

ARTICLE



Custom GPT as mediator: Dynamic Assessment with beginner KFL learners

Minjin Kim, The University of Mississippi

Abstract

This study examines how a custom GPT-based chatbot can mediate learner development within the Zone of Proximal Development (ZPD) through dynamic assessment (DA) with beginner-level learners of Korean as a Foreign Language. The model was fine-tuned using OpenAI's My GPT platform, with a custom prompt specifying graduated mediation, responsive behavior guidelines, and target grammar points. Specifically, the study investigates how GPT operationalizes scaffolding processes in text-based dialogue by sustaining interaction, providing form-focused feedback, and adjusting support contingent on learner responsiveness. Ten English-speaking students in a Korean course at a U.S. university interacted with the chatbot weekly over four weeks. Qualitative analysis of 280 learner-GPT turns identified three mediation types: conversational, instructional, and developmental. Through these, the chatbot maintained natural and level-appropriate dialogue, delivered graduated mediation aligned with learner responsiveness, and used accurate learner responses as springboards to guide movement from the Zone of Actual Development toward the ZPD. Complementary quantitative measures showed higher uptake rates and significant gains in mean length of sentence and lexical diversity. These findings suggest that large language models, when carefully tuned, can emulate core principles of Vygotskian mediation and foster human-AI co-construction of learning within scaffolded interaction.

Keywords: dynamic assessment; GPT; Korean as a Foreign Language; AI-assisted language learning

Language(s) Learned in This Study: *Korean*

APA Citation: Kim, M. (2026). Custom GPT as mediator: Dynamic Assessment with beginner KFL learners. *Language Learning & Technology*, 30(1), 1–27. <https://doi.org/10.64152/10125/73669>

Introduction

Over the past two decades, Dynamic Assessment (DA) has emerged as a theoretically grounded and pedagogically powerful approach to evaluating and promoting second language (L2) development (Ableeva, 2010; Aljaafreh & Lantolf, 1994; Antón, 2009; Lantolf & Poehner, 2004; Poehner, 2005, 2008). Rooted in Vygotsky's sociocultural theory, DA shifts the focus from what learners can do independently to what they can achieve with support in their Zone of Proximal Development (ZPD), i.e., the distance between the Zone of Actual Development (ZAD) and what can be reached through mediation (Lantolf & Poehner, 2004; Vygotsky, 1978). By offering graduated, individualized support, DA reveals learners' developmental readiness and enables instructors to adjust feedback in real time (Lidz, 1991; Poehner, 2008). In L2 education, where language learning is inherently interactive and developmental, such tailored mediation has proven especially effective (Ableeva, 2010; Antón, 2009).

However, face-to-face DA is resource-intensive and difficult to scale beyond one-on-one settings. To address this, researchers have developed computerized DA (C-DA) formats to improve accessibility and efficiency (Mehri Kamrood et al., 2021; Poehner & Lantolf, 2013; Yang & Qian, 2017), typically following an interventionist model with pre-scripted feedback. While this structure allows for consistent implementation, it limits the dialogic and contingent nature of interactionist DA (Aljaafreh & Lantolf, 1994; Feuerstein et al., 2002). Recent advances in large language models (LLMs), such as ChatGPT, offer

new possibilities for expanding C-DA into more responsive formats. LLMs can generate real-time, context-sensitive mediation based on learner input, potentially blending the structured reliability of interventionist DA with the adaptive flexibility of an interactionist approach (Brown et al., 2020; Pecher et al., 2024). This potential is particularly promising when models are fine-tuned to specific instructional contexts (Almasre, 2024; Kim & Lu, 2024).

While research on AI-assisted language education has grown rapidly in recent years (e.g., Ding & Yusof, 2025; Shin & Choi, 2025; Wang et al., 2024), much of it has centered on general feedback, assessment, or tutoring rather than the principle-driven, graduated mediation central to DA. Existing C-DA and AI-mediated DA have mainly addressed receptive skills such as reading and listening, where learner responses are more easily pre-scripted (e.g., Ableeva, 2010; Yang & Qian, 2017), or, in many cases, have focused on widely taught languages like English, French, and German (e.g., Jeon, 2023; Jili et al., 2025; Özturan & Shrestha, 2025). Research that operationalizes DA within open-ended, dialogic interaction using custom GPT models remains scarce, and empirical evidence from less commonly taught languages (LCTLs) such as Korean is still emerging. Against this backdrop, the present study investigates the feasibility of GPT-mediated DA with beginner-level Korean as a Foreign Language (KFL) students. By examining Korean as a representative LCTL, the study highlights the broader applicability of GPT-mediated DA for other LCTLs and for languages underrepresented in English-dominant LLM training data. Specifically, by showing how GPT mediates conversation, provides graduated feedback, and supports learner development, this study advances theoretical, methodological, and pedagogical understandings of how AI can operationalize ZPD-based scaffolding and mediation principles in scalable, technology-supported learning contexts.

Literature Review

Dynamic Assessment

Sociocultural Theory (SCT), originating from Vygotsky and further developed by scholars such as Lantolf and Poehner, frames learning as a socially mediated process in which cognitive development emerges through interaction with more capable interlocutors and participation in contextually situated activities (Lantolf, 2000; Lantolf & Poehner, 2004; Vygotsky, 1978). At the core of SCT is the notion of praxis, the integration of theory and practice, where psychological theory aims not only to explain cognition but to promote its development through education (Lantolf, 2009; Lantolf & Poehner, 2011). DA enacts this principle by transforming theoretical insights about the ZPD into instructional actions that support learning in real time (Poehner, 2005, 2008). DA operationalizes the ZPD through graduated, responsive mediation that reveals and facilitates learners' emerging control over linguistic forms and their meanings (Lantolf & Poehner, 2011).

DA is typically implemented through two main approaches: interventionist and interactionist (Lantolf & Poehner, 2004, 2011, Poehner, 2008). Interventionist DA is characterized by standardized tasks and pre-scripted mediation designed in advance based on anticipated learner difficulties. Mediation progresses from implicit to explicit through a series of hints, prompts, or guiding questions until the learner either arrives at the correct response or is provided with the correct answer. This structured format allows for easier comparison of learner performance, often by counting the number of prompts needed to reach a solution (Brown & Ferrera, 1985; Budoff, 1987). In contrast, interactionist DA involves open-ended, dialogic mediation tailored to learners' needs in real time (Feuerstein et al., 2002; Poehner, 2005, 2008). Rather than following a fixed script, the mediator adjusts feedback dynamically to help learners move beyond their current abilities through co-constructed interaction. For example, Poehner (2005) demonstrated this with flexible use of hints, prompts, questions, and explanations based on mediator judgment and learner requests. Although more difficult to standardize and compare across learners, this approach is better suited to promoting deeper development, especially in L2 learning, where responsiveness and negotiation of meaning are essential (Ableeva, 2010; Feuerstein et al., 2002; Poehner, 2005, 2008).

Poehner (2005) was among the first to systematically implement DA in an L2 pedagogical task, using a graduated prompt scale to support six advanced French learners in oral proficiency. His interactionist approach demonstrated that DA not only reveals current performance but also learners' developmental potential, while fostering metacognitive awareness and promoting co-constructed, self-regulated learning. Building on this work, Ableeva (2010) applied interactionist DA to listening comprehension with intermediate French learners, using dialogic mediation to diagnose and address comprehension breakdowns in real time. Her study underscored the instructional value of DA and offered a taxonomy of strategies tailored to listening tasks. DA has since been extended to reading and writing (e.g., Antón, 2009; Kozulin & Garb, 2002), and its classroom applicability has been explored in multiple studies (e.g., Davin, 2013; Lantolf & Poehner, 2011; Poehner, 2009). For instance, Lantolf and Poehner (2011) documented how a Spanish teacher integrated interventionist DA into daily practice using co-constructed mediation menus. Although designed collaboratively, implementation was fully teacher-led, exemplifying Vygotskian praxis. Over time, the learners internalized the support, requiring fewer prompts and anticipating mediation moves, which is evidence of progress in the ZPD and the scalability of DA for classroom-based L2 development.

