and offering further suggestions if needed. This reciprocal interaction transforms feedback into a co-constructive process, enhancing students' understanding, engagement, and ultimately improving their writing quality (Zhao et al., 2024).

Moreover, recent technological advancements have pushed forward new insights into approaches to understanding interactions in feedback practices. Noticeably, the application of technology-enhanced tools for feedback was framed as a manifestation of the sociometrical perspective of feedback (Gravett, 2022), a posthuman and poststructural approach to include human-machine communication under the conceptual umbrella of feedback interactions. Comparatively, according to a more traditional conceptualization of feedback interaction, the focus on feedback interaction has predominantly been limited to the discussion between human-based agents such as teachers, peers, and individual learners in the existing body of literature. Efforts examining the effects of human-machine interaction in feedback practices could be found in very recent literature. For example, Han et al. (2024) established a corpus of student-ChatGPT interaction in EFL writing classrooms. Analysis on the dialogue data in the corpus facilitated the understanding of students' use intentions and satisfaction with the GenAI-based tool. In another study by Yan (2024c), the researcher explored the effects of different approaches to processing ChatGPT-generated feedback on L2 writing on writing performance. However, the interactional details in students' usage of ChatGPT was not disclosed and studied in the paper. In summary, compared to our rich understanding in the mechanism and effects of human-based feedback dialogues, we possess rather limited knowledge pertinent to similar conversational or dialogic interactions between human and non-human agents in feedback practices. Therefore, the interactions between individual L2 writers and GenAI in the present study could be seen as sociomaterial feedback encounters and events. Such a perspective is a continuation of the longstanding view of feedback, expanding the notion of feedback providers to include non-human agents such as GAI-based tools.

Feedback Seeking Behavior

Feedback-seeking behavior (FSB), which refers to the ways in which learners pursue feedback information, is a key factor influencing the effectiveness of feedback (Leenknecht et al., 2019). This dual concept, consisting of feedback inquiry and monitoring, is adapted from Ashford (2003) to describe the intentional gathering of feedback from various sources, including the environment and other individuals. In the context of language learning, Papi et al. (2019) redefined the term to highlight how students employ both inquiry and monitoring strategies to seek feedback from teachers, peers, and their learning context.

The incorporation of feedback-seeking behavior (FSB) into language learning highlights several key aspects: 1) an emphasis on the individual differences among learners regarding feedback (Winstone et al., 2021); 2) a transition from viewing feedback primarily as a teaching tool to recognizing it as a valuable learning resource (Papi et al., 2019); and 3) a focus on encouraging proactive participation from learners rather than passive reception learners (Malecka et al., 2022). Within the existing literature on FSB in second language (L2) learning, two main research directions can be identified. First, many studies have examined the relationships between FSB, its motivational factors, and the resulting learning outcomes (see Xu & Wang, 2023; Yao & Zhu, 2022). Second, a smaller number of recent studies have investigated L2 learners' FSB in contexts enhanced by technology. For instance, Kessler (2023) suggested that L2 learners in online communities actively sought feedback by making clarification and confirmation requests, displaying distinct strategic behaviors. Likewise, Xu's (2021) study explored the feedback-seeking behaviors of L2 writers during the Covid-19 pandemic and found a positive relationship between feedback-seeking orientations and the application of self-regulatory strategies.

Despite existing research, there is still limited understanding of how students seek feedback in new educational contexts that emphasize innovative feedback paradigms, where the role of teachers is less prominent and the ability of students to generate and interpret technology-based feedback is crucial (Jensen et al., 2023). Current studies on feedback-seeking behavior largely adhere to traditional paradigms, in which students act as receivers of feedback, balancing direct requests for feedback from teachers and peers with indirect monitoring of their learning environment (Papi et al., 2019). In contrast, the evolving educational landscape and the role of learners as active feedback generators highlight the need for further investigation into how students engage in feedback-seeking behaviors within GenAI-enhanced settings.

Research Questions

Based on the above reviewed literature, the study aims at probing into the identified knowledge gaps through the following two research questions (RQs):

RQ1: What prompting strategies have the L2 writers used in seeking feedback from GenAI on writing products?

RQ2: What are the interactional patterns of the conversational interactions between L2 writers and GenAI in seeking feedback on writing products?

Methodology

This research adopted a mixed-method multiple case study design, following the framework outlined by Yin (2013). Specifically, the study aimed at being a collective and multiple case study which enabled broader generalization across multiple cases, offering insights that could be extended to a wider context (Stake, 1995). Case study methodology is widely used in applied linguistics (Duff, 2010) and feedback research (e.g., Zhang & Hyland, 2023), recognized for its ability to gather comprehensive data on students’ learning experiences. However, case study research has frequently been criticized for its small sample size. Adams (2019) argued that the smaller sample sizes were typical in case studies, noting that they allowed for a deeper understanding of how students utilize feedback beyond its mere design. In this study, each case was defined by the students’ prompting efforts to seek feedback from ChatGPT and the patterns of the learners’ interaction with ChatGPT.

The study was conducted within an undergraduate English as a Foreign Language (EFL) program at a Chinese university, where students were required to complete two writing courses that integrated automated and peer feedback systems. A purposeful sampling approach was adopted to select three participants with varying levels of language proficiency and technological competence, ensuring a diverse representation of L2 writers within the program. The screening criteria included: (1) mean scores from prior L2 writing assessments, which provided a reliable indicator of language proficiency; (2) prior experience with and expressed interest in computer-assisted language learning and feedback, gathered through a pre-project survey; (3) performance in two digital humanities courses, where students engaged with technology-enhanced learning tools; and (4) self-assessed technological competence. These criteria were designed to recruit participants with a broad range of abilities and experiences, making them suitable for investigating how L2 learners at different proficiency levels interact with feedback technologies.

