ReCALL (2026), 1–18
doi:10.1017/S095834402610055X

RESEARCH ARTICLE

Effects of dialogic and monologic AI/non-AI feedback on L2 speaking development: Exploring learner achievements, feedback literacy, and interactional moves

Hanieh Shafiee Rad
English Department, Shahrekord University, Shahrekord, Islamic Republic of Iran (haniyeshafieerad@gmail.com)

Abstract

Although prior studies have examined the role of artificial intelligence (AI) in supporting second language (L2) speaking development, limited attention has been paid to how feedback modes (i.e., dialogic vs. monologic) and delivery media (i.e., AI vs. non-AI) differentially shape learners' oral performance and feedback literacy. Addressing this gap, the present study investigated a total of 120 intermediate learners of English who were recruited from language institutes and randomly assigned to four equal groups: AI-dialogic (ChatGPT), AI-monologic (ELSA Speak), non-AI-dialogic (Zoom), and non-AI-monologic (Audacity). Data were collected through a pretest–posttest speaking assessment, a feedback literacy questionnaire, semi-structured interviews, and analysis of recorded dialogic interactions. Quantitative findings revealed that the AI-dialogic group outperformed the other groups in both speaking proficiency and feedback literacy, followed by the non-AI-dialogic group, whereas the monologic groups made comparatively weaker gains. Qualitative analysis revealed that dialogic conditions facilitated clarification requests, confirmation checks, and self-repair, whereas monologic conditions hindered learner uptake and interaction. These findings demonstrate that the mode of feedback and its medium of delivery play a crucial role in oral development, carrying both theoretical and pedagogical implications for designing hybrid feedback practices that integrate AI's precision with the co-constructive power of dialogue.

Keywords: artificial intelligence; feedback; monologic feedback; feedback literacy; L2 speaking

1. Introduction

In second language (L2) learning, corrective feedback is not merely a mechanism for error correction but functions as a pedagogical bridge that mediates learners' movement from their current performance toward their potential developmental level (Morgana, 2024). Within speaking instruction, where immediacy, interactivity, and learner engagement are paramount, the way feedback is delivered can significantly shape both skill acquisition and learner motivation (Shadiev et al., 2026). The increasing availability of artificial intelligence (AI)-driven feedback tools, ranging from automated speech evaluation systems to conversational agents, has added a new dimension to this process, raising critical questions about how such feedback compares to

Email: haniyeshafieerad@gmail.com

Cite this article: Shafiee Rad, H. (2026). Effects of dialogic and monologic AI/non-AI feedback on L2 speaking development: Exploring learner achievements, feedback literacy, and interactional moves. *ReCALL* FirstView, 1–18. https://doi.org/10.1017/S095834402610055X

© The Author(s), 2026. Published by Cambridge University Press on behalf of EUROCALL, the European Association for Computer-Assisted Language Learning. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

more conventional, human-delivered alternatives (Liu et al., 2025a, 2025b). More recently, the distinction between dialogic and monologic feedback constitutes a fundamental aspect of feedback in teaching/learning (Kazemi et al., 2025). Dialogic feedback involves a reciprocal exchange in which learners and feedback providers jointly negotiate meaning, clarify misunderstandings, and co-construct knowledge (Er et al., 2021). In contrast, monologic feedback is unidirectional, evaluative, and offers little scope for learner input or negotiation (Chinpakdee, 2026).

While feedback has long been recognized as central to communicative language teaching (Ellis, 2017; Lyster & Ranta, 1997), research on AI-mediated speaking feedback is still in its infancy, especially in terms of how the mode of delivery (dialogic vs. monologic) interacts with the source (AI vs. human) to influence learning. Much of the existing work has focused on writing, pronunciation scoring, or discrete error correction, leaving the broader dynamics of spoken interaction underexplored (Bibauw et al., 2019). Furthermore, studies that address L2 speaking often concentrate on measurable outcomes, such as gains in accuracy or fluency, without examining the process-oriented dimensions of how learners interpret, respond to, and act upon the feedback they receive (Timpe-Laughlin et al., 2022). This omission is critical because learner uptake, self-regulation, and affective responses are shaped not only by the content of feedback but also by the conversational architecture in which it is embedded (Du, 2025).

This study seeks to address these gaps by systematically comparing the repercussions of dialogic and monologic feedback, delivered by both AI and non-AI sources, on L2 speaking skill development. Drawing on the interaction hypothesis by Long (1996) and sociocultural perspectives on mediated learning by Vygotsky (1978) and Lantolf and Thorne (2006), it examines not only measurable gains in speaking performance but also the dialogic moves learners employ and their perceptions of the feedback process. By integrating performance data with learner-reported experiences, the study aims to provide a richer account of how feedback mode and source jointly shape speaking development. In doing so, it offers both empirical evidence and pedagogical implications for the design of feedback systems – human, AI, or hybrid – that are capable of fostering sustained learner engagement and communicative competence.

2. Literature review

2.1. Artificial intelligence and speaking skill

The integration of AI into L2 speaking instruction has been increasingly framed within socio-constructivist and interactionist perspectives, which view language development as emerging through meaningful interaction and feedback. AI-driven tools such as automated speech recognition (ASR), intelligent tutoring systems, and conversational agents operationalize these theories by creating opportunities for learners to engage in extended dialogue, receive adaptive feedback, and rehearse speaking in low-anxiety environments (Hou & Min, 2026; Wu et al., 2025). From a sociocultural perspective, AI can be understood as a mediational tool that scaffolds learners' communicative competence within their zone of proximal development (ZPD; Vygotsky, 1978), providing individualized prompts, corrections, and models that enhance fluency and accuracy. Furthermore, dialogic interaction with AI systems aligns with Long's (1996) interaction hypothesis, as learners are exposed to the negotiation of meaning and are prompted to produce output, while corrective feedback supports the noticing of gaps in interlanguage (Swain, 2000). These theoretical foundations suggest that AI technologies, far from being mere substitutes for human teachers, embody pedagogical agents that can facilitate adaptive, learner-centered oral practice in ways that conventional monologic or teacher-dominated instruction often cannot.

Empirical studies increasingly confirm the potential of AI-enhanced environments to support speaking skill development in L2 learning (e.g., Fathi et al., 2024). Research on ASR-based systems indicates measurable gains in pronunciation, fluency, and accuracy, as learners benefit from real-time feedback and repeated opportunities for self-correction (Amrate & Tsai, 2025; Evers & Chen,

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

2022). More recent work demonstrates that AI-powered chatbots and intelligent tutoring systems promote communicative confidence and reduce learner anxiety by offering authentic yet low-stakes spaces for oral practice (Wu et al., 2025). Large-scale studies also suggest that AI-driven dialogic feedback fosters greater learner engagement and sustained oral production compared to conventional corrective approaches that refer to teacher- or task-initiated forms of error correction, such as recasts, prompts, or explicit corrections (Hou & Min, 2026). However, scholars caution that technological affordances alone do not guarantee learning outcomes; learner motivation, feedback literacy, and pedagogical integration remain crucial mediating factors (Zhang & Zou, 2022). Collectively, empirical evidence highlights both the promise and the challenges of AI-assisted speaking instruction, underscoring the need for research that balances technological innovation with pedagogical intentionality to optimize its role in L2 contexts. Liu et al. (2025a) examined the effectiveness of AI-integrated systems on learners' speaking skills, but their study did not account for the role of different feedback modes. Therefore, there is a noticeable gap in the literature concerning the examination of monologic and dialogic feedback within AI-mediated environments and their specific impact on the development of speaking skills.