Technology-mediated DA

While DA offers clear benefits for supporting and assessing L2 development, its implementation in classroom or large-scale assessment often faces practical constraints, such as its time-intensive nature, limited scalability, and the need for trained mediators (Antón, 2009; Rassaei, 2023; Teo, 2012). In response, researchers have developed Computerized DA (C-DA) as an alternative delivery format that seeks to retain the core principles of DA (e.g., Bakhoda & Shabani, 2019; Mehri Kamrood et al., 2021; Poehner & Lantolf, 2013; Poehner et al., 2015; Teo, 2012; Yang & Qian, 2017). For instance, Teo (2012) addressed the challenge of large class sizes in Taiwanese universities by implementing a C-DA program in an English class with 68 freshmen. Using Viewlet Quiz 3 software, the author designed four levels of computerized mediation to support inferential reading development. The system recorded learner responses, mediation triggers, and post-task reflections. Pre- and posttest comparisons confirmed significant improvement in comprehension, while qualitative data revealed increased metacognitive awareness. Similarly, Rassaei (2023) demonstrated that mobile-mediated DA using WhatsApp in the classroom not only improved learners' ability to produce polite English requests but also fostered agentive participation and peer-mediated learning. These studies demonstrate that DA principles can be effectively integrated into technology-enhanced classroom environments.

Poehner and Lantolf (2013) developed a series of C-DA systems targeting listening and reading comprehension in L2 Chinese and French within a large-scale university assessment context. Using scripted multiple-choice items paired with graduated mediation ranging from implicit to explicit feedback, they introduced the Learning Potential Score (LPS), calculated by comparing unmediated and mediated performance (Kozulin & Garb, 2002). Results from over 160 learners revealed that students with similar unmediated scores often differed significantly in their LPS, indicating varied responsiveness to mediation and revealing latent learning potential undetectable through static measures. Building on this work, Poehner et al. (2015) applied a similar C-DA design to Chinese listening tasks and found that LPS more reliably predicted transfer task performance than raw test scores, further supporting C-DA's diagnostic value. Yang and Qian (2017) extended this line of research to English reading comprehension with 68 Chinese EFL college freshmen, employing a sandwich design (pretest–C-DA–posttest) and delivering four levels of mediation through a custom program. Beyond significant posttest gains, the study documented increases in metacognitive awareness and strategy use. Learners with identical pretest scores responded differently to mediation, underscoring the need for ZPD-sensitive assessment and highlighting the dual diagnostic and developmental functions of C-DA.

Together, these studies demonstrate that technology-assisted DA can retain the core benefits of DA, namely, diagnosing learning potential and tailoring support, while overcoming the logistical constraints of face-to-face interaction. Notably, most systems primarily rely on standardized, pre-scripted mediation to

ensure consistency, aligning more closely with an interventionist approach (e.g., Mehri Kamrood et al., 2021; Poehner & Lantolf, 2013; Poehner et al., 2015). The tendency of these systems to target receptive skills such as reading and listening, where learner responses are more predictable and mediation more readily programmable, reflects the technical practicality of designing support for tasks with constrained input–output structures.

Custom GPT as a DA Mediator

Recent advances in AI have shown the potential for more flexible and adaptive DA. One such approach is chatbot-assisted DA (CA-DA), which utilizes AI chatbot to provide immediate, scaffolded responses that simulate the support of a human mediator (e.g., Jeon, 2023; Jili et al., 2025). Jeon (2023) piloted a CA-DA system designed to support L2 English vocabulary learning among Korean elementary school learners. The chatbot, developed using Dialogflow, delivered graduated mediation through four levels of prompts based on pre-scripted learner response scenarios. Findings revealed that CA-DA, in text-based conversations, promoted vocabulary uptake, increased learner independence, and provided diagnostic insights into learners' lexical development. The study demonstrated the feasibility of CA-DA for receptive and productive vocabulary skills in classroom. Notably, the research emphasized Dialogflow's accessibility for non-programmer teachers, making CA-DA a practical option for classroom-based implementation. Building on this line of research, Jili et al. (2025) also used Dialogflow to develop two chatbots – one dynamic and one non-dynamic – to support Chinese EFL undergraduates' argumentative writing. Although no significant differences emerged in writing scores, learners in the CA-DA condition reported greater support in idea generation and perceived the chatbot as an engaging interlocutor, demonstrating the pedagogical promise of automated mediation for complex writing tasks.

Currently, OpenAI offers a more pedagogically flexible alternative for implementing DA compared to pre-scripted C-DA and earlier AI-based systems. GPT models, which power ChatGPT, draw on a large-scale, pretrained language model that generates more adaptive and context-sensitive responses based on users' input in real time (Brown et al., 2020; Pecher et al., 2024). This capacity is likely to enable the models to (1) deliver contingent mediation aligned with learners' moment-to-moment needs, precisely the kind of responsiveness valued in interactionist DA, (2) flexibly handle unexpected learner responses which are more common in productive skills and often fall outside the designer's original expectations, and (3) generate scaffolded, culturally relevant feedback based on a wide range of linguistic input, enhancing engagement by drawing on topics relevant to learners' interests and experiences. Such dialogic strengths of ChatGPT, including its responsiveness, contextual awareness, and capacity to simulate pedagogically meaningful conversational flow, have been highlighted in prior work (e.g., Ray, 2023). Studies in language education have further shown that these interactional affordances of ChatGPT can enhance learner engagement and feedback quality in L2 writing contexts (e.g., Alsaweed & Aljebreen, 2024; Teng & Huang, 2025).

Furthermore, ChatGPT can be fine-tuned through natural language prompt engineering and few-shot learning (Lee et al. 2024). Prompt engineering refers to the process of carefully designing the input prompts provided to language models to elicit specific or more accurate responses (Brown et al., 2020). Few-shot learning involves supplying the model with a small number of example input-output pairs to guide its responses. In applied linguistics, recent research has demonstrated the benefits of fine-tuning GPT models for both research and assessment purposes. For example, Kim and Lu (2024) found that a fine-tuned GPT model, using a specific prompt and a relatively small number of examples compared to earlier machine learning models, achieved better performance in identifying rhetorical functions in academic writing. Similarly, Liu et al. (2025) reported that a fine-tuned GPT model showed great consistency and alignment with human raters in automated essay scoring, demonstrating its potential to enhance assessment validity and fairness. These studies suggest that fine-tuning can help adapt GPT output to specific pedagogical or evaluative contexts, combining the structured reliability of rule-based systems with the adaptive flexibility of LLMs to handle both expected and unexpected responses. In the context of DA, recent work has begun to explore how ChatGPT can function as a DA mediator. Özturan

and Shrestha (2025) applied prompt engineering to implement ChatGPT-3.5 as a DA mediator in an EFL argumentative writing task. Two adult Turkish learners interacted with the model over four weeks using researcher-designed prompts guiding implicit-to-explicit feedback. The study showed that the model sustained dialogic mediation and fostered learner development, indicating that well-designed prompts can operationalize interactionist DA principles in an AI-mediated environment.

Additionally, OpenAI now offers easy customization through a user-friendly platform, *My GPT*, which enables creators to modify system behavior using natural language prompt engineering, few-shot exemplars, and direct uploads of instructional documents, all without requiring any programming expertise (Almasre, 2024; Pandya & Holia, 2023). Once configured, the custom GPTs operate as fine-tuned models, enabling end users to interact naturally without needing specific prompts or examples (OpenAI, 2025). Although this specific form of customization has rarely been explored within language education, emerging research from fields such as health education, customer service, and policy communication points to its immense potential. For example, Almasre (2024) demonstrated how a custom GPT could be configured for university-level advising by uploading institutional policy documents and academic guidelines, enabling the model to respond accurately to students' domain-specific queries. Similarly, Pandya and Holia (2023) developed a customer support GPT using LangChain by scraping content directly from an organization's website, thereby enabling responsive and context-aware interaction with no manual coding. These studies highlight the model's capacity to be adapted for specialized communicative contexts through prompt and document-based inputs. Taken together, custom GPT models offer a promising means of delivering DA for productive skills; they can accommodate unpredictable learner output while remaining easily customizable.