Out of an initial pool of volunteers, two participants were selected based on their adherence to the outlined criteria, as well as recommendations from their instructors, who evaluated their engagement with feedback and technological tools in classroom settings. This process ensured that the selected participants reflected distinct levels of both language proficiency and technological competence, offering varied yet representative cases for analysis. Each student’s characteristics are outlined in Table 14.1, with pseudonyms used to maintain anonymity. Both participants provided informed consent, fully aware of the study’s purpose, methodology, and ethical protections.

Table 14.1 Basic information of the participants

Pseudonym	Gender	Age	Language Proficiency	Digital Humanities Performance	Self-rated Technological Competence
Daniel	Male	21	83/100	88/100	9/10
Lily	Female	19	72/100	78/100	7/10

The final selection ensured a balance of students across a spectrum of competencies, with Daniel representing those with high proficiency in both language and technology, and Lily embodying moderate language skills and technological abilities. This mutually complementary nature of the two cases provided a comprehensive foundation for analyzing how students with different backgrounds approach and benefit from feedback systems in an L2 writing context.

The study was embedded in a 16-week course titled “Intermediate English Writing” where students used ChatGPT for feedback on their writing products for the three writing tasks in the course. The writing tasks were designed by the first author according to the test guidelines of the Test for English Majors band 4 (TEM-4) exam, a national English proficiency test for English language majors in China (Yan, 2024a). The writing task required students to summarize a provided passage and provide commentary on specific points, with a minimum word count of 200. The TEM-4 writing tasks aimed to evaluate key writing skills, including: (1) relevance of content, (2) adequacy of content, (3) structure and organization, and (4) effective language usage. For each writing task, the participants sought initial feedback from ChatGPT using a tailored prompt template (Table 14.2), which was created in general alignment with the task requirement of TEM-4 writing tasks. After receiving the initial feedback, the students could continue to discuss with ChatGPT for clarification and refining the feedback.

To collect data of students’ prompts and interactional pattern, the study applied two different instruments respectively. First, throughout the project, students were required to document the prompts they generated to engage with ChatGPT in their learning logs. These logs were flexible and could include supplementary multimedia materials such as screenshots, video recordings, or other relevant files. Moreover, to comprehensively capture students’ learning behaviors, the project was conducted in computerized language laboratories equipped with advanced technologies like real-time screen recording, keylogging, and video-based classroom monitoring systems. After each session, students were also asked to review their learning logs and

Table 14.2 Sample prompt template

1. Task requirement, feedback requirement, and context
Sample: <i>“Please provide feedback and evaluation on an essay writing by a second-year university English as a foreign language (EFL) learner. For the rating, you could consult the rating criteria that follows. For the feedback for writing improvement, try to focus on grammatical correctness, fluency, coherency, and the quality of the language. Try to give detailed feedback and explain any suggestions for improvement.”</i>
2. Rating criteria
[Rating criteria in the official TEM-4 test specifications including the following three dimensions:]
(a) Content (10 marks)
(b) Organization (3 marks)
(c) Language use (7 marks)
3. Writing product
[The draft writing product to be assessed by ChatGPT]

perform member-checking to ensure accuracy and completeness of the entries. Second, the instructors were tasked with capturing the students' conversational interactions with ChatGPT. Students participated in all feedback and revision sessions in language labs outfitted with keylogging and screen recording technology to support the data collection process. All recorded data, including logs and screen captures, were collected, processed and analyzed for further coding and analyses.

Two additional coding schemes were utilized to translate the prompts recorded in the learning logs and the video-recorded student-ChatGPT interactions into numerical data. First, for the coding of prompt writing patterns, a coding scheme developed by the first author in a previous work was used (Yan, 2024b). The coding was performed by four coders recruited from the teaching faculty. The coders first individually coded learning logs from 50 students. As good inter-rater reliability was attained (Fleiss' Kappa = 0.88, 95%CI [0.76, 0.93]), they split the remaining coding duties. Second, another coding scheme was applied to convert the video-recorded interactions between students and ChatGPT into numerical data. This scheme was based on the three-phase dialogical peer feedback framework developed by Zhao et al. (2024) but was adapted for this study. The coding scheme consisted of five categories: initial feedback (IF), feedback-on-feedback (FF), feedback reception without inquiry (NI), re-feedback (RF), and non-response to feedback-on-feedback (NR). Compared to the original three-step framework in Zhao et al. (2024), two additional categories—feedback reception without inquiry and ignoring feedback-on-feedback—were added to identify episodes where students demonstrated superficial engagement or disengagement with feedback generated by ChatGPT. Similarly, three co-researchers first coded the interactional data of 30 students individually. Then based on the good inter-rater reliability (Fleiss' Kappa = .78, 95%CI [0.68, 0.86]), the remainder of the coding duties were split among the coders.

After data codification, the numerical data were analyzed. First, we used descriptive statistics and visualization techniques to display the prompt patterns and their development across the tasks. Second, to analyze students' interaction patterns with ChatGPT, a lag sequential analysis (LSA) was conducted using GSEQ 5.1 software. LSA is a statistical technique designed to uncover the sequences and transitions between behaviors by calculating the likelihood of one event following another within a specified time frame or lag (Bakeman & Quera, 2011). In this study, LSA was applied to the coded interactions between students and ChatGPT, derived from the five categories in the adapted coding scheme. GSEQ software was used to compute adjusted residuals based on the transitional probability matrix generated from these coded sequences of interaction (Pohl et al., 2016). This matrix examined the conditional probabilities of students moving between various stages of the feedback process. The significance of these transitions was determined using the Z-scores of the adjusted residuals, where $Z > 1.96$ indicated statistically significant transitions. Visual representations of these behavioral transitions were generated to illustrate the interaction patterns, with particular attention to how students engaged with or disengaged from the feedback.