2.2. Dialogic and monologic feedback

Dialogic peer feedback has been increasingly acknowledged as a dynamic and socially mediated process that enriches learning by transforming feedback into a two-way interaction rather than a passive reception of comments. Unlike monologic feedback, which often positions learners as recipients of instructor judgments, dialogic approaches emphasize shared meaning-making, negotiation, and co-construction of knowledge among participants (Bibauw et al., 2019; Steen-Utheim & Wittek, 2017). Rooted in sociocultural theory, this perspective views learning as fundamentally social, where scaffolding within the ZPD enables learners to extend their skills and understanding with the support of peers (Vygotsky, 1978). Importantly, dialogic peer feedback is not limited to hierarchical exchanges from more proficient to less proficient learners; rather, it can operate reciprocally, allowing all students to engage in mutual scaffolding (Donato, 1994).

Recent studies have further shown that factors such as feedback literacy (Carless & Boud, 2018), learner motives (Evers & Chen, 2022), and proficiency differences (Dai & Wu, 2022) influence how learners participate in and benefit from feedback dialogues. When learners actively assume roles as givers, receivers, or even coordinators, they not only refine their understanding but also develop collaborative competencies that are crucial in higher education contexts (Forbes et al., 2024). These insights underscore the pedagogical value of dialogic peer feedback as a learner-centered practice that fosters reflection, confidence, and deeper engagement with language learning (Li & Su, 2025). In addition to peer-mediated feedback, research on computer-mediated communication shows that online dialogue, whether synchronous or asynchronous, supports negotiation of meaning, clarification requests, and co-construction of knowledge (Blake, 2017; Dooly, 2008). Learners often use interactional moves such as confirmation checks and reformulations in text- or video-based interactions, mirroring face-to-face dialogic feedback (Pellettieri, 2000; Storch, 2002). This evidence underscores that feedback is most effective when embedded in interactive, co-constructive discourse, regardless of the medium. Research shows that teacher- or peer-facilitated dialogic sessions via platforms like Zoom provide real-time interaction similar to face-to-face negotiation (Forbes et al., 2024). Learners can make clarification requests, confirm understanding, and collaboratively repair errors – key dialogic moves in L2 learning (Lyster & Ranta, 1997; Storch, 2002). Non-AI-dialogic feedback emphasizes scaffolding, turn-taking, and role rotation, fostering co-construction of knowledge and supporting feedback literacy, clearly distinguishing it from monologic and AI-mediated conditions (Er et al., 2021).

In contrast to dialogic approaches, monologic feedback has long been the conventional mode of response in language education, operating through a predominantly unidirectional transmission of teacher authority to the learner. In this model, feedback is typically evaluative and corrective,

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

with instructors identifying errors, providing explicit solutions, and expecting students to internalize these without sustained dialogue (Bibauw & Dizon, 2025; Carless, 2006). Although such feedback offers efficiency and ensures accuracy in contexts where expert judgment is paramount (Evans, 2013), it often positions learners as passive recipients, thereby constraining opportunities for self-regulation, reflection, and deeper engagement with learning processes (Boud & Molloy, 2013). Recent studies, including those by Li and Su (2025), highlight that while monologic practices may secure immediate corrective outcomes, they risk reducing learners' agency and feedback literacy by neglecting the co-construction of meaning central to sustainable learning. Moreover, the absence of dialogic negotiation may contribute to learners' misinterpretations or superficial uptake of teacher comments, a phenomenon well-documented in L2 writing research (Myers & Buchanan, 2025). Nevertheless, monologic feedback should not be dismissed entirely; in high-stakes assessment or novice learning contexts, its clarity, directness, and authoritative tone may serve as valuable scaffolding (Cai et al., 2023). The challenge, therefore, lies in reconceptualizing monologic feedback as part of a broader continuum, where it can be complemented by dialogic approaches that foster learner agency, motivation, and feedback literacy (Carless & Boud, 2018; Winstone & Carless, 2020). However, to the best of the researcher's knowledge, there remains a paucity of research investigating the role of dialogic and monologic AI and non-AI feedback environments within the L2 context.

2.3. This study

Although prior studies have explored the role of feedback in L2 speaking skills (e.g., Liu et al., 2025a), limited research has systematically compared dialogic and monologic feedback across both AI and non-AI environments in this area. This gap is significant because the mode of feedback and the mediating platform can shape learner engagement, perception, and the quality of dialogic moves during oral practice. To address this, the current study investigates how different feedback modes influence L2 learners' speaking proficiency and perceptions, thereby offering pedagogical insights into optimizing AI and non-AI feedback for oral language development. The following research questions were addressed in this study:

1. To what extent do dialogic and monologic feedback in AI and non-AI environments differentially influence L2 learners' speaking proficiency?
2. To what extent does the feedback literacy of L2 learners vary across dialogic versus monologic AI and non-AI feedback environments?
3. How do L2 learners perceive and experience dialogic moves during AI- and non-AI-mediated feedback on their speaking performance?

3. Method

3.1. Design and participants

This study used a mixed-methods design, beginning with quantitative data collection and analysis, followed by qualitative data to gain a deeper understanding. A total of 120 intermediate Persian-speaking English learners (ages 19–25, $M = 20.57$, $SD = 1.78$; 2–4 years of prior study) were randomly selected from language institutes and assigned to four groups: two experimental (AI-monologic, AI-dialogic) and two control (non-AI-monologic, non-AI-dialogic). None had visited an English-speaking country.

3.2. Materials

The study used *Four Corners Level 2* (Richards & Bohlke, 2011) for its communicative approach and balanced focus on fluency and accuracy. Over 12 weeks, all groups covered 12 units with real-

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

life themes (e.g., daily routines, travel, work, technology, culture), providing authentic speaking practice relevant to learners' academic, social, and professional lives.

3.3. Tools

3.3.1. Speaking test

For the pretest and posttest, participants completed three speaking tasks based on the IELTS speaking band: a Part 1 interview with personal questions, a Part 2 long-turn (cue card) task requiring them to describe a specific topic, and a Part 3 discussion involving broader, opinion-based questions (see Appendix A for the tests; all appendices appear in the supplementary material). The pretest focused on free-time activities and travel experiences, while the posttest focused on timing. The participants' speaking skills were evaluated using the IELTS speaking band descriptors (see Appendix B), which assess fluency and coherence, vocabulary use, grammatical range and accuracy, and pronunciation. Scores from each area (1–9) were averaged to produce an overall speaking score. To ensure accuracy, a second experienced EFL instructor rescored the recordings, resulting in an acceptable interrater reliability of 0.83 (Pearson correlation).