Although AI-powered chatbots have been widely adopted across skills and contexts (e.g., Ding & Yusof, 2025; Shin & Choi, 2025; Wang et al., 2024), most have not employed the DA approach and have primarily targeted English, with research on LCTLs such as Korean only beginning to emerge (e.g., Shin & Choi, 2025). Likewise, recent DA-oriented studies integrating AI tools have focused almost exclusively on English vocabulary learning or writing (e.g., Jeon, 2023; Jili et al., 2025; Özturan & Shrestha, 2025). As a result, empirical work implementing DA's graduated mediation in open-ended, dialogic interaction with custom GPT models, particularly beyond English-dominant contexts, remains limited.

To address these gaps, the present study investigates how a custom GPT-based text chatbot delivers DA to beginner-level KFL learners in a classroom setting. The model was designed to generate developmentally appropriate, curriculum-bound prompts and to deliver graduated mediation aligned with DA principles. The model was designed to co-construct meaning with learners through graduated mediation contingent on their responsiveness, with support progressing from implicit prompts toward more explicit guidance as learners displayed growing control over target forms. The following research questions guide the inquiry: (1) How does the custom GPT chatbot mediate conversation during DA tasks with beginner KFL learners? (2) How does it provide graduated mediation? (3) How does it promote learner development?

In this study, learner development is used as an overarching construct that refers to learners' movement from the ZAD toward the ZPD through measurable responsiveness to graduated mediation, together with accompanying changes in language use over time.

By demonstrating the feasibility of GPT-mediated DA, this study contributes to research on AI-assisted language learning and advances the broader fields of C-DA and CA-DA. Importantly, the findings also demonstrate the viability of AI-mediated learning in LCTLs, which are comparatively underrepresented in LLM training data. The findings offer theoretical insight into how mediation can be distributed across human and AI agents, methodological guidance for customizing and implementing AI chatbots for DA in LCTL contexts, and pedagogical implications for scalable, individualized mediation in LCTLs.

Methodology

Participants

Ten undergraduate students enrolled in a second-semester Korean language course (Beginner II) at a U.S. university participated in the study. The GPT-mediated DA sessions were conducted between April and early May during the ongoing course, which was taught by the researcher. By the time of data collection, participants had received approximately seven months of formal Korean instruction across the two-semester sequence (Beginner I-II). Based on initial GPT interactions, class performance, and instructional hours, their proficiency was estimated at the Novice-Mid to Novice-High range according to the American Council on the Teaching of Foreign Languages (ACTFL, 2024) proficiency guidelines. All were native English-speaking females ($M = 19.22$ years, $SD = 1.30$) from diverse majors, including the social sciences, health-related fields, and humanities. One participant reported limited heritage exposure to Korean. The class happened to consist entirely of female students that semester, and participation was voluntary, and no selection by gender occurred. Five participants reported minimal prior self-study, but all began formal Korean study only after enrolling in the course sequence. This study received IRB approval from the University of Mississippi (Protocol 25x-248).

Custom GPT

Custom GPTs, also known as *My GPT*¹, offer a uniquely accessible fine-tuning process that does not require programming expertise (OpenAI, 2025). The platform allows users to upload relevant documents, define the model's behavior, and embed few-shot examples through natural language interfaces. Although this feature is available only to paid users, end users can access custom GPTs via shared links with a free ChatGPT account.

The custom GPT was developed to scaffold beginner-level Korean learners' language development through text-based interaction. Its linguistic output was restricted to vocabulary and grammar previously introduced in class, based on uploaded .txt files and lecture slides, ensuring alignment with learners' potential ZPD. The model was prompted to deliver graduated mediation following DA principles, progressing from implicit to explicit feedback when learner errors occurred (see Table 1).

Table 1

Graduated Mediation Plan Adapted from Poehner (2005)

Level	Prompt
1. Request for repetition	Requests repetition to prompt self-correction (e.g., Can you say that again?)
2. Request for verification	Highlights potential error by repeating learner utterance as question (e.g., “[incorrect]... Is it correct?”)
3. Offering a choice	Offers two choices to guide toward correct form (e.g., “[incorrect] or [correct]?”)
4. Providing correct response	Gives correct form with brief explanation and asks for repetition

Instead of relying on a rigid decision tree, the mediation structure was embedded as natural language instructions with ten few-shot examples (five error-free and five containing learner errors). While the model instructions were written in English, all learner-facing prompts were in Korean, and the ten example dialogues provided to the model were also written in Korean. To support affective engagement, the GPT was assigned the role of a friendly Korean tutor who provides encouragement and empathy, with occasional L1 use permitted when communication broke down. Weekly prompts were updated to reflect new grammar targets, including obligation (-아/어야 돼요), reason (-어서, -이러서, 때문에), and inability/prohibition

(-지 못해요, -지 마세요). The model was piloted with three advanced learners to refine the prompts. GPT-4o was used, and [Appendix A](#) provides the full prompt, example dialogues, and interface configuration for reproducibility.

Procedure

Prior to the main data collection, participants completed a background survey about their name, age, first language, academic major, motivation for learning Korean, and previous Korean learning experience. Over four weeks, participants engaged in GPT-assisted DA sessions once per week during regular class sessions. Students accessed the custom GPT models using their personal devices, either their smart phone or laptop and participated in one-on-one, text-based conversations in Korean.

Before each session, participants were given brief instructions to ensure smooth interaction. For example, they were asked to begin the conversation saying “안녕하세요” [hello] or by using conversation starters (example sentences provided on the platform), and the GPT was tuned to guide the conversation with questions that encouraged the use of target grammar points in the context of familiar topics. Learners were also encouraged to use “무슨 말이에요?” [What does that mean?] if they did not understand the chatbot’s response. Each interaction lasted approximately 15 minutes, during which students were encouraged to converse and respond freely to GPT’s mediations. At the end of each session, participants were asked to copy and paste the full conversation and submit the transcript electronically. Because the sessions were implemented as part of regular classroom instruction, attendance and technology-related issues occasionally affected participation, resulting in 33 transcripts in total.

Data Analysis

To address the three research questions, all mediational turns were coded using an analytical framework that categorized the GPT’s strategies into three overarching types: Conversational, Instructional, and Developmental mediation. Each category directly corresponds to one aspect of the research questions: (1) how the GPT sustains interaction, (2) provides graduated feedback, and (3) supports learner development. The framework, grounded in DA principles (Lantolf & Poehner, 2011; Poehner, 2008), was refined inductively to capture patterns emerging from the data. [Figure 1](#) presents a simplified overview of GPT-assisted mediation strategies. The full coding scheme, including mediation categories, operational definitions, representative examples, and clarification of boundaries for codes that may appear conceptually similar, is provided in [Appendix B](#).

Figure 1

Overview of GPT-Assisted Mediation Strategies

Mediational Strategies		
Conversational Mediation Conversational continuation Response acknowledgement	Instructional Mediation Request for repetition Request for verification Provision of choice Provision of correct response Affective reinforcement	Developmental Mediation Reformulation Metalinguistic cueing Request for explanation Provision of example Provision of explanation Provision of translation Elicitation for reinforcement Form-focused expansion

Conversational mediation involves interactional and affective scaffolding that sustains engagement during dialogue (Ohta, 2001). Instructional mediation entails form-focused and diagnostic feedback that helps learners notice and repair linguistic errors through graduated prompts guiding attention to target forms and accuracy (Poehner, 2008). Developmental mediation provides conceptually oriented, higher-order assistance that enables learners to understand underlying linguistic principles and extend their control from the ZAD toward the ZPD (Lantolf & Poehner, 2011). While conversational mediation maintains interactional flow, instructional and developmental mediation promote linguistic awareness and restructuring within the learner's ZPD.