Results

Prompt Patterns

Table 14.3 shows an overview of the ChatGPT prompts created by Daniel. It provides details about the number of prompts, words generated, words per prompt, rounds of interactions, average number of prompts per round, and time spent on each task. Across the three tasks, Daniel generated a total of 974 prompts, leading to a combined word count of 11,820. The mean number of words per prompt was 12.8, with an average of 18 rounds per task. The total time spent on these tasks was 465 min, averaging 155 min per task. Similarly, Table 14.4 shows the overview of ChatGPT prompts Lily has created. Lily generated a total of 608 prompts across the three tasks, with a combined word count of 6542. The average number of words per prompt was 10.77, and the tasks involved an average of 34.33 rounds. Lily spent a total of 289 min, averaging 96.3 min per task. Daniel and Lily’s efforts to prompt ChatGPT for feedback differed in several key areas. Daniel generated more prompts (974 vs. 608) and a higher total word count (11,820 vs. 6542) across tasks. His prompts were slightly longer on average (12.8 vs. 10.77 words per prompt) and involved fewer rounds (18 vs. 34.33) but more prompts per round (18 vs. 5.9). Additionally, Daniel spent more time overall (465 min vs. 289 min), averaging 155 min per task compared to Lily’s 96.3 min. In summary, Daniel’s approach involved more intensive and prolonged interactions, while Lily used fewer but more concise prompts over more rounds in shorter sessions.

Figure 14.1 illustrates Daniel’s evolving prompt patterns across three tasks, showing a shift from basic to more complex strategies. In Task 1, Daniel predominantly used barebone (minimal) prompts (BP), with 112 instances, which significantly decreased in Tasks 2 and 3 (93 and 70, respectively). Over time, he

Table 14.3 Overview of ChatGPT prompts created by Daniel

	Prompt	Words	Words per prompt	Rounds	Prompts per round	Time spent (min)
Task 1	238	2500	12.5	15	15.9	198
Task 2	331	3778	13.1	21	15.8	141
Task 3	405	5542	12.8	18	22.5	126
Total	974	11,820	–	54	–	465
Mean	324.67	3940	12.8	18	18	155

Table 14.4 Overview of ChatGPT prompts created by Lily

	Prompt	Words	Words per prompt	Rounds	Prompts per round	Time spent (min)
Task 1	180	1885	9.5	31	5.7	108
Task 2	211	2400	12.1	38	6.1	97
Task 3	217	2257	10.7	34	5.8	84
Total	608	6542	–	103	–	289
Mean	202.67	2180.67	10.77	34.33	5.9	96.3

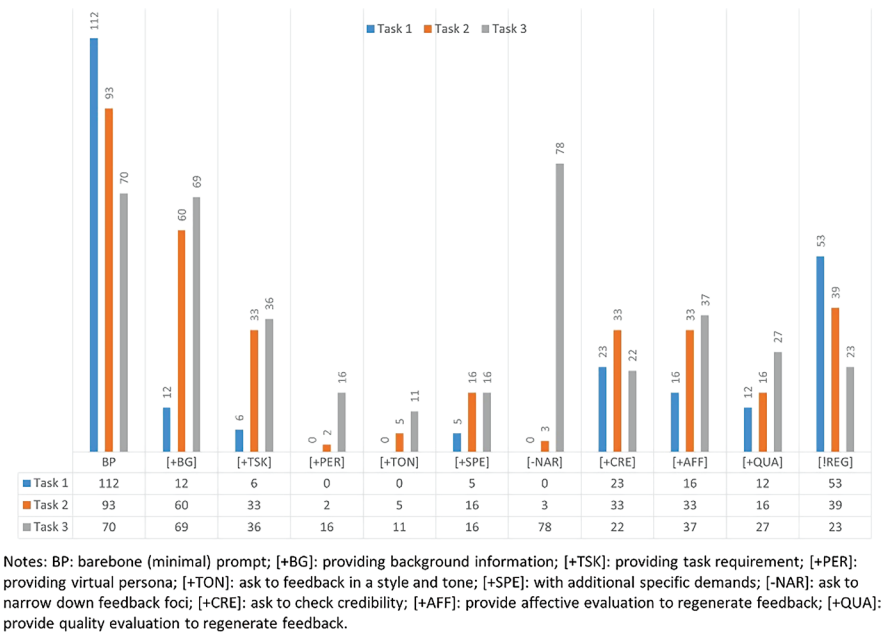


Fig. 14.1 Prompt patterns of Daniel

increasingly provided background information (+BG), growing from 12 in Task 1 to 69 in Task 3, and consistently included task requirements (+TSK). Advanced strategies became more prominent in later tasks, such as narrowing down feedback foci (-NAR), which rose dramatically from 0 in Task 1 to 78 in Task 3, and affective evaluation prompts (+AFF), increasing from 23 to 39. Additionally, prompts for checking credibility (+CRE), requesting feedback in a specific style or tone (+TON), and ensuring quality evaluation (+QUA) showed upward trends.

Figure 14.2 illustrates Lily’s evolving prompt patterns across the same three tasks, showcasing her progression in prompt complexity. In Task 1, Lily primarily employed barebone prompts (BP), with 79 instances, similar to Daniel but slightly lower. This usage decreased in Tasks 2 and 3, with 54 and 38 instances, respectively, indicating a shared trend of moving away from minimal prompts. Lily demonstrated a significant increase in providing background information (+BG), from 15 in Task 1 to 58 in Task 3, which reflects a strategic shift towards more context-rich prompting. Task requirements (+TSK) were consistently included, with a rise from 20 in Task 1 to 42 in Task 3. Advanced strategies emerged prominently in later tasks, particularly in narrowing down feedback foci (-NAR), which escalated from 1 in Task 1 to 60 in Task 3. Affective evaluation prompts (+AFF) saw an increase as well, growing from 15 to 31. Additionally, prompts for checking credibility (+CRE), requesting feedback in a specific style or tone (+TON), and ensuring quality evaluation (+QUA) also displayed upward trends.