3.3.2. Feedback literacy scale

The student feedback literacy scale used in the current study was developed and validated by Zhan (2022) and includes 24 items (see Appendix C). The items are rated on a 6-point Likert scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*), with no neutral midpoint. The scale measures six dimensions: eliciting, processing, enacting, appreciation, readiness, and commitment. Each dimension consists of four items, totaling 24 questions. The scale was translated into the learners' mother tongue to ensure clarity and prevent any potential misconceptions. It was validated by conducting forward and backward translation, expert review for content equivalence, and a pilot test to ensure clarity and cultural appropriateness. Cronbach's alpha was calculated and found to be satisfactory at 0.86.

3.3.3. Recording for dialogic moves

As part of examining learners' dialogic moves, all participants were asked to record their spoken interactions and submit the recordings to the teacher for subsequent analysis. Dialogic moves were analyzed using a customized coding scheme adapted from Lyster and Ranta's (1997) corrective feedback uptake framework and Storch's (2002) interaction patterns (see Appendix D). The instrument classified learners' responses such as clarification requests, confirmation checks, uptake with or without repair, and self-repair, alongside broader interactional patterns including collaborative, expert/novice, dominant/passive, and dominant/dominant exchanges. This framework enabled systematic examination of how learners engaged with feedback in both AI-enhanced and non-AI-enhanced dialogic conditions. The recordings were transcribed verbatim for thematic analysis, allowing the identification of common patterns and differences across the groups. The reliability and validity of them were established through expert review for content validity and confirmed by satisfactory Cronbach's alpha coefficients for internal consistency of 0.91 and 0.89.

3.4. AI-enhanced applications for monologic and dialogic feedback

In this study, ELSA Speak (https://elsaspeak.com) was selected as the monologic AI feedback application for speaking, while ChatGPT (https://chatgpt.com) was employed as the dialogic AI feedback application. ELSA Speak was chosen because it provides automated one-way corrective feedback on learners' pronunciation, intonation, and fluency, offering precise phoneme-level analysis and accuracy scores without requiring extended interaction, thus representing a

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

prototypical form of monologic feedback. By contrast, ChatGPT was selected for the dialogic condition due to its capacity to sustain interactive, back-and-forth conversations that allow learners to negotiate meaning, request clarification, and receive context-sensitive scaffolding. Importantly, the AI-dialogic condition involved individual, one-to-one interactions between each learner and ChatGPT, rather than group-based engagement. Learners interacted independently in real time with ChatGPT as a personalized conversational partner. To ensure consistency across participants and tasks, all interactions followed a standardized system prompt and a fixed learner-initiated prompt, which defined ChatGPT's role, prioritized meaning before form, limited corrective feedback to one or two issues per turn, and encouraged self-repair through clarification rather than immediate correction. Learners responded orally to ChatGPT's follow-up questions and scaffolding moves. The exact prompts and interactional constraints are provided in Appendix E, ensuring internal validity and replicability of the AI-dialogic condition.

In both the ELSA Speak (AI-monologic) and human-monologic conditions, feedback targeted the same aspects of speaking proficiency, including pronunciation, intonation, fluency, vocabulary use, and grammatical accuracy, ensuring comparability across modalities. Both provided one-way, evaluative, and corrective feedback without extended interaction. ELSA Speak delivered automated monologic feedback, while the human instructor provided one-way oral feedback; in both cases, the focus, scope, and evaluative nature of the feedback were consistent. The use of these two applications ensured a clear contrast between evaluative, corrective feedback in a monologic mode and interactive, co-constructive feedback in a dialogic mode, thereby aligning with the aims of the present research. This individual configuration in the AI-dialogic condition was designed to align with the AI's individualized feedback function and minimize peer-interaction effects that could confound outcomes. Although differences in format (one-to-one AI vs. small-group human interaction) may affect internal validity, task type, instructional goals, and feedback focus were held constant across conditions. Accordingly, observed differences are interpreted cautiously as reflecting both feedback modality (AI vs. human) and the affordances and constraints of the respective interactional configurations.

3.5. Non-AI-enhanced applications for monologic and dialogic feedback

In the non-AI conditions of this study, Audacity (https://www.audacityteam.org/) was selected as the monologic feedback application for speaking, while Zoom (https://www.zoom.com) was employed as the dialogic feedback application. Audacity was chosen because it enables learners to record and replay their speech samples, after which the teacher provides corrective comments on aspects such as pronunciation, accuracy, and fluency; this process is unidirectional and evaluative in nature, thereby reflecting the characteristics of monologic feedback. By contrast, Zoom was selected for the dialogic condition because it facilitates real-time interactive communication between teachers and learners or among peers, allowing for clarification requests, negotiation of meaning, and co-construction of knowledge. The use of these two applications created a clear contrast between non-interactive corrective feedback and interactive meaning-focused feedback, thus supporting the objectives of the current study.

3.6. Procedures for experimental and control groups

The study was carried out over a 12-week period, during which both experimental and control groups received identical instructional content, but their feedback was mediated through different applications. In the experimental groups, feedback was provided through AI-enhanced platforms: learners in the monologic AI condition used ELSA Speak, where they practiced speaking tasks and received immediate, automated corrective feedback on pronunciation, intonation, and fluency. This feedback was evaluative and one-directional, enabling students to recognize and adjust their errors independently. In the dialogic AI condition, learners engaged with ChatGPT in interactive

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

Table 1. Procedure by group

Group	Feedback mode	Platform	Feedback process	Session structure
Experimental Group 1 (AI-dialogic)	Dialogic (interactive, conversational)	ChatGPT	Learners engaged in interactive speaking tasks with ChatGPT, receiving context-sensitive prompts, clarifications, and scaffolding. Feedback was dialogic and negotiated.	Pre-task (20 min), task + ChatGPT interaction and dialogic feedback (50 min), reflection and peer discussion (20 min).
Experimental Group 2 (AI-monologic)	Monologic (one-way corrective)	ELSA Speak	Learners completed speaking tasks and received automated feedback on pronunciation, intonation, and fluency. Feedback was evaluative and corrective, without extended interaction.	Pre-task (20 min), task + ELSA practice and automated feedback (50 min), reflection on errors and re-practice (20 min).
Control Group 1 (non-AI-dialogic)	Dialogic (teacher/peer interaction)	Zoom	Learners participated in speaking tasks on Zoom with peers and the teacher, receiving real-time dialogic feedback through clarification, negotiation, and co-construction.	Pre-task (20 min), task + Zoom interactive session (50 min), reflection and follow-up discussion (20 min).
Control Group 2 (non-AI-monologic)	Monologic (teacher corrective)	Audacity	Learners recorded their speech in Audacity. Teachers later provided corrective comments on pronunciation, grammar, and fluency. Feedback was one-directional.	Pre-task (20 min), task + speech recording (50 min), teacher feedback reviewed and reflection (20 min).

speaking tasks that simulated real-life communication. Also, two-way interaction with ChatGPT gave qualitative feedback on pronunciation, intonation, and fluency. Here, learners were able to negotiate meaning, request clarification, and reformulate their speech based on context-sensitive prompts, creating an ongoing dialogue that mirrored authentic conversational exchanges (see Table 1 for more details).