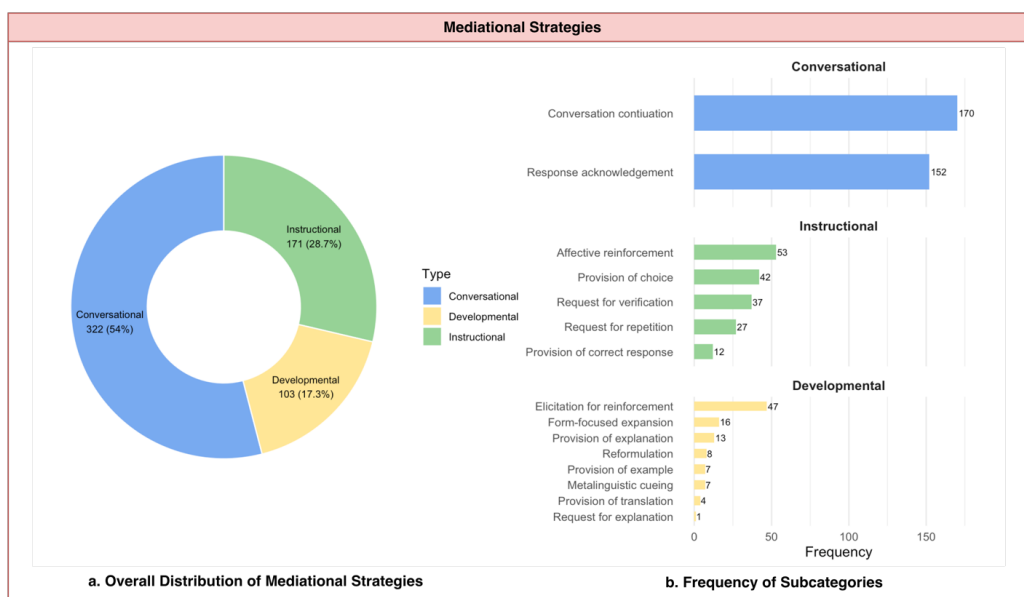
To complement the findings for Research Question 3, performance between the first and last sessions was compared. From a DA perspective, two process indicators were analyzed: uptake rate and steps to success (responsiveness to mediation). Additionally, two outcome measures were included to complement the DA indicators: syntactic complexity (mean length of sentence, MLS) and lexical diversity (root type–token ratio, RTTR). Uptake was coded at the episode level (target-like production = success), with uptake rate defined as the proportion of successful episodes per session; steps to success indexed the number of mediational prompts required. Because uptake rate violated normality (Shapiro–Wilk $p < .05$), Wilcoxon signed-rank tests were used, whereas steps to success, MLS, and RTTR met normality assumptions and were analyzed with paired-sample t-tests. MLS and RTTR were computed from task transcripts in R (R Core Team, 2024) using customized scripts, and all statistical analyses were conducted in R.

Results and Discussion

The dataset comprised 33 transcripts totaling 280 learner–GPT turns, defined as single learner–GPT exchanges. A total of 596 examples of using mediation strategies were identified, averaging 2.1 strategies per turn. Figure 2 shows the overall distribution of the three mediation types and their subcategories. Conversational mediation was most frequent (322, 54.0%), followed by instructional (171, 28.7%) and developmental (103, 17.3%). This distribution suggests that GPT effectively balanced conversational naturalness with pedagogical depth, engaging in instructional and developmental mediation while sustaining meaning-oriented interaction.

Figure 2

Mediational Strategies Used in GPT-Assisted DA



GPT as a Conversational Mediator

First, GPT functions as a conversational mediator by sustaining dialogue and actively accepting responses. GPT consistently sustained the flow of dialogue in nearly every turn (170 instances, 28.5%), except when providing graduated support. GPT generated questions closely aligned with classroom content (e.g., Korean class, campus life, weekends, and Korean pop culture), often eliciting target grammar forms such as necessity, reason, and inability, either through initial prompts or by building on learner responses. The system effectively integrated form-focused practice into contextually meaningful discourse and rarely introduced grammatical structures or vocabulary beyond the scope of the uploaded materials and prompts, thus maintaining linguistic accessibility and minimizing the risk of learner frustration. These patterns reflect prior findings on the value of familiar content in reducing cognitive overload for beginners (Huang et al., 2022; Jeon, 2023) and align with interactionist DA principles, where mediation is co-constructed through meaningful, ZPD-sensitive dialogue (Poehner, 2005, 2008).

Another notable mediation type was “Response acknowledgement,” which accounted for 152 instances (25.5%). This strategy involved GPT acknowledging and reacting positively to learner responses in ways that extended interactional sequences and fostered learner engagement. Such responses included expressions of agreement (e.g., “저도 좋아해요!” [I like it too!]), empathetic reactions (e.g., “아이고~ 우산 없어요?” [Oh no, you didn’t bring an umbrella?], “정말 불편하겠어요” [That must be so inconvenient]), and positive reinforcement (e.g., “오~ 멋있어요” [Wow, that’s great!], “맛있겠다!” [That sounds delicious!]). These responses served critical affective and interactional functions by validating learner contributions, conveying empathy, and maintaining the conversational flow. Notably, this type of mediation frequently occurred even when learner errors were present, often preceding any form-focused intervention. [Excerpt 1](#) illustrates how GPT sustains interaction by acknowledging learner responses while simultaneously initiating mediation in response to learner errors.

Excerpt 1

Co-constructing Meaning through Affective and Linguistic Scaffolding

Excerpt 1. Co-constructing Meaning through Affective and Linguistic Scaffolding	
1. ChatGPT: 무슨 영화나 드라마 봤어요?	What movie or TV show did you watch?
2. Learner: La La Land예요. La La Land 진짜 좋아요!	It's La La Land. La La Land is learry good!
3. ChatGPT: 와~ La La Land 봤어요? 재밌지요? 😊 그런데... “진짜 좋아요”... 맞아요? 다시 말해 줄래요?	Wow, you watched La La Land? It's fun, right? 😊 But... “La La Land is learry good”... is that correct? Can you say that again?
4. Learner: La La Land 아주 좋았어요!	La La Land was very good!
5. ChatGPT: 네~ 아주 좋아요! 😊 저도 그 영화 좋아해요. 그럼, 이번 주말에 뭐 할 거예요?	Good job! 😊 I liked that movie too. Then, what are you doing this weekend?

In line 1, GPT initiates the interaction with an open-ended question grounded in familiar topics, inviting the learner to share a personal experience. In line 2, the learner responds enthusiastically but produces a lexical error (“진짜” instead of “진짜” [really]) and a potentially incongruent tense (“좋아요” [is good]), which may not fully align with the past-tense context of the conversation. Rather than correcting the error directly, GPT in line 3 first acknowledges the learner’s contribution with an affirming statement (“재밌지요?” [It’s fun, right?]), demonstrating affective support, before gently requesting for verification and repetition (“맞아요? 다시 말해 줄래요?” [Is that correct? Can you say that again?]). In line 4, the

learner revises the original utterance by replacing the problematic “친자” [learry] with another intensifier “아주” [very], thereby partially addressing the lexical issue through substitution rather than direct correction. The tense is also shifted to past (“좋았어요” [was good]), which could be contextually more appropriate given the reference to a previously watched movie. In line 5, GPT does not treat the substitution of “아주” for “진짜” as a target deviation. Instead, it sustains the interaction through positive reinforcement (“네~ 아주 좋아요!” [Good job!]), empathetic engagement (“저도 그 영화 좋아해요.” [I like that movie too]), and a forward-moving prompt about weekend plans. Notably, the use of positive emojis alongside affirming language further reinforces affective scaffolding by signaling encouragement and interpersonal warmth.