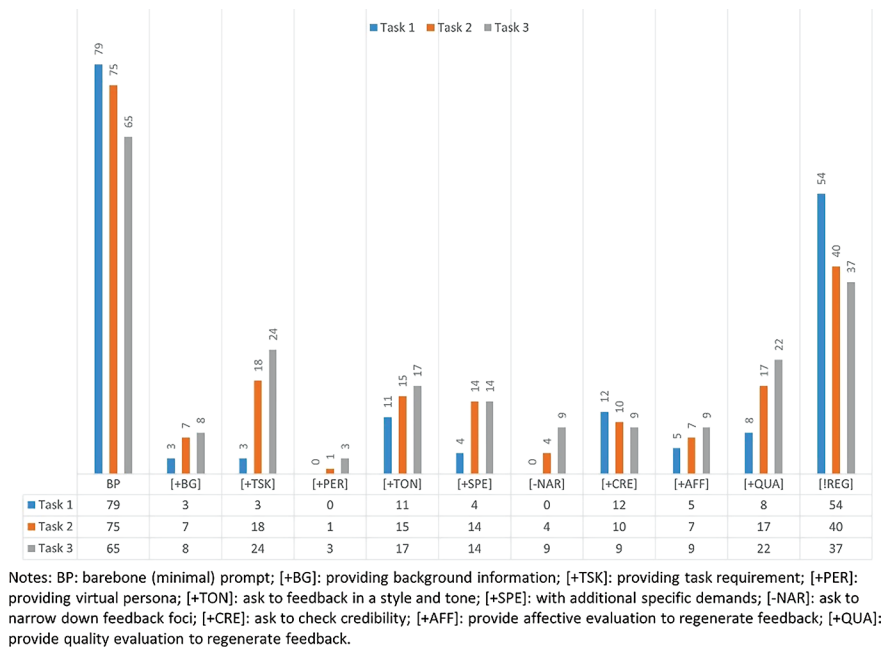


Fig. 14.2 Prompt patterns of Lily

Overall, Daniel showed a more pronounced adaptation to complex strategies, while Lily maintained a consistent approach to task requirements and background information, reflecting the differences in their evolving capabilities to utilize ChatGPT for feedback on writing products.

Interactional Sequences

The results of LSA for both students are shown in Tables 14.5 and 14.6 respectively. Each row of the table represents an initial behavior while each column shows a subsequent behavior. For example, the IF→FF sequence displayed the behavioral transition of ‘after the reception of *initial feedback* from ChatGPT, the student continued to engage in conversation with ChatGPT for *feedback-on-feedback*’. When such a behavioral transition was statistically significant ($Z_{IF \rightarrow FF} = 9.699$), it indicated a strong trend to formulate back evaluation of the received feedback from ChatGPT.

Figure 14.3 illustrates the contrasting interactional sequences in the feedback processes of Daniel and Lily as reflected in the LSA results. Daniel employs a robust three-phase approach characterized by initial feedback (IF),

Table 14.5 Frequency transition and corresponding adjusted residual of Daniel

		IF	FF	NI	RF	NR
IF	Freq.	15	142	17	12	0
	AR	3.42	7.51	0.86	−9.95	−2.02
FF	Freq.	0	0	9	78	2
	AR	−2.44	−13.19	0.85	14.81	1.53
NI	Freq.	0	3	0	0	0
	AR	−0.38	1.46	−0.51	−1.08	−0.17
RF	Freq.	0	45	0	0	1
	AR	−1.61	5.83	−2.16	−4.54	0.95
NR	Freq.	0	0	0	0	0
	AR	0	0	0	0	0

Note: * $Z > 1.96$; *Freq.* Frequency, *AR* Adjusted residual for the behavioral transition

Table 14.6 Frequency transition and corresponding adjusted residual of Lily

		IF	FF	NI	RF	NR
IF	Freq.	16	49	42	8	0
	AR	4.74	8.7	4.48	−6.77	−7.90
FF	Freq.	0	0	19	68	63
	AR	−4.67	−8.88	−4.63	6.92	8.02
NI	Freq.	0	0	0	0	0
	AR	0	0	0	0	0
RF	Freq.	0	1	1	0	0
	AR	−0.36	1.14	0.90	−0.90	−0.80
NR	Freq.	0	0	0	0	0
	AR	0	0	0	0	0

Note: * $Z > 1.96$; *Freq.* Frequency, *AR* Adjusted Residual for the behavioral transition

feedback-on-feedback (FF), and re-feedback (RF), with solid transitions indicating effective engagement (IF → FF → RF). His dialogic communication with ChatGPT is further enhanced by a bidirectional channel (IF ↔ RF) that facilitates the exchange of feedback messages requiring explanation or clarification. In contrast, Lily’s feedback process is marked by weaker transitions, particularly from RF to FF (−4.63), indicating a lack of engagement in revisiting feedback-on-feedback. Additionally, her strong transitions towards feedback reception without inquiry (NI) and non-response to feedback-on-feedback (NR) (8.7 and 8.02, respectively) suggest a tendency to terminate dialogue prematurely, limiting the potential for comprehensive feedback cycles. Overall, the diagram highlights Daniel’s proactive, iterative approach versus Lily’s more static engagement when seeking feedback on writing performance from ChatGPT.