In the control groups, non-AI tools were employed to parallel the feedback modes. Learners in the monologic non-AI condition used Audacity to record and replay their speech, after which the teacher provided corrective feedback on their recordings, focusing on accuracy, pronunciation, and fluency. Feedback was delivered in a unidirectional manner, consistent with the nature of monologic feedback. In contrast, learners in the dialogic non-AI condition engaged in speaking tasks via Zoom, where real-time interaction with peers and the instructor fostered dialogic feedback. These exchanges encouraged clarification, confirmation, and co-construction of knowledge, ensuring that feedback was embedded in dynamic, two-way communication. Across all conditions, speaking tasks were matched in type, topic, and difficulty. Dialogic groups completed the same communicative tasks, interacting either with ChatGPT or with teachers/peers, while monologic groups completed identical speaking prompts individually and received one-way feedback. Also, all conditions targeted the same speaking features, pronunciation, fluency, and grammatical accuracy, to ensure alignment and control for task effects.

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

Before the intervention, all participants attended a single 60-minute in-person training session. During this session, learners were introduced to the assigned platform, practiced completing sample speaking tasks, and learned how to interpret and respond to feedback. This standardized training aimed to reduce technological and procedural variability, consistent with computer-assisted language learning research highlighting the role of learner preparation in effective technology-mediated instruction. To minimize confounding effects, speaking tasks were carefully matched across all groups in type, topic, and difficulty, and all groups targeted the same speaking features, pronunciation, fluency, and grammatical accuracy. Pre- and posttests were designed with comparable prompts and assessment criteria to ensure measurement consistency. Interactional roles were standardized within each feedback mode (AI or human, dialogic or monologic), and instructional content, time on task, and feedback episodes were held constant, ensuring that observed differences could be attributed to feedback mode and mediation rather than task or procedural variability. This parallel design allowed for a systematic analysis of AI-enhanced and non-AI-enhanced monologic and dialogic feedback modes on learners' speaking development.

3.7. Study procedure

At the outset of the study, several language institutes were contacted and informed that a free 12-week speaking course would be offered to interested learners. Those who registered were first required to take the Oxford Placement Test, and only learners who demonstrated an intermediate level of English proficiency were selected to participate. In total, 120 learners met the criteria and were then randomly assigned to four groups: two experimental groups and two control groups. Before the instructional phase, all participants completed two baseline assessments: a speaking pretest and a feedback literacy questionnaire. These measures ensured comparability across groups before treatment. To maintain instructional consistency, the same teacher was assigned to all four groups, and all participants followed the same coursebook, which provided weekly speaking tasks and practice opportunities.

During the treatment, the experimental groups received AI-enhanced feedback: one group worked with monologic feedback using ELSA Speak, and the other group worked with dialogic feedback using ChatGPT. In contrast, the control groups received non-AI feedback: one group practiced with monologic feedback via Audacity, while the other engaged in dialogic feedback using Zoom. This design allowed for direct analysis of AI and non-AI tools, as well as between dialogic and monologic feedback modes. Across all conditions, feedback was operationalized by mode (dialogic vs. monologic) and delivery medium (AI vs. non-AI). Monologic feedback (AI and non-AI) was unidirectional, explicitly corrective, and form-focused; AI-monologic feedback (ELSA Speak) was immediate and multimodal, whereas non-AI-monologic feedback (Audacity) was written and delayed. Dialogic feedback (both AI and non-AI) was interactive and embedded in ongoing communication, primarily immediate, task-relevant, and combining implicit prompts with optional explicit explanations. AI-dialogic feedback (ChatGPT) occurred through individualized learner–AI exchanges, while non-AI-dialogic feedback (Zoom) involved real-time teacher or peer interaction (see Appendix F).

Across the 12-week intervention, all four groups were designed to receive the same number of feedback episodes per learner per session, thereby ensuring equivalence of feedback quantity across conditions. What differed was the distribution of feedback moves: in the dialogic conditions (AI and non-AI), feedback episodes were realized mainly through clarification requests, confirmation checks, negotiated reformulations, uptake with repair, and self-repair, whereas in the monologic conditions (AI and non-AI), the same number of feedback episodes consisted primarily of evaluative and explicit corrective feedback, with uptake occurring mostly as uptake with or without repair and limited opportunity for negotiation. In addition, time on task was fully equivalent across groups, with all participants completing 90-minute sessions comprising identical pre-task, task-plus-feedback, and reflection phases, resulting in comparable exposure to input and

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

output; thus, differences in outcomes are attributable to feedback mode and interactional configuration rather than unequal feedback quantity or instructional time. After the intervention, participants completed the speaking posttest and the feedback literacy questionnaire again to measure progress. Additionally, semi-structured interviews were conducted with a sample of learners from each group to gain deeper insights into their perceptions of feedback processes and the role of AI in shaping their speaking development.

3.8. Data analysis

Data analysis followed a mixed-methods approach aligned with the three research questions. For the first two questions, ANCOVAs were conducted on pretest and posttest scores, controlling for initial differences, with post hoc pairwise comparisons and effect sizes calculated where significant effects were found. For the third question, learners' recorded interactions were coded using a customized dialogic move scheme adapted from Lyster and Ranta (1997) and Storch (2002), which captured uptake patterns and interactional dynamics, and semi-structured interviews were thematically analyzed to explore perceptions of dialogic feedback across AI- and non-AI-mediated conditions.

4. Results

4.1. Answering the first and second research questions

Table 2 reports descriptive statistics for speaking performance and feedback literacy across groups. Experimental Group 1 (AI-enhanced dialogic feedback) showed the largest gains, with speaking performance rising from $M = 3.59$ ($SD = .259$) to 8.90 ($SD = .088$) and feedback literacy from 1.29 ($SD = .091$) to 5.44 ($SD = .099$); skewness and kurtosis values were within acceptable ranges, indicating approximate normality. Experimental Group 2 (AI-enhanced monologic feedback) improved moderately in both speaking performance ($M = 3.70$ to 5.26) and feedback literacy ($M = 1.36$ to 1.78), with distributions showing slight deviations but remaining acceptable. Control Group 1 (non-AI-enhanced dialogic feedback) also progressed (speaking: $M = 3.60$ to 6.33; feedback literacy: $M = 1.30$ to 4.23), though with somewhat higher variability, while Control Group 2 (non-AI-enhanced monologic feedback) displayed minimal improvement in either measure (speaking: $M = 3.74$ to 3.87; feedback literacy: $M = 1.31$ to 1.78).