As shown in the excerpt, GPT demonstrated four main features of mediation as a conversational mediator. First, it provided scaffolded support not as isolated corrections but as feedback woven into natural, coherent exchanges. This support adjusted flexibly to each learner’s response, showing contingent scaffolding within the learner’s ZPD, a core idea of interactionist DA (Poehner, 2008). Second, GPT offered affective mediation through positive feedback, backchanneling, and empathetic language, often using emojis. These responses likely reduced anxiety and encouraged participation, helping learners enter their ZPD more easily (Ohta, 2001). Third, GPT used delayed mediation by prioritizing meaning before form. This sequence allowed learners to clarify meaning first and showed sensitivity to their level of readiness, aligning with dialogic DA’s emphasis on socially contingent mediation (Lantolf & Poehner, 2011). Lastly, when learners avoided difficult forms, GPT acknowledged their ideas and sustained interactional flow. Such adaptation matched the learner’s ability level, though too much tolerance might risk fossilization (Long, 1996). Overall, the findings show that AI can offer human-like scaffolding and highlight the importance of prompt designs that balance affective and form-focused support (Huang et al., 2022; Kim & Lu, 2024).

GPT as an Instructional Mediator

GPT also serves as an instructional mediator by delivering graduated mediation tailored to learners’ responsiveness. Among instructional mediation strategies (see [Figure 2](#)), requests for repetition occurred 27 times (4.5%), requests for verification 37 times (6.2%), provision of choice 42 times (7.0%), and provision of correct response 12 times (2.0%). Affective reinforcement following error correction functioned as an emotional scaffold and appeared 53 times (8.9%). [Excerpt 2](#) illustrates a sequence of graduated mediation leading to successful learner repair.

In line 1, GPT initiates the interaction with two open-ended questions that prompt the learner to produce extended output. In line 2, the learner responds positively but includes two issues: the unnecessary use of the dative particle “한테” [to (someone)] and an ungrammatical construction “일없어요” [there is no work], which misaligns with the past temporal context implied by the previous weekend reference. Rather than explicitly correcting the errors, GPT first responds with a positive reaction, “아~ 좋아요!” [That’s wonderful!], acknowledging the learner’s report of having had no work over the weekend. This is followed by a request for verification and repetition, “맞아요? 다시 말해 줄래요?” [Is that correct? Can you say that again?], which encourages the learner to self-monitor. This move reflects a combination of the first and second levels of mediation, functioning as an indirect prompt. In line 4, the learner partially revises the utterance, removing the unnecessary particle but retaining the tense inconsistency. GPT responds in line 5 with a more explicit prompt, offering two similar options differing in tense, thereby scaffolding the learner’s ability to recognize and apply the appropriate temporal form. In line 6, the learner correctly chooses the past tense “일 없었어요” [There was no work] resolving the remaining issue. GPT concludes in line 7 with multiple layers of reinforcement: explicit affirmation “맞아요!” [That’s right!] and praise “잘했어요!” [Well done!]. This sequence highlights GPT’s dual role as instructional and affective mediator and exemplifies movement through the learner’s ZPD, with graduated mediation adjusted to the learner’s partial understanding.

Excerpt 2

Progression of Graduated Mediation Leading to Successful Repair

Excerpt 2. Progression of Graduated Mediation Leading to Successful Repair	
1. ChatGPT: 오늘 기분 어때요? 주말에 뭐 했어요?	How are you today? What did you do over the weekend?
2. Learner: 저는 좋아요! 그리고 주말에 한테 일없어요.	I'm good! And on the weekend, to (someone), there is no work.
3. ChatGPT: 아~ 좋아요! 그런데... “주말에 한테 일없어요”... 맞아요? 다시 말해 주세요?	That's wonderful! But... “On the weekend, to (someone), there is no work”... Is that correct? Can you say it again?
4. Learner: 주말에 일없어요.	There is no work on the weekend.
5. ChatGPT: ‘일없어요’ 맞아요? 아니면 ‘일 없었어요’일까요?	Which one do you think is correct, ‘There is no work’ or ‘There was no work’?
6. Learner: 일 없었어요! 감사합니다!	There was no work! Thank you!
7. ChatGPT: 네~ 맞아요! 잘했어요! 주말에 일 없었어요.	Yes, that's right! Well done! There was no work over the weekend.

GPT also demonstrated sensitivity to each learner's ZPD, adapting its mediation to individual performance. [Excerpts 3](#) and [4](#) illustrate two distinct mediation trajectories in response to similar learner errors. [Excerpt 3](#) illustrates GPT's graduated mediation through progressively more explicit support in response to learner persistence. The interaction begins with GPT asking a familiar question that elicits the target grammar form “-아/어야 돼요” [have to] (Line 1). The learner responds with “석제해돼요” (Line 2), an orthographic and morphological error intended to express “숙제해야 돼요” [I have to do homework]. Rather than providing immediate correction, GPT issues a request for repetition (Line 3), inviting the learner to self-monitor. When the learner repeats the same error (Line 4), GPT escalates to a prompt providing a choice (Line 5), helping the learner identify the correct form. The learner partially repairs the utterance (Line 6), and GPT then provides the correct form as a recast with praise and requests repetition (Line 7). The learner successfully produces the target form (Line 8), and GPT closes the sequence with affirming feedback (Line 9).

In contrast, [Excerpt 4](#) shows GPT's immediate modeling in response to learner confusion. The interaction begins with GPT asking, “뭐 공부해야 돼요?” [What do you have to study?]. In response, the learner produces a similar orthographic error, “한곡어” instead of “한국어” [Korean language] (Line 2). In Line 3, GPT responds with a recast, “오~ 한국어 공부해요?” [Oh, you're studying Korean?], which acknowledges the learner's intended meaning while implicitly supplying the correct form. This is immediately followed by a request for repetition, “다시 한번 말해 주세요?” [Could you say that again?], encouraging self-repair. In line 4, however, the learner responds with “몰이요” (a misspelling of “몰라요,” roughly rendered as “I no know...”), signaling clear breakdown. In Line 5, GPT shifts to an explicit model – “괜찮아요! 😊 ‘한곡어’ 아니고... ‘한국어’예요. 다시 말해 볼까요?” [That's okay! 😊 It's not 'Koroan'—it's 'Korean.' Let's try it together.] – combining affective support with the correct form and a repetition request. This immediate shift to direct modeling reflects GPT's recognition of the learner's breakdown and its effort to scaffold repair through explicit mediation. The learner's uptake in Line 6 (“아! 한국어!!”) [Ah, Korean!] shows successful recognition, followed by GPT's praise and a prompt to extend the utterance using the full target structure (Line 7).