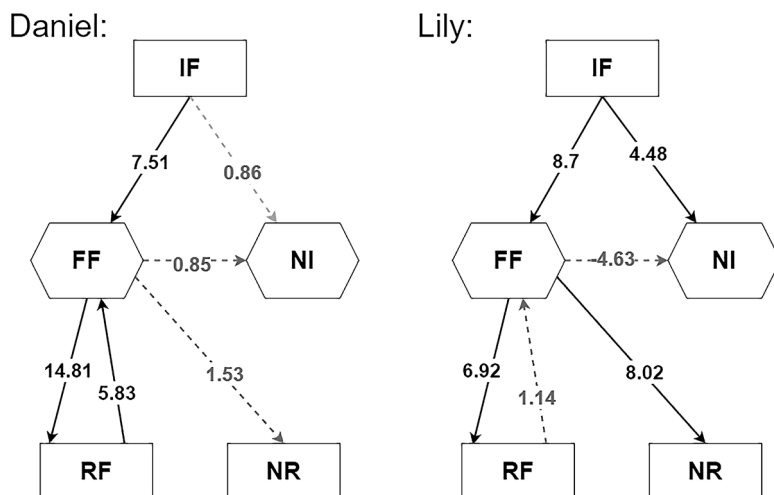


Fig. 14.3 Interactional patterns of students' dialogic use of ChatGPT for feedback

Discussion and Conclusions

The present multiple case study explored how two L2 learners with varying language proficiency and technological competence engaged with ChatGPT when seeking feedback on their writing performance. The findings revealed distinct patterns of feedback-seeking behavior, where students who actively made use of sophisticated prompts also displayed a higher tendency to engage more actively with the dialogic interaction with ChatGPT. By investigating these dynamics through a behavior-oriented perspective, this research extends existing literature on L2 writing pedagogy and feedback practices, offering novel insights into how FSB and learner characteristics interact in the digital learning environment. The following discussion will connect these findings with existing studies on feedback practices, examining how they align with or diverge from previous research.

Regarding **RQ1**, the research findings highlight notable differences in the prompt patterns used by Daniel and Lily. Daniel's approach involved generating a larger number of prompts with a focus on complexity and depth, indicating a proactive engagement with ChatGPT. He progressively incorporated more advanced prompting strategies, facilitating richer feedback interactions. In contrast, Lily's prompts were generally more concise and consistent, emphasizing clarity and context. While her method allowed for thorough exploration across multiple rounds, it may have limited the depth of engagement compared to Daniel's more intensive approach. Overall, these differences underscore how varied prompting strategies can significantly impact the quality of feedback received from AI systems. The findings have

extended our understanding of FSB in an L2 writing context. Previously, FSB were understood of students' explicit/direct or implicit/indirect efforts to elicit support from the learning environment (Papi et al., 2020). Comparatively, the finding of the present study revealed that, in a technology-enhanced learning environment, students' efforts to seek feedback from technology-based agent in feedback practices were generally active, explicit, and direct. However, the two participants still displayed different strategical use and considerations when devising their individual ChatGPT prompts. This variation in strategic use highlights the importance of individual differences, particularly domain knowledge and AI/technology literacy (Yan et al., 2023; Yan & Wang, 2022), in how students interact with AI tools (Koltovskaia, 2020; Yan, 2023). Specifically, Daniel's more complex prompting strategy reflects a willingness to engage deeply with the technology, potentially leading to more nuanced feedback. In contrast, Lily's approach, while effective, illustrates a tendency towards a more straightforward engagement that may not fully leverage the capabilities of AI. In the context of GenAI and technology-supported language learning, the approach used by Daniel aligned with existing argument on the recommend use of ChatGPT by utilizing its iterative and conversational nature and mechanism (White et al., 2023).

Underneath the frequencies of prompt patterns and the underlying prompting strategies, the two participants varied in their metacognitive awareness, which determined how they formulated the prompts in different feedback scenarios. Specifically, Daniel's ability to navigate complex prompting likely stems from a higher level of metacognitive awareness regarding his feedback needs and the capabilities of ChatGPT. This awareness enables him to reflect on the feedback received and adjust his prompts accordingly, leading to a more refined interaction. In contrast, Lily's more straightforward prompting might reflect a lower level of metacognitive engagement. If she is less aware of how to assess the feedback or how to ask for clarification, her prompts may not evolve as effectively. This echoed previous literature on the role of metacognition on FSB (de Chantal & Organisciak, 2023). Noticeably, the study was in tandem with a previous study by Yan and Zhang (2024) that students' behavioral engagement with technology-supported feedback was related to their (meta)cognitive engagement. Following this line of reasoning, we could infer that students who possessed higher level of metacognitive or self-regulatory capabilities were more likely to rely on advanced or sophisticated approaches to prompting GenAI systems and more effectively control their interactions with AI.

Regarding **RQ2**, the research findings reveal crucial insights into the interactional sequences employed by Daniel and Lily when engaging with ChatGPT for feedback. Daniel exhibited a robust interaction model which highlights his proactive approach, as he actively sought clarification and further elaboration on the feedback received. In contrast, Lily's interactional sequences were less dynamic, suggesting a tendency to receive feedback without adequately revisiting or

questioning it, which may limit her ability to fully engage with the feedback process. The identification of the two different approaches to engage with dialogic feedback validated and supported the three-step model proposed by Zhao et al. (2024). Moreover, it also extended the scope of applicability for the framework, which was primarily used in the setting of human-based feedback practices. From this perspective, the findings of the current study not only indicated that human learners could engage in dialogic interactions with both human-based and technology-supported feedback agent, but also unveiled the potential for GenAI-based systems to somewhat replace or augment the role played by human learners in peer feedback settings.

Specifically, the study revealed that the two participants adhered to two different approaches to participating in the dialogic interaction with ChatGPT. Therefore, it testified to the understanding that the use of GenAI for feedback was related to students' modes of involvement. This was in agreement with the notion of the "Human in the loop" approach proposed by Liu et al. (2024), which highlighted the significance of human roles in AI-assisted language instructions. Additionally, it was also in alignment with the argument on the significance of students' agency in feedback practices, particularly in the "new feedback paradigms", where student-centeredness and activeness were highlighted (Jensen et al., 2023).