Additionally, the Kolmogorov–Smirnov and Shapiro–Wilk results indicated that all distributions of speaking and feedback scores at both the pretest and posttest were not significantly different from normality ($p > .05$), confirming the general suitability of the data for parametric tests. Then, an initial ANOVA was conducted to examine whether the four groups (Experimental Group 1, Experimental Group 2, Control Group 1, and Control Group 2) differed before the intervention. ANOVA results indicated no significant baseline differences among the groups on speaking performance and feedback literacy, with F values of 0.65 and 0.48 and corresponding p values (0.71, 0.59) all above 0.05. A repeated measures ANOVA was then conducted, and its findings are summarized in Table 3.

The ANCOVA findings demonstrated that group membership exerted a statistically significant effect on both speaking performance, $F(3, 115) = 286.09$, $p < .001$, $\eta^2 = .882$, and feedback literacy, $F(3, 115) = 208.34$, $p < .001$, $\eta^2 = .845$, after controlling for pretest scores. While the speaking pretest emerged as a significant covariate, $F(3, 115) = 64.55$, $p < .001$, $\eta^2 = .360$, the feedback pretest did not, $F(3, 115) = 1.03$, $p = .312$, $\eta^2 = .009$, suggesting that the substantial posttest differences were primarily attributable to the instructional group conditions rather than initial feedback literacy levels. Subsequently, post hoc tests were conducted to identify the specific group differences underlying the overall effects, and the results are summarized in Table 4.

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

Table 2. Descriptive statistics

Groups			*N*	*M*	*SD*	Skewness	Kurtosis
Experimental Group 1	Pretest	Speaking performance	30	3.59	.259	−.009	.415
	Posttest	Speaking performance	30	8.90	.088	−1.121	1.270
	Pretest	Feedback literacy	30	1.29	.091	−.383	−.259
	Posttest	Feedback literacy	30	5.44	.099	−.037	−1.011
Experimental Group 2	Pretest	Speaking performance	30	3.70	.341	.000	1.936
	Posttest	Speaking performance	30	5.26	.949	−.310	−.529
	Pretest	Feedback literacy	30	1.36	.124	−.483	.219
	Posttest	Feedback literacy	30	1.78	.300	−.399	−.507
Control Group 1	Pretest	Speaking performance	30	3.60	.576	.184	1.341
	Posttest	Speaking performance	30	6.33	1.165	−.251	−.621
	Pretest	Feedback literacy	30	1.30	.0970	−1.216	1.149
	Posttest	Feedback literacy	30	4.23	1.216	−.941	.660
Control Group 2	Pretest	Speaking performance	30	3.74	.877	.532	−.306
	Posttest	Speaking performance	30	3.87	.903	.670	−.053
	Pretest	Feedback literacy	30	1.31	.141	.190	.050
	Posttest	Feedback literacy	30	1.78	.363	.048	−1.025

Table 3. Tests of between-subjects effects of ANCOVA for four groups

Source	Type III sum of squares	*df*	Mean square	*F*	Sig.	Partial eta squared
Corrected model	439.404[a]	4	109.851	220.705	.000	.885
Intercept	10.648	1	10.648	21.393	.000	.157
Speaking pretest	32.128	1	32.128	64.549	.000	.360
Groups	427.188	3	142.396	286.093	.000	.882
Error	57.239	115	.498			
Total	4959.891	120				
Corrected total	496.643	119				
Source	Type III sum of squares	*df*	Mean square	*F*	Sig.	Partial eta squared
Corrected model	302.508[b]	4	75.627	176.757	.000	.860
Intercept	12.963	1	12.963	30.298	.000	.209
Feedback pretest	.441	1	.441	1.031	.312	.009
Groups	267.419	3	89.140	208.339	.000	.845
Error	49.204	115	.428			
Total	1666.907	120				
Corrected total	351.712	119				

[a]R squared = .885 (adjusted R squared = .881).
[b]R squared = .860 (adjusted R squared = .855).

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

Table 4. Pairwise comparison through post hoc test

	(I) Groups	(J) Groups	Mean difference (I-J)	Std. error	Sig.	95% confidence interval for difference	95% confidence interval for difference
						Lower bound	Upper bound
Speaking performance	Experimental Group 1	Experimental Group 2	3.746*	.183	.000	3.256	4.236
		Control Group 1	2.669*	.183	.000	2.179	3.159
		Control Group 2	5.185*	.183	.000	4.694	5.677
	Experimental Group 2	Control Group 1	−1.077*	.182	.000	−1.566	−.588
		Control Group 2	1.439*	.182	.000	.950	1.929
Feedback literacy	Control Group 1	Control Group 2	2.516*	.182	.000	2.027	3.006
	Experimental Group 1	Experimental Group 2	3.599*	.180	.000	3.115	4.083
		Control Group 1	1.183*	.172	.000	.721	1.645
		Control Group 2	3.624*	.173	.000	3.160	4.088
	Experimental Group 2	Control Group 1	−2.416*	.172	.000	−2.877	−1.955
		Control Group 2	.025	.171	.896	−.434	.484
	Control Group 1	Control Group 2	2.441*	.169	.000	1.987	2.895

Note. * $p < .001$

Table 4 presents the results of post hoc pairwise comparisons, highlighting significant group differences in both speaking performance and feedback literacy. In terms of speaking performance, the AI-enhanced dialogic group (Experimental Group 1) outperformed all other groups, while the AI-enhanced monologic group (Experimental Group 2) scored higher than the non-AI-monologic group (Control Group 2) but lower than the non-AI-dialogic group (Control Group 1); moreover, Control Group 1 significantly exceeded Control Group 2. For feedback literacy, the AI-enhanced dialogic group (Experimental Group 1) again achieved the highest outcomes, surpassing all groups, while the non-AI-dialogic group (Control Group 1) performed better than both the AI-enhanced monologic group (Experimental Group 2) and the non-AI-monologic group (Control Group 2), with no significant difference observed between Experimental Group 2 and Control Group 2.

4.2. Answering the third research question

The analysis of learners' spoken interactions, based on the integrated coding scheme of Lyster and Ranta's (1997) corrective feedback uptake framework and Storch's (2002) interaction patterns, revealed distinct patterns across the four groups (see Table 5 for more details). Clarification requests and confirmation checks were more prevalent in the dialogic conditions, particularly in the AI-dialogic group, where learners were able to engage in sustained negotiation of meaning with ChatGPT. Uptake with repair also emerged strongly in both AI conditions, with ELSA Speak supporting phoneme-level correction and ChatGPT facilitating reformulation in context. By contrast, uptake without repair and no uptake were more common in the monologic non-AI condition, where limited opportunity for interaction often restricted learners' engagement with corrective feedback. These findings suggest that dialogic environments, especially when supported by AI, can create more space for active learner agency and deeper processing of feedback.