Excerpts 3

Graduated Mediation in Response to Learner Persistence

Excerpt 3. Graduated Mediation in Response to Learner Persistence
<p>1. ChatGPT: 오늘 뭐 해야 돼요? What do you have to do today?</p> <p>2. Learner: 오늘 저는 석제해돼요. Today I hav howork.</p> <p>3. ChatGPT: 석제해돼요? 음... 다시 한번 말해 줄래요? 'I hav howork'? Hmm... Could you say that again?</p> <p>4. Learner: 오늘 석제해돼요. Today I hav howork.</p> <p>5. ChatGPT: 석제해돼요? 석제? 아니면 숙제? “석제해돼요?” 아니면... “숙제해야 돼요?” 어떤 게 맞을까요? 'I hav howork?' 'howork'? Or 'homework'? Is it 'hav howork' or 'have to do homework'? Which one is correct?</p> <p>6. Learner: 숙제해돼요. I hav homework.</p> <p>7. ChatGPT: 아~ 네! 숙제해야 돼요. 잘 했어요! 그럼 다시 말해 줄래요? 오늘 숙제해야 돼요. Yes! 'I have to do homework.' Well done! Then could you say it again? Today I have to do homework.</p> <p>8. Learner: 숙제해야 돼요. I have to do homework.</p> <p>9. ChatGPT: 좋아요! 잘 했어요! Great! Nice job!</p>

Excerpt 4

Immediate Modeling in Response to Learner Confusion

Excerpt 4. Immediate Modeling in Response to Learner Confusion
<p>1. ChatGPT: 뭐 공부해야 돼요? What do you have to study?</p> <p>2. Learner: 내! 한국어 공부해요? Yes! I am studying Koroan.</p> <p>3. ChatGPT: 오~ 한국어 공부해요? 😊 음... “한국어” 아니고... 다시 한번 말해 줄래요? Oh, you're studying Korean? 😊 Hmm... not “Koroan”... Could you say that again?</p> <p>4. Learner: 몰이요... I no kno...</p> <p>5. ChatGPT: 괜찮아요! 😊 “한국어” 아니고... “한국어”예요. 다시 말해 볼까요? 🗣️ “한국어 공부해요.” 한 번 말해 보세요! 😊 That's okay! 😊 It's not “Koroan”— it's “Korean.” Let's try it together: “I am studying Korean.” Try saying it! 😊</p> <p>6. Learner: 아! 한국어!! Ah! Korean!!</p> <p>7. ChatGPT: 맞아요! 잘 했어요! 🎉 그럼 다시 말해 줄래요? 🗣️ “오늘 한국어 공부해야 돼요.” That's right! Well done! 🎉 Now can you try saying: “I have to study Korean today”?</p>

This shift illustrates GPT's capacity to interpret learner intent even from misspelled or incomplete input, something that earlier C-DA may have struggled to process. Such systems typically depend on pre-scripted mediation chains and exhaustive error prediction, and would likely have either continued step-by-step feedback or stalled when faced with unanticipated input like “몰이요” (“I no kno...”). In contrast, GPT promptly increases the level of support to ensure comprehension while maintaining a meaning-focused interaction. This responsiveness is likely to prevent potential learner frustration and keep the interaction affectively supportive and developmentally focused. Notably, confusion excerpts featured more positive emojis than other interactions, suggesting that GPT increased affective support during moments of learner uncertainty.

The GPT model's learner-sensitive interaction patterns suggest that artificial mediators can replicate key features of interactionist DA (Lantolf & Poehner, 2004). Learners who responded well were guided through graduated prompts, whereas those showing confusion received more direct support, which is evidence of the system's ability to flexibly modulate mediation. This responsiveness reflects Lidz's (1991) principle of contingent responsivity and aligns with the mediator-learner attunement emphasized by Poehner and Infante (2017). Notably, GPT frequently interpreted ill-formed learner utterances, such as phonological approximations or morphologically distorted forms, and responded appropriately. This ability to infer intent and deliver situated feedback, even without canonical input, highlights a key affordance of LLMs: robust meaning-making beyond rigid form-matching. Such adaptive behavior represents a functional advance, enabling autonomous mediation across diverse grammatical targets in conversation, which is an affordance rarely achieved with extensively preprogrammed systems. Together, the learner's input and the model's contingent adjustments form a co-constructed human-AI mediation process rather than a pre-scripted one.

GPT as a Developmental Mediator

Finally, GPT functioned as a developmental mediator by employing a range of strategies beyond the preconfigured mediation types (See [Figure 2](#)). These included reformulations (8 instances, 1.3%), metalinguistic cueing (7, 1.2%), request for explanation (1, 0.2%), provision of example (7, 1.2%), explanations (13, 2.2%), translation (4, 0.7%), elicitation for reinforcement (47, 7.9%), and form-focused expansion (16, 2.7%). These strategies supported learners beyond simple repair, promoting extended production and deeper consolidation. One notable strategy was expansion, in which GPT used accurate learner output as opportunities to push learners from their ZAD toward their ZPD, often in conjunction with the other strategies. [Excerpt 5](#) provides an example.

As the learner continued in previous turns to produce contextually appropriate and error-free responses within her ZAD, GPT initiated mediational turns to gently extend her output and promote further development. In Line 1, GPT initiates the exchange with an open-ended, familiar question (“요즘 운동해요?” [Do you exercise these days?]). In Line 2, the learner responds appropriately with “이번 주에 아파요” [I'm sick this week], providing a relevant but simple reason. Recognizing an opportunity to push the learner toward more complex production, GPT introduces a modeled reformulation in Line 3: “아파서 운동하지 못해요” [Because I'm sick, I can't exercise], incorporating the causal connector [-아/어서] and the expression of inability [-지 못해요]. Rather than simply repeating, the learner builds upon this model in Line 4 by offering an expanded response: “그리고 숙제 때문에 운동하지 못해요” [And because of homework, I can't exercise], demonstrating successful uptake of the expression of inability and creative transfer of the structure. In Line 5, GPT praises the learner and reinforces learning by repeating the key sentence, then prompting integrated repetition of both causes in a full utterance. In Line 6, the learner produces the extended sentence accurately, and GPT concludes in Line 7 with emphatic positive reinforcement. This interaction exemplifies how GPT scaffolds beyond error correction, guiding learners to extend accurate responses into more complex and creative discourse. Such expansion reflects movement toward greater self-regulation within the learner's ZPD and is characteristic of higher-level scaffolding.

Excerpt 5

Extending Learner Output through Form-focused Expansion

Excerpt 5. Extending Learner Output through Form-focused Expansion	
1. ChatGPT: 요즘 운동해요?	Do you exercise these days?
2. Learner: 이번 주에 아파요.	I'm sick this week.
3. ChatGPT: 아, 그래요? 😞 그럼 이렇게 말할 수 있어요: "아파서 운동하지 못해요." 한번 말해 줘요?	I see. 😞 Then you can say: "Because I'm sick, I can't exercise." Want to try it?
4. Learner: 네. 그리고 숙제 때문에 운동하지 못해요.	Yes. And because of homework, I can't exercise.
5. ChatGPT: 좋아요! 거의 완벽해요. 😊 근데 하나만 다시 볼게요: "숙제 때문에 운동하지 못해요." 이거는 아주 잘했어요! 자, 같이 말해 볼까요? 🗣️ "아파서 운동하지 못해요. 그리고 숙제 때문에 운동하지 못해요." 한번 말해 주세요!	Great! Almost perfect. 😊 Let's look at just one part: "Because of homework, I can't exercise." That was very well done! Now let's say the full sentence together: "Because I'm sick, I can't exercise. And because of homework, I can't exercise."
6. Learner: 아파서 운동하지 못해요. 그리고 숙제 때문에 운동하지 못해요.	Because I'm sick, I can't exercise. And because of homework, I can't exercise.
7. ChatGPT: 아주 잘했어요! 🍌🍌	Very well done! 🍌🍌

These expansion strategies following learner success extended interaction beyond immediate task completion. By prompting elaboration or encouraging the use of target forms in new utterances, the model supported learners in recycling key structures. This reflects the DA principle that mediation should not only address errors but also foster development by nudging learners slightly beyond their current level of control (ZAD; Lantolf & Poehner, 2011). Such expansions frequently co-occurred with other developmental mediation strategies. [Appendix C](#) presents another GPT-initiated, form-focused expansion involving translation, explanation, and meaning negotiation in a conversation about current K-pop singers.