From another perspective, the findings about Lily's interactional pattern showcased that, in technology-supported feedback, students still had a tendency to passively or ineffectively engage with dialogic feedback. This was in line with previous observation of students' feedback engagement, where a similar scenarios has been addressed as a relatively negative pole of the feedback engagement continuum (Wu & Schunn, 2023). Specifically, Lily's interactions often involved minimal follow-up inquiries or clarification requests, resulting in a surface-level adoption of feedback without deeper reflection or revision. The finding revealed the complexity of feedback practices and students' engagement with it. The findings echoed Wu and Schunn's (2023) and Yan's (2024a) findings that some learners, despite having access to dialogic feedback, may struggle to fully capitalize on its potential for deeper learning and critical self-assessment. Such passive engagement patterns could be attributed to both cognitive and affective factors. For example, the students' lack of prior experiences with feedback negotiation or their over-reliance on the automated feedback itself might cause their inability to sustainably involve in the human-AI dialogs (Koltovskaia, 2020). Similarly, internal resistance to discuss with other agents in feedback practices, particularly when facing negative comments or relatively a harsh tone in presenting feedback, would restricted their interest and courage to continue discussing with AI about the feedback (Bharuthram & van Heerden, 2023; Papi, 2021).

The study had several implications for educators and researchers aiming to promote the effective use of GenAI in language learning. First, it revealed that active

feedback seeking is crucial; educators should encourage students to adopt proactive strategies and formulate sophisticated prompts to maximize their interactions with AI tools. However, in the current body of literature, the understanding of FSB in a technology-supported language learning context remains severely under-explored. Researchers and educators should collaborate to design studies that explore various dimensions of feedback engagement, including the effectiveness of different prompting strategies and the impact of individual learner characteristics on interaction quality. Second, there is a need for targeted training in metacognitive strategies through integrating metacognitive awareness activities into the curriculum can empower students to reflect on their feedback needs and enhance their engagement with GenAI systems. From a standpoint of self-regulated learning, fostering metacognitive skills allows students to become more aware of their cognitive processes, enabling them to set specific goals, monitor their progress, and adjust their strategies when interacting with AI tools (Yan et al., 2025). Finally, recognizing individual differences in technology literacy and language proficiency is essential. Thus, tailored instruction can ensure all students are equipped to engage meaningfully with AI feedback, while emphasizing the dialogic nature of interactions can foster critical self-assessment skills. Educators should attempt to construct an inclusive and equitable learning environment in which all students could make effective use of and reap learning gains from GenAI-based language learning scaffolds.

The present study has several limitations that warrant consideration in future research. First, as a multiple case study, the generalizability of the findings is inherently limited. Future research should aim to explore the mechanisms of feedback-seeking behavior from GenAI within a larger sample, employing diverse methodologies or designs to enhance the external validity of the results. Second, the reliance on quantitative methods in this study restricts our understanding of the discursive nuances within human-GenAI dialogues in L2 learning contexts. Future investigations could employ qualitative approaches, such as naturalistic or conversational analysis, to provide richer insights into the dynamics of these interactions and how they shape feedback processes. Finally, the current study didn't offer insights into the longitudinal effect of GenAI-mediated dialogic feedback on learning in authentic learning environments. Therefore, the researchers would suggest future scholarship to examine the longitudinal effects of the intervention through experimental studies.

Competing Interests The authors have no competing interests to declare.

Disclaimer The manuscript consists of content developed with the assistance of an artificial intelligence tool, Alibaba Qwen, 2024. Every section was carefully reviewed, revised, and approved by the author(s) to ensure that the content is appropriate, scientifically accurate, and maintains the integrity of the original content.

References

- Adams, G. (2019). A narrative study of the experience of feedback on a professional doctorate: 'A kind of flowing conversation'. *Studies in Continuing Education*, 41(2), 191–206. <https://doi.org/10.1080/0158037X.2018.1526782>
- Ashford, S. (2003). Reflections on the looking glass: A review of research on feedback-seeking behavior in organizations. *Journal of Management*, 29(6), 773–799. [https://doi.org/10.1016/S0149-2063\(03\)00079-5](https://doi.org/10.1016/S0149-2063(03)00079-5)
- Bakeman, R., & Quera, V. (2011). *Sequential analysis and observational methods for the behavioral sciences*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139017343>
- Bang, Y., Cahyawijaya, S., Lee, N., Dai, W., Su, D., Wilie, B., Lovenia, H., Ji, Z., Yu, T., Chung, W., Do, Q. V., Xu, Y., & Fung, P. (2023). A multitask, multilingual, multimodal evaluation of ChatGPT on reasoning, hallucination, and interactivity (arXiv:2302.04023). arXiv. <https://doi.org/10.48550/ARXIV.2302.04023>
- Bharuthram, S., & van Heerden, M. (2023). The affective effect: Exploring undergraduate students' emotions in giving and receiving peer feedback. *Innovations in Education and Teaching International*, 60(3), 379–389. <https://doi.org/10.1080/14703297.2022.2040567>
- Carless, D. (2012). Trust and its role in facilitating dialogic feedback. In *Feedback in higher and professional education* (pp. 90–103). Routledge.
- Carless, D. (2022). From teacher transmission of information to student feedback literacy: Activating the learner role in feedback processes. *Active Learning in Higher Education*, 23(2), 143–153. <https://doi.org/10.1177/1469787420945845>
- Chong, S. W. (2021). Reconsidering student feedback literacy from an ecological perspective. *Assessment & Evaluation in Higher Education*, 46(1), 92–104. <https://doi.org/10.1080/002602938.2020.1730765>
- de Chantal, P.-L., & Organisciak, P. (2023). Automated feedback and creativity: On the role of metacognitive monitoring in divergent thinking. *Psychology of Aesthetics Creativity and The Arts*, No Pagination Specified-No Pagination Specified. <https://doi.org/10.1037/aca0000592>
- Dizon, G., & Gayed, J. (2021). Examining the impact of Grammarly on the quality of mobile L2 writing. *The JALT CALL Journal*, 17(2), 74–92. <https://doi.org/10.29140/jaltcall.v17n2.336>
- Duff, P. (2010). *Case study research in applied linguistics*. Routledge.
- Durak, G., Çankaya, S., Özdemir, D., & Can, S. (2024). Artificial intelligence in education: A bibliometric study on its role in transforming teaching and learning. *The International Review of Research in Open and Distributed Learning*, 25(3), 219–244. <https://doi.org/10.19173/irrodl.v25i3.7757>
- Er, E., Dimitriadis, Y., & Gašević, D. (2021). A collaborative learning approach to dialogic peer feedback: A theoretical framework. *Assessment & Evaluation in Higher Education*, 46(4), 586–600. <https://doi.org/10.1080/002602938.2020.1786497>
- Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research*, 83(1), 70–120. <https://doi.org/10.3102/0034654312474350>
- Fang, T., Yang, S., Lan, K., Wong, D. F., Hu, J., Chao, L. S., & Zhang, Y. (2023). Is ChatGPT a highly fluent grammatical error correction system? A comprehensive evaluation (arXiv:2304.01746). arXiv. <https://doi.org/10.48550/ARXIV.2304.01746>
- Gravett, K. (2022). Feedback literacies as sociomaterial practice. *Critical Studies in Education*, 63(2), 261–274. <https://doi.org/10.1080/17508487.2020.1747099>
- Han, J., Yoo, H., Myung, J., Kim, M., Lee, T. Y., Ahn, S.-Y., & Oh, A. (2024). RECIPE4U: Student-ChatGPT interaction dataset in EFL writing education (arXiv:2403.08272). arXiv. <https://doi.org/10.48550/arXiv.2403.08272>
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112. (arXiv:2403.08272). arXiv. <https://doi.org/10.3102/003465430298487>
- Hounsell, D. (2007). Towards more sustainable feedback to students. In *Rethinking assessment in higher education*. Routledge.