In terms of interaction patterns, collaborative exchanges dominated in dialogic groups, reflecting learners' willingness to co-construct meaning and jointly solve language problems. The expert/novice pattern was observed frequently in the non-AI-dialogic group, where the teacher assumed a guiding role, and in the AI-monologic group, where learners followed system-generated corrections. Dominant/passive interactions occurred mainly in monologic contexts, with learners acknowledging feedback but contributing minimally to the exchange, while dominant/dominant interactions were rare, surfacing occasionally in peer-to-peer disputes in the non-AI-dialogic group. Overall, the coding results highlight how feedback mode (monologic vs. dialogic) and mediation (AI vs. non-AI) influence not only the form of uptake but also the quality of interaction, with dialogic AI feedback providing the richest opportunities for sustained engagement and self-repair.

Clarification requests were most frequent in dialogic groups (AI-dialogic: 18; non-AI-dialogic: 18), while confirmation checks appeared across all groups but were higher in dialogic conditions (AI-dialogic: 22; non-AI-dialogic: 22). Uptake with repair was prominent in both AI conditions (AI-monologic: 30; AI-dialogic: 30), whereas uptake without repair and no uptake occurred mainly in monologic groups (AI-monologic: 15; non-AI-monologic: 15; non-AI-monologic: nine). Self-repair/reformulation was observed primarily in dialogic contexts (AI-dialogic: 20; non-AI-dialogic: 20). Interaction patterns further reflected these trends: collaborative exchanges dominated dialogic groups, expert/novice patterns were common in non-AI-dialogic and AI-monologic groups, dominant/passive patterns were frequent in monologic contexts, and dominant/dominant interactions were rare.

Learners' perceptions of feedback highlighted its usefulness, particularly in terms of clarity, precision, and relevance. Participants appreciated when feedback clearly explained errors and offered alternatives, with one learner noting, "ChatGPT explained why my sentence was wrong, not only that it was wrong" (nine interviewees, AI-dialogic and non-AI-dialogic). Accuracy in identifying pronunciation or grammatical issues was valued, exemplified by an AI-monologic

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

Table 5. Dialogic moves identified across groups

Theme	Subtheme	Description	Quotation	Frequency
Feedback response	Clarification request	Learners requested clarification when the AI/teacher feedback was unclear. Most frequent in dialogic groups, as interaction invites questions.	"Do you mean I should stress the first syllable?"	18 (mostly AI-dialogic, non-AI-dialogic)
Feedback response	Confirmation check	Learners repeated feedback to confirm understanding. Appeared across all groups but most in dialogic ones.	"So, it's went, not goed, right?"	22 (all groups, higher in dialogic)
Feedback response	Uptake with repair	Learners incorporated corrections successfully into their speech. Strong in AI-monologic (pronunciation drills) and AI-dialogic.	"I went to the park yesterday."	30 (esp. AI-monologic, AI-dialogic)
Feedback response	Uptake without repair	Learners acknowledged feedback but did not attempt correction. More common in monologic groups.	"Okay, I see."	15 (mostly AI-monologic, non-AI-monologic)
Feedback response	No uptake	Learners ignored or continued without adjusting. Limited presence, mostly in control groups.	(Learner continues speaking unchanged)	9 (mostly non-AI-monologic)
Feedback response	Self-repair/ reformulation	Learners corrected themselves after feedback without direct prompting. Strong in AI-dialogic where negotiation was emphasized.	"I go – sorry, I went yesterday."	20 (esp. AI-dialogic, non-AI-dialogic)
Interaction patterns	Collaborative	Both participants contributed equally, negotiating and co-constructing meaning. Dominant in dialogic groups.	AI/peer: "What's another way to say it?" → Learner: "Should I use the past tense?"	25 (mostly AI-dialogic, non-AI-dialogic)
Interaction patterns	Expert/novice	One participant guided while the learner followed. Common in teacher-led Zoom (non-AI-dialogic) and AI-monologic.	Peer: "Say went." → Learner: "Okay, went."	21 (AI-monologic, non-AI-dialogic)
Interaction patterns	Dominant/passive	The teacher/AI led feedback with minimal learner participation. More frequent in monologic groups.	Teacher: "Your pronunciation is wrong here." → Learner: "Yes … okay."	14 (mostly AI-monologic, non-AI-monologic)
Interaction patterns	Dominant/ dominant	Both interlocutors tried to control the interaction. Rare but found in some peer–peer exchanges.	Learner: "No, I think I'm right!" → Peer: "No, listen, it's went."	6 (few in non-AI-dialogic)

learner: "ELSA showed the exact sound I mispronounced, which was helpful" (six interviewees). Learners also considered whether feedback aligned with their speaking goals, emphasizing support for fluency, accuracy, or confidence, as reflected in the comment, "The Zoom sessions helped me speak more naturally, not just correctly" (eight interviewees, non-AI-dialogic and AI-dialogic).

The interactional experience and perceived learning impact further shaped learner perceptions. Opportunities to negotiate meaning and clarify feedback were especially appreciated in dialogic contexts, with one learner stating, "I could ask ChatGPT to explain again in simpler words" (12 interviewees, AI-dialogic and non-AI-dialogic). Emotional comfort emerged as significant, with learners feeling safer and more confident in AI-dialogic settings: "I wasn't shy with ChatGPT; it felt safe to make mistakes" (six interviewees). Learners also reported enhanced awareness of language gaps, better uptake and application of feedback, and growth in feedback literacy. For instance, an AI-dialogic learner noted, "ChatGPT made me think about why the change was needed, not just copy it" (six interviewees), while another emphasized immediate application: "After the Zoom session, I could fix my sentence right away because the teacher guided me step by step" (10 interviewees, non-AI-dialogic and AI-dialogic).

5. Discussion

This study set out to investigate the differential effects of dialogic and monologic feedback in AI- and non-AI-mediated environments on L2 learners' speaking proficiency, feedback literacy, and dialogic moves. By adopting a mixed-methods design and systematically comparing four groups, AI-dialogic, AI-monologic, non-AI-dialogic, and non-AI-monologic, the research provided a nuanced account of how feedback mode and mediation interact to shape learners' oral development. Drawing on interactionist (Long, 1996) and sociocultural (Vygotsky, 1978) perspectives, the study not only measured proficiency gains but also analyzed learner perceptions and recorded dialogic behaviors, thereby extending the scope of prior studies that have focused predominantly on monologic teacher-led or AI-based corrective approaches.

The quantitative findings revealed that learners in the AI-dialogic group achieved the greatest gains in both speaking proficiency and feedback literacy, significantly outperforming all other groups. This result can be explained theoretically through the interaction hypothesis, which underscores the importance of negotiation of meaning in facilitating L2 development. ChatGPT, as an AI-dialogic platform, afforded opportunities for sustained conversational exchanges where learners could request clarification, confirm understanding, and engage in self-repair. Such dialogic moves promoted deeper processing of input and greater noticing of interlanguage gaps (Swain, 2000), ultimately leading to measurable improvements in oral proficiency. Practically, this finding highlights the pedagogical potential of conversational AI systems to simulate authentic communicative environments that promote learner agency and reduce the affective barriers often associated with speaking practice.