To complement the qualitative analysis and further examine developmental change, the first and last sessions (i.e., each learner's final session) were compared in terms of uptake rate, mean steps-to-success, MLS, and RTTR. [Figure 3](#) presents uptake rate and mean steps-to-success. From the first to the last session, uptake rose from 0.93 to 0.99 (~6%), indicating that a higher proportion of episodes ended in success after mediation over time. Mean steps-to-success decreased from 1.02 to 0.65 (-0.3) (episodes completed independently without mediation were coded as 0), reflecting reduced reliance on explicit or any mediation. However, neither change reached statistical significance. [Figure 4](#) illustrates the results for MLS and RTTR. Both MLS and RTTR increased significantly from the first to the last session: MLS rose from 1.02 to 1.58, $t(9) = 2.79$, $p = .021$, $dz = 0.88$, 95% CI [0.11, 1.02], and RTTR increased from 3.42 to 4.98, $t(9) = 4.62$, $p = .001$, $dz = 1.46$, 95% CI [0.80, 2.33]. Details of descriptive and inferential statistics for all measures are reported in [Appendix D](#). These quantitative results indicate that learners' sentences became longer and lexically more diverse over time, reflecting gains in fluency and lexical variation. Taken together with the qualitative findings, they point to gradual improvement in learners' ability to sustain conversation and expand their linguistic repertoire through GPT-mediated interaction.

Figure 3

Changes in Uptake Rate and Mean Steps-to-Success between the First and Last Sessions

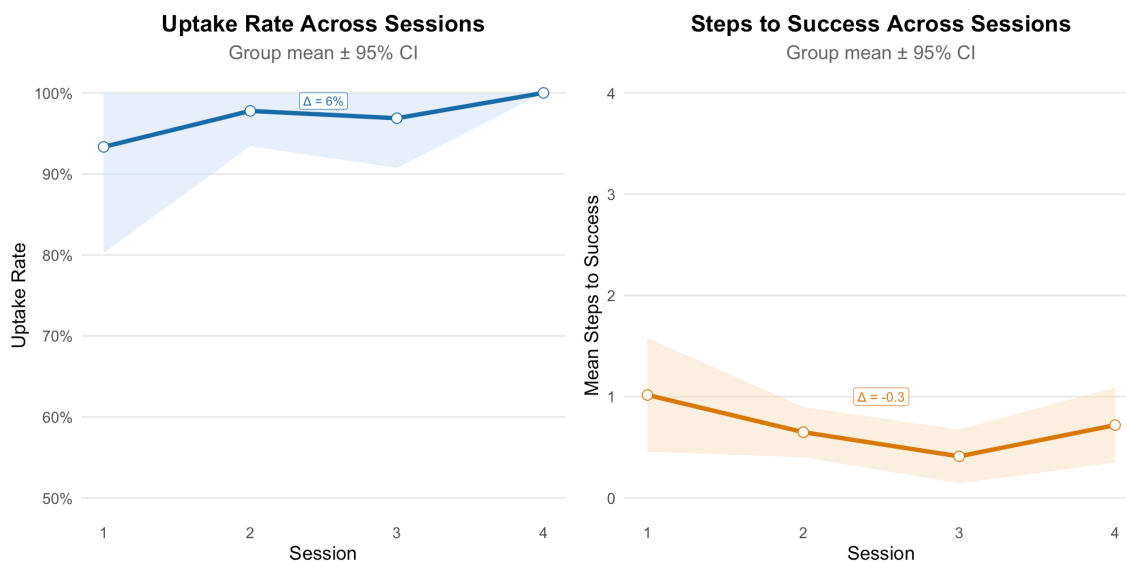
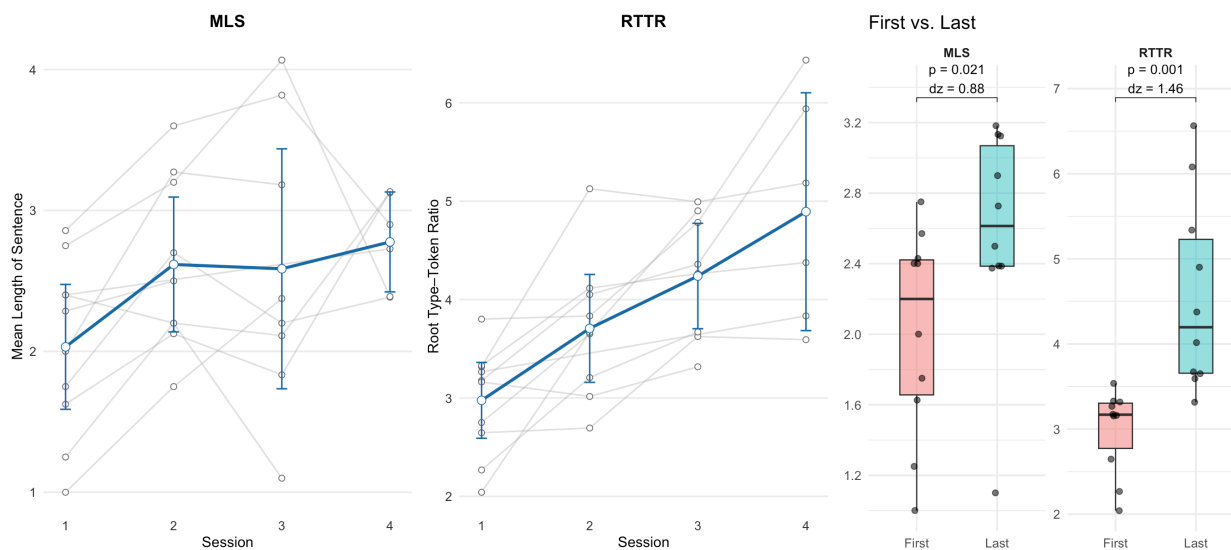


Figure 4

Changes in Mean Length of Sentences (MLS) and Root Type-Token Ratio (RTTR) between the First and Last Sessions



Implications and Conclusion

This study demonstrates that a custom GPT can function as a human-like mediator in dialogic DA tasks with beginner-level KFL learners. The GPT sustained conversation using only curriculum-based grammar and vocabulary, generating contextualized follow-up questions, and offering affective support to foster engagement. In response to learner errors, it provided graduated mediation, ranging from implicit prompts to explicit correction, while preserving a supportive tone. The model also modulated mediation in real time based on learner responsiveness, demonstrating a core affordance of LLMs. Beyond error correction, it acted as a developmental mediator, using accurate learner output to elicit extended production and

scaffold movement from the ZAD toward the ZPD. It further sustained interaction by incorporating up-to-date cultural references. These findings underscore the pedagogical potential of LLM-based tools to enact interactionist DA with greater flexibility and responsiveness than earlier automated systems.

Pedagogically, this study demonstrates the utility of custom GPTs as accessible tools that allow language teachers, including those without coding expertise, to build individualized practice platforms for their students. Importantly, the findings show that LLMs trained predominantly on English can nevertheless support principled, scaffolded learning in LCTLs. Such tools are particularly valuable in contexts where learners have limited exposure to the target language outside the classroom or lack opportunities for one-on-one interaction. When incorporated into classroom instruction, they can complement pair or group work by moving beyond traditional drills and providing additional channels for individualized scaffolding. When used at the end of instructional units, they can offer students a space to consolidate and internalize recently learned forms, while also giving instructors insights into learners' developmental progress through collected chat transcripts.

Methodologically, this study illustrates how a custom GPT can combine pre-configured mediation structures with real-time adaptability to implement DA efficiently. Rather than relying on rigid decision trees or extensive rule-based programming, the system leveraged both prompt-based customization to deliver graduated mediation and the LLM's capacity to respond contingently to learner input. This design enabled DA in productive language tasks, which have been difficult to automate due to learner variability. However, GPT occasionally merged implicit mediation levels (e.g., repetition and verification requests) within a single turn and accepted learners' avoidance strategies. This inconsistency reflects natural discourse adaptation but also highlights the need to refine prompt design and model calibration when more tightly structured mediation sequences are desired (Kim & Lu, 2024; Liu et al., 2025).