- Hyland, K., & Hyland, F. (2019). Contexts and issues in feedback on L2 writing. In F. Hyland (Ed.), *Feedback in second language writing: Contexts and issues* (2nd ed., pp. 1–22). Cambridge University Press. <https://doi.org/10.1017/9781108635547.003>
- Jensen, L. X., Bearman, M., & Boud, D. (2023). Feedback encounters: Towards a framework for analysing and understanding feedback processes. *Assessment & Evaluation in Higher Education*, 48(1), 121–134. <https://doi.org/10.1080/02602938.2022.2059446>
- Kessler, M. (2023). Written corrective feedback in an online community: A typology of English language learners' requests and interlocutors' responses. *Computers and Composition*, 67, 102752. <https://doi.org/10.1016/j.compcom.2023.102752>
- Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*, 44, 100450. <https://doi.org/10.1016/j.asw.2020.100450>
- Leenknecht, M., Hompus, P., & Van Der Schaaf, M. (2019). Feedback seeking behaviour in higher education: The association with students' goal orientation and deep learning approach. *Assessment & Evaluation in Higher Education*, 44(7), 1069–1078. <https://doi.org/10.1080/002602938.2019.1571161>
- Leenknecht, M. J. M., & Carless, D. (2023). Students' feedback seeking behaviour in undergraduate education: A scoping review. *Educational Research Review*, 40, 100549. <https://doi.org/10.1016/j.edurev.2023.100549>
- Lipnevich, A. A., McCallen, L. N., Miles, K. P., & Smith, J. K. (2014). Mind the gap! Students' use of exemplars and detailed rubrics as formative assessment. *Instructional Science*, 42(4), 539–559. <https://doi.org/10.1007/s11251-013-9299-9>
- Liu, Z.-M., Hwang, G.-J., Chen, C.-Q., Chen, X.-D., & Ye, X.-D. (2024). Integrating large language models into EFL writing instruction: Effects on performance, self-regulated learning strategies, and motivation. *Computer Assisted Language Learning*, 1–25. <https://doi.org/10.1080/09588221.2024.2389923>
- Loncar, M., Schams, W., & Liang, J.-S. (2023). Multiple technologies, multiple sources: Trends and analyses of the literature on technology-mediated feedback for L2 English writing published from 2015–2019. *Computer Assisted Language Learning*, 36(4), 722–784. <https://doi.org/10.1080/09588221.2021.1943452>
- Malecka, B., Boud, D., & Carless, D. (2022). Eliciting, processing and enacting feedback: Mechanisms for embedding student feedback literacy within the curriculum. *Teaching in Higher Education*, 27(7), 908–922. <https://doi.org/10.1080/13562517.2020.1754784>
- Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, 2(2), 100050. <https://doi.org/10.1016/j.rmal.2023.100050>
- Ngo, T. T.-N., Chen, H. H.-J., & Lai, K. K.-W. (2022). The effectiveness of automated writing evaluation in EFL/ESL writing: A three-level meta-analysis. *Interactive Learning Environments*, 1–18. <https://doi.org/10.1080/10494820.2022.2096642>
- Nieminen, J. H., Tai, J., Boud, D., & Henderson, M. (2022). Student agency in feedback: Beyond the individual. *Assessment & Evaluation in Higher Education*, 47(1), 95–108. <https://doi.org/10.1080/02602938.2021.1887080>
- Papi, M. (2021). The role of motivational and affective factors in L2 writing performance and written corrective feedback processing and use. In R. M. Manchón & C. Polio (Eds.), *The Routledge handbook of second language acquisition and writing* (pp. 152–165). Routledge.
- Papi, M., Bondarenko, A. V., Wawire, B., Jiang, C., & Zhou, S. (2020). Feedback-seeking behavior in second language writing: Motivational mechanisms. *Reading and Writing*, 33(2), 485–505. <https://doi.org/10.1007/s11145-019-09971-6>
- Papi, M., Rios, A., Pelt, H., & Ozdemir, E. (2019). Feedback-seeking behavior in language learning: Basic components and motivational antecedents. *The Modern Language Journal*, 103(1), 205–226. <https://doi.org/10.1111/modl.12538>