The second major finding was that the non-AI-dialogic group also demonstrated notable improvement, performing better than both monologic groups, though not to the same extent as the AI-dialogic group. This suggests that dialogic interaction, whether mediated by AI or a human teacher, inherently supports learners' feedback literacy and speaking growth by embedding corrective input within co-constructed discourse. From a sociocultural perspective, teacher-led Zoom sessions provided scaffolding within learners' ZPD, fostering collaborative engagement and active uptake of feedback. This result reinforces prior research showing that dialogic peer and teacher interactions enhance reflection and confidence (Bibauw et al., 2019). In practice, it demonstrates that even without technological enhancement, dialogic feedback remains a powerful mechanism for improving oral proficiency and cultivating feedback literacy.

In contrast, the AI-monologic group improved moderately, with noticeable gains in speaking performance but weaker outcomes in feedback literacy compared to dialogic groups. This pattern

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

aligns with the evaluative and one-way nature of ELSA Speak, which emphasizes phoneme-level precision but provides limited space for negotiation or metacognitive engagement. While such systems are effective for targeted corrective practice, they risk reinforcing passive uptake without fostering learner agency. This finding resonates with critiques of monologic feedback as efficient but limited in promoting sustainable learning (Boud & Molloy, 2013). Nonetheless, in practical terms, the results confirm that AI-monologic tools can serve as valuable supplements for focused accuracy training, particularly in pronunciation-sensitive contexts.

Finally, the non-AI-monologic group exhibited the least improvement across both speaking performance and feedback literacy. Learners in this group often displayed uptake without repair or no uptake at all, reflecting the restricted opportunities for interaction when corrective comments were delivered unidirectionally by the teacher after recordings. Theoretically, this outcome illustrates the limitations of monologic feedback in supporting active noticing and self-regulation, as learners were positioned as passive recipients of evaluative input (Zhu & To, 2022). Practically, the result signals that conventional teacher-led corrective methods, while authoritative and accurate, may not suffice for sustained oral development unless complemented by dialogic strategies that engage learners in meaning-making and active repair.

The qualitative analysis of dialogic moves further enriched these findings by illustrating how learners engaged differently with feedback across conditions. The AI-dialogic group demonstrated the highest frequency of clarification requests, confirmation checks, and self-repair, underscoring the role of interactive AI in fostering collaborative, learner-driven exchanges. Conversely, dominant/passive patterns were more common in monologic groups, reflecting reduced learner participation. These patterns confirm that the quality of interaction is as critical as the accuracy of corrective input, and that dialogic environments, particularly AI-mediated ones, create richer opportunities for collaborative problem-solving and sustained engagement.

6. Conclusion

This study demonstrated that the mode of feedback (dialogic vs. monologic) and the medium of delivery (AI vs. non-AI) play a decisive role in shaping L2 learners' speaking proficiency, feedback literacy, and interactional engagement. The findings showed that AI-dialogic feedback, exemplified through ChatGPT, yielded the greatest improvements, highlighting the pedagogical value of interactive, scaffolded exchanges that promote clarification, self-repair, and learner agency. Non-AI-dialogic feedback also proved effective, reinforcing the enduring importance of co-constructed dialogue in conventional settings. By contrast, monologic feedback, particularly in non-AI contexts, was less impactful, underscoring the limitations of one-way corrective input. Together, these results suggest that fostering dialogic engagement, especially when enhanced by AI, is key to supporting sustainable oral development and empowering learners to take ownership of their feedback processes.

The findings of this study carry both theoretical and practical implications for L2 pedagogy and feedback research. Theoretically, the results reinforce sociocultural and interactionist perspectives by showing that dialogic feedback, particularly in AI-mediated contexts, facilitates deeper engagement with corrective input, scaffolding within the learner's ZPD, and opportunities for negotiation of meaning that advance interlanguage development. Practically, the study suggests that teachers and curriculum designers should prioritize dialogic modes of feedback, whether through AI tools like ChatGPT or conventional platforms such as Zoom, to foster active learner participation and feedback literacy. At the same time, monologic AI applications such as ELSA Speak can be strategically integrated as supplementary tools to strengthen pronunciation and fluency, while conventional non-AI-monologic feedback should be complemented with interactive tasks to enhance learner agency. Collectively, these insights emphasize the need for

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

hybrid pedagogical models that combine the precision of AI-driven correction with the richness of dialogic interaction to maximize learning outcomes.

Despite its contributions, this study is not without limitations. First, the sample size, although sufficient for statistical analysis, was restricted to intermediate-level of Iranian L2 learners within language institutes, which may limit the generalizability of the findings to other proficiency levels, cultural contexts, or educational settings. Second, the intervention was relatively short in duration, spanning only 12 weeks, which constrains the ability to assess the long-term sustainability of dialogic and monologic feedback effects on speaking development and feedback literacy. Third, while the study employed validated instruments and a mixed-methods design, the qualitative coding of dialogic moves inevitably involved a degree of subjectivity, despite the use of established frameworks and interrater checks. Finally, the research examined only one AI application for dialogic feedback (ChatGPT) and one for monologic feedback (ELSA Speak), which, although representative, may not capture the full range of functionalities and learner experiences afforded by other AI tools. Future studies with larger, more diverse populations, extended timelines, and multiple AI platforms are therefore necessary to deepen and broaden the current findings.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/S095834402610055X

Data availability statement. The data that support the findings of this study are not available due to privacy/ethical restrictions.

Authorship contribution statement. Hanieh Shafiee Rad: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Visualization, Validation.

Funding disclosure statement. This research did not receive any specific funding.

Competing interests statement. The author declares no competing interests.

Ethical statement. Ethical review and approval were not required for the study, as per local legislation and institutional requirements. Informed consent was obtained from all participants.

GenAI use disclosure statement. AI assistance was limited to grammar and language editing. No generative AI was used in the research or writing process.

References

Amrate, M., & Tsai, P. (2025). Computer-assisted pronunciation training: A systematic review. *ReCALL*, 37(1), 22–42. https://doi.org/10.1017/S0958344024000181

Bibauw, S., & Dizon, G. (2025). Conversational AI and chatbots. In L. Ohashi, M. Hillis, & R. Dykes (Eds.), *Artificial intelligence in our language learning classrooms* (pp. 19–41). Candlin & Mynard ePublishing. https://doi.org/10.47908/38/1

Bibauw, S., François, T., & Desmet, P. (2019). Discussing with a computer to practice a foreign language: Research synthesis and conceptual framework of dialogue-based CALL. *Computer Assisted Language Learning*, 32(8), 827–877. https://doi.org/10.1080/09588221.2018.1535508

Blake, R. J. (2017). Technologies for teaching and learning L2 speaking. In C. A. Chapelle & S. Sauro (Eds.), *The handbook of technology and second language teaching and learning* (pp. 107–117). John Wiley & Sons. https://doi.org/10.1002/9781118914069.ch8

Boud, D., & Molloy, E. (Eds.). (2013). *Feedback in higher and professional education: Understanding it and doing it well.* Routledge. https://doi.org/10.4324/9780203074336

Cai, Z., Gui, Y., Mao, P., Wang, Z., Hao, X., Fan, X., & Tai, R. H. (2023). The effect of feedback on academic achievement in technology-rich learning environments (TREs): A meta-analytic review. *Educational Research Review*, 39, Article 100521. https://doi.org/10.1016/j.edurev.2023.100521

Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education*, 31(2), 219–233. https://doi.org/10.1080/03075070600572132

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Chinpakdee, M. (2026). Talking feedback: Fostering student feedback literacy through a dialogic approach. *RELC Journal*, 57(1), 158–171. https://doi.org/10.1177/00336882251346206

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

Dai, Y., & Wu, Z. (2022). Mobile-assisted peer feedback on EFL pronunciation: Outcome effects, interactional processes, and shaping factors. *System*, 111, Article 102953. https://doi.org/10.1016/j.system.2022.102953

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 Publishing Corporation.

Dooly, M. (Ed.). (2008). *Telecollaborative language learning: A guidebook to moderating intercultural collaboration online*. Peter Lang.

Du, Q. (2025). How artificially intelligent conversational agents influence EFL learners' self-regulated learning and retention. *Education and Information Technologies*, 30(15), 21635–21701. https://doi.org/10.1007/s10639-025-13602-9

Ellis, R. (2017). Oral corrective feedback in L2 classrooms: What we know so far. In H. Nassaji & E. Kartchava (Eds.), *Corrective feedback in second language teaching and learning: Research, theory, applications, implications* (pp. 3–18). Routledge. https://doi.org/10.4324/9781315621432-2

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/10.1080/02602938.2020.1786497

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

Evers, K., & Chen, S. (2022). Effects of an automatic speech recognition system with peer feedback on pronunciation instruction for adults. *Computer Assisted Language Learning*, 35(8), 1869–1889. https://doi.org/10.1080/09588221.2020.1839504

Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System*, 121, Article 103254. https://doi.org/10.1016/j.system.2024.103254

Forbes, D., Daly, N., & Li, L. (2024). *Designing discussions for online and blended courses: A forum for learning in higher education*. Springer. https://doi.org/10.1007/978-981-97-6196-8

Hou, Z., & Min, S. (2026). Dialogue-based computer-assisted language learning systems for second language speaking development: A three-level meta-analysis. *ReCALL*, 38(1), 40–56. https://doi.org/10.1017/S0958344025100268

Kazemi, M., Saadat, M., Wilson, J., & Rahimi, M. (2025). A comparative study on monologic vs. dialogic technology-mediated feedback modalities: Students' uptake and perspectives. *The JALT CALL Journal*, 21(1), Article 1758. https://doi.org/10.29140/jaltcall.v21n1.1758

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford University Press.

Li, K., & Su, W. (2025). Exploring student roles in mixed-proficiency triads for dialogic peer feedback learning. *Innovations in Education and Teaching International*, 62(3), 909–924 https://doi.org/10.1080/14703297.2024.2347492

Liu, X., Huang, J., Deng, Y., & Spiridakis, J. (2025a). AI versus human assessment in EFL speaking classrooms: A comparative study in China. *Computer Assisted Language Learning*. Advance online publication. https://doi.org/10.1080/09588221.2025.2530555

Liu, X. J., Wang, J., & Zou, B. (2025b). Evaluating an AI speaking assessment tool: Score accuracy, perceived validity, and oral peer feedback as feedback enhancement. *Journal of English for Academic Purposes*, 75, Article 101505. https://doi.org/10.1016/j.jeap.2025.101505

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19(1), 37–66. https://doi.org/10.1017/S0272263197001034

Morgana, V. (2024). Fostering English speaking and writing subskills for the Cambridge B2 First through technology-mediated tasks. *ReCALL*, 36(2), 119–134. https://doi.org/10.1017/S0958344023000277

Myers, T., & Buchanan, J. (2025). Dialogism in feedback literacies: A critical review. *Assessment & Evaluation in Higher Education*, 50(6), 846–860. https://doi.org/10.1080/02602938.2025.2478159

Pellettieri, J. (2000). Negotiation in cyberspace: The role of chatting in the development of grammatical competence. In M. Warschauer & R. Kern (Eds.), *Network-based language teaching: Concepts and practice* (pp. 59–86). Cambridge University Press.

Richards, J. C., & Bohlke, D. (2011). *Four corners level 2 full contact with self-study CD-ROM*. Cambridge University Press.

Shadiev, R., Feng, Y., Zhussupova, R., & Altinay, F. (2026). Effects of speech-enabled corrective feedback technology on EFL speaking skills, anxiety and confidence. *Computer Assisted Language Learning*, 39(4), 878–914. https://doi.org/10.1080/09588221.2024.2430761

Steen-Utheim, A., & Wittek, 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

Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning*, 52(1), 119–158. https://doi.org/10.1111/1467-9922.00179

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 97–114). Oxford University Press.

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press

Timpe-Laughlin, V., Sydorenko, T., & Daurio, P. (2022). Using spoken dialogue technology for L2 speaking practice: What do teachers think? *Computer Assisted Language Learning*, 35(5–6), 1194–1217. https://doi.org/10.1080/09588221.2020.1774904

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Winstone, N., & Carless, D. (2020). *Designing effective feedback processes in higher education: A learning-focused approach*. Routledge. https://doi.org/10.4324/9781351115940

Wu, T.-T., Hapsari, I. P., & Huang, Y.-M. (2025). Effects of incorporating AI chatbots into think–pair–share activities on EFL speaking anxiety, language enjoyment, and speaking performance. *Computer Assisted Language Learning*. Advance online publication. https://doi.org/10.1080/09588221.2025.2478271

Zhan, Y. (2022). Developing and validating a student feedback literacy scale. *Assessment & Evaluation in Higher Education*, 47(7), 1087–1100. https://doi.org/10.1080/02602938.2021.2001430

Zhang, R., & Zou, D. (2022). Types, purposes, and effectiveness of state-of-the-art technologies for second and foreign language learning. *Computer Assisted Language Learning*, 35(4), 696–742. https://doi.org/10.1080/09588221.2020.1744666

Zhu, Q., & To, J. (2022). Proactive receiver roles in peer feedback dialogue: Facilitating receivers' self-regulation and co-regulating providers' learning. *Assessment & Evaluation in Higher Education*, 47(8), 1200–1212. https://doi.org/10.1080/02602938.2021.2017403

About the author

Hanieh Shafiee Rad received her PhD in ELT. Her research interests are educational psychology, computer-assisted language learning (CALL), and L2 teaching methodology. She has published in international journals such as *System*, *Computer Assisted Language Learning*, *Interactive Learning Environment*, and *Language Learning & Technology*.

https://doi.org/10.1017/S095834402610055X Published online by Cambridge University Press