- Pohl, M., Wallner, G., & Kriglstein, S. (2016). Using lag-sequential analysis for understanding interaction sequences in visualizations. *International Journal of Human-Computer Studies*, 96, 54–66. <https://doi.org/10.1016/j.ijhcs.2016.07.006>
- Ranalli, J. (2021). L2 student engagement with automated feedback on writing: Potential for learning and issues of trust. *Journal of Second Language Writing*, 52, 100816. <https://doi.org/10.1016/j.jslw.2021.100816>
- Stake, R. E. (1995). *The Art of Case Study Research*. Sage Publications.
- Steen-Utheim, A., & Witteke, A. L. (2017). Dialogic feedback and potentialities for student learning. *Learning, Culture and Social Interaction*, 15, 18–30. <https://doi.org/10.1016/j.lcsi.2017.06.002>
- Tonmoy, S. M. T. I., Zaman, S. M. M., Jain, V., Rani, A., Rawte, V., Chadha, A., & Das, A. (2024). *A comprehensive survey of hallucination mitigation techniques in large language models* (arXiv:2401.01313). arXiv. <https://doi.org/10.48550/arXiv.2401.01313>
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). *A prompt pattern catalog to enhance prompt engineering with ChatGPT* (arXiv:2302.11382). arXiv. <https://doi.org/10.48550/ARXIV.2302.11382>
- Winstone, N., Hepper, E. G., & Nash, R. (2021). Individual differences in self-reported use of assessment feedback: The mediating role of feedback beliefs. *Educational Psychology*, 41(7), 844–862. <https://doi.org/10.1080/01443410.2019.1693510>
- Wu, Y., & Schunn, C. D. (2023). Passive, active, and constructive engagement with peer feedback: A revised model of learning from peer feedback. *Contemporary Educational Psychology*, 73, 102160. <https://doi.org/10.1016/j.cedpsych.2023.102160>
- Xu, J. (2021). Chinese university students' L2 writing feedback orientation and self-regulated learning writing strategies in online teaching during COVID-19. *The Asia-Pacific Education Researcher*, 30(6), 563–574. <https://doi.org/10.1007/s40299-021-00586-6>
- Xu, J., & Wang, Y. (2023). Individual differences in L2 writing feedback-seeking behaviors: The predictive roles of various motivational constructs. *Assessing Writing*, 56, 100698. <https://doi.org/10.1016/j.asw.2023.100698>
- Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, 28(11), 13943–13967. <https://doi.org/10.1007/s10639-023-11742-4>
- Yan, D. (2024a). Rubric co-creation to promote quality, interactivity and uptake of peer feedback. *Assessment & Evaluation in Higher Education*, 49(8), 1017–1034. <https://doi.org/10.1080/002602938.2024.2333005>
- Yan, D. (2024b). Feedback seeking abilities of L2 writers using ChatGPT: A mixed method multiple case study. *Kybernetes*. <https://doi.org/10.1108/K-09-2023-1933>
- Yan, D. “Alex.” (2024c). Comparing individual vs. collaborative processing of ChatGPT-generated feedback: Effects on L2 writing task improvement and learning. *Language Learning & Technology*, 28(1), 1–19.
- Yan, D. “Alex,” Jia, C., & Tian, F. (2025). Automated feedback complements online dialogic peer feedback: Fostering L2 writers' feedback quality and self-regulatory writing strategy use. *Interactive Learning Environments*, 1–15. <https://doi.org/10.1080/10494820.2025.2491620>
- Yan, D., Kasuma, S. A. A., & Amini, M. (2023). Gender difference in digital literacy among translation trainees: Self-perceptions and application abilities. *International Journal of Academic Research in Progressive Education and Development*, 12(3), 2383–2399. <https://doi.org/10.6007/IJARPEd/v12-i3/19611>
- Yan, D., & Wang, J. (2022). Teaching data science to undergraduate translation trainees: Pilot evaluation of a task-based course. *Frontiers in Psychology*, 13, 939689. <https://doi.org/10.3389/fpsyg.2022.939689>
- Yan, D., & Zhang, S. (2024). L2 writer engagement with automated written corrective feedback provided by ChatGPT: A mixed-method multiple case study. *Humanities and Social Sciences Communications*, 11(1), 1–14. <https://doi.org/10.1057/s41599-024-03543-y>

- Yao, Y., & Zhu, X. (2022). Modeling Chinese high school EFL writers' language mindsets, goal orientations, and feedback-seeking behavior: The interaction effect between language mindsets and writing competence. *Reading and Writing*. <https://doi.org/10.1007/s11145-022-10339-6>
- Yin, R. K. (2013). *Case study research: Design and methods* (5th ed.). SAGE Publications, Inc..
- Zhang, S. (2021). Review of automated writing evaluation systems. *Journal of China Computer-Assisted Language Learning*, 1(1), 170–176. <https://doi.org/10.1515/jccall-2021-2007>
- Zhang, Z., & Hyland, K. (2023). Student engagement with peer feedback in L2 writing: Insights from reflective journaling and revising practices. *Assessing Writing*, 58, 100784. <https://doi.org/10.1016/j.asw.2023.100784>
- Zhao, Y., Zhang, F., Schunn, C. D., He, P., Li, D., & Zhao, Y. (2024). Feedback, feedback-on-feedback and re-feedback: Effects of written dialogic peer feedback on English as a foreign language writing. *Assessment & Evaluation in Higher Education*, 49(4), 523–538. <https://doi.org/10.1080/02602938.2023.2278017>