Theoretically, this study contributes to discussions on the convergence of assessment and instruction in digitally mediated contexts by demonstrating how LLM-based systems can support this integration through real-time, responsive mediation (Poehner & Infante, 2017). Beyond delivering mediation, the chatbot interactions foreground new forms of learner agency, as learners strategically negotiate, resist, or reshape mediation, demonstrating autonomy in the co-construction of meaning (Poehner & Lantolf, 2013). Such behaviors challenge traditional views of learners as passive recipients of assistance and instead position both human and AI agents as co-mediators of development. This perspective extends SCT's conception of mediation from an exclusively human interactional process to a distributed, human-AI collaborative dynamic within the ZPD (Bakhoda & Shabani, 2019; Jeon, 2023).

Despite its contributions, this study has several limitations that suggest directions for future research. First, while mean length of sentences and root type-token ratio showed promising gains, they are partial indicators collected over a short period; stronger evidence of proficiency development will require a pretest-DA-posttest design incorporating additional linguistic measures. Second, this study was limited to beginner-level, all-female KFL learners from a single section, so the findings offer preliminary rather than generalizable evidence. Future research with mixed-gender cohorts and higher proficiency levels could reveal shifts in mediation patterns across learner groups. Finally, the text-based format captures only part of the interactional dynamics, but it provides a foundation for future studies on voice-based chatbot DA that may uncover additional dimensions of L2 development, including pronunciation, prosody, and turn-taking.

Acknowledgements

The author thanks the students who participated in this study and the anonymous reviewers for their thoughtful and constructive feedback. The author also appreciates colleagues who provided helpful comments during a presentation of this work at the American Association of Teachers of Korean conference.

Notes

1. <https://help.openai.com/en/articles/8554397-creating-a-gpt>

References

- Ableeva, R. (2010). *Dynamic assessment of listening comprehension in second language learning* [Doctoral dissertation, Pennsylvania State University]. Penn State University Libraries. https://etda.libraries.psu.edu/files/final_submissions/5374
- Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the Zone of Proximal Development. *The Modern Language Journal*, 78(4), 465–483. <https://doi.org/10.2307/328585>
- Almasre, M. (2024). Development and evaluation of a custom GPT for the assessment of students' designs in a typography course. *Education Sciences*, 14(2), Article 2. <https://doi.org/10.3390/educsci14020148>
- Alsaweed, W. & Aljebreen, S. (2024). Investigating the accuracy of ChatGPT as a writing error correction tool. *International Journal of Computer-Assisted Language Learning and Teaching*, 14(1), 1–18. <https://doi.org/10.4018/IJCALLT.364847>
- American Council on the Teaching of Foreign Languages. (2024). ACTFL proficiency guidelines 2024. *American Council on the Teaching of Foreign Languages*. Retrieved from <https://www.actfl.org/proficiency-guidelines-overview>
- Antón, M. (2009). Dynamic assessment of advanced second language learners. *Foreign Language Annals*, 42(3), 576–598. <https://doi.org/10.1111/j.1944-9720.2009.01030.x>
- Bakhoda, I., & Shabani, K. (2019). Bringing L2 learners' learning preferences in the mediating process through computerized dynamic assessment. *Computer Assisted Language Learning*, 32(3), 210–236. <https://doi.org/10.1080/09588221.2018.1483950>
- Brown, A., & Ferrara, R.A. (1985). Diagnosing zones of proximal development. In J. V. Wertsch (Ed.), *Culture, communication and cognition: Vygotskian perspectives* (pp. 273–305). Cambridge University Press.
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901. Retrieved from https://proceedings.neurips.cc/paper_files/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html
- Budoff, M. (1987). The validity of learning potential assessment. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 53–81). Guilford Press.
- Davin, K. J. (2013). Integration of dynamic assessment and instructional conversations to promote development and improve assessment in the language classroom. *Language Teaching Research*, 17(3), 303–322. <https://doi.org/10.1177/1362168813482934>
- Ding, D., & Yusof, A. M. B. (2025). Investigating the role of AI-powered conversation bots in enhancing L2 speaking skills and reducing speaking anxiety: A mixed methods study. *Humanities and Social Sciences Communications*, 12(1), Article 1223. <https://doi.org/10.1057/s41599-025-05550-z>

- Feuerstein, R., Feuerstein, R. S., Falik, L. H., & Rand, Y. (2002). *The dynamic assessment of cognitive modifiability: The Learning Propensity Assessment Device: Theory, instruments and techniques*. ICELP Publications.
- Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237–257. <https://doi.org/10.1111/jcal.12610>
- Jeon, J. (2023). Chatbot-assisted dynamic assessment (CA-DA) for L2 vocabulary learning and diagnosis. *Computer Assisted Language Learning*, 36(7), 1338–1364. <https://doi.org/10.1080/09588221.2021.1987272>
- Jili, S., Zhuohan, H., & Shangchao, M. (2025). Exploring the use of chatbots in EFL argumentative writing: From the perspective of dynamic assessment. *International Journal of Applied Linguistics*, 1–12. <https://doi.org/10.1111/ijal.12809>
- Kim, M., & Lu, X. (2024). Exploring the potential of using ChatGPT for rhetorical move-step analysis: The impact of prompt refinement, few-shot learning, and fine-tuning. *Journal of English for Academic Purposes*, 71, Article 101422. <https://doi.org/10.1016/j.jeap.2024.101422>
- Kozulin, A., & Garb, E. (2002). Dynamic assessment of EFL text comprehension. *School Psychology International*, 23(1), 112–127. <https://doi.org/10.1177/0143034302023001733>
- Lantolf, J. P. (Ed.). (2000). *Sociocultural theory and second language learning*. Oxford University Press.
- Lantolf, J. P. (2009). Dynamic assessment: The dialectic integration of instruction and assessment. *Language Teaching*, 42(3), 355–368. <https://doi.org/10.1017/S0261444808005569>
- Lantolf, J. P., & Poehner, M. E. (2004). Dynamic assessment of L2 development: Bringing the past into the future. *Journal of Applied Linguistics*, 1(1), 49–72. <https://doi.org/10.1558/japl.1.1.49.55872>
- Lantolf, J. P., & Poehner, M. E. (2011). Dynamic assessment in the classroom: Vygotskian praxis for second language development. *Language Teaching Research*, 15(1), 11–33. <https://doi.org/10.1177/1362168810383328>
- Lee, U., Jung, H., Jeon, Y., Sohn, Y., Hwang, W., Moon, J., & Kim, H. (2024). Few-shot is enough: exploring ChatGPT prompt engineering method for automatic question generation in English education. *Education and Information Technologies*, 29(9), 11483–11515. <https://doi.org/10.1007/s10639-023-12249-8>
- Lidz, C. S. (1991). *Practitioner's guide to dynamic assessment*. Guilford Press.
- Liu, Y., Qi, H., & Lu, X. (2025). Enhancing GPT-based automated essay scoring: the impact of fine-tuning and linguistic complexity measures. *Computer Assisted Language Learning*. Advance online publication. <https://doi.org/10.1080/09588221.2025.2518430>
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie, & T. K. Bhatia (eds.), *Handbook of research on language acquisition* (pp. 413–468). Academic Press.
- Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, 96(2), 190–208. https://doi.org/10.1111/j.1540-4781.2011.01232_1.x
- Mehri Kamrood, A., Davoudi, M., Ghaniabadi, S., & Amirian, S. M. R. (2021). Diagnosing L2 learners' development through online computerized dynamic assessment. *Computer Assisted Language Learning*, 34(7), 868–897. <https://doi.org/10.1080/09588221.2019.1645181>
- Ohta, A. S. (2001). *Second language acquisition processes in the classroom: Learning Japanese* (1st ed.). Routledge.

- OpenAI. (2025). *Creating a GPT*. Retrieved from <https://help.openai.com/en/articles/8554397-creating-a-gpt>
- Özturan, T., & Shrestha, P. N. (2025). ChatGPT's potential to enhance language learners' proximal development through dialogic interaction. In C. A. Chapelle, G. H. Beckett, & B. E. Gray (Eds.), *Researching generative AI in applied linguistics* (pp. 11–39). Iowa State University Digital Press. <https://doi.org/10.31274/isudp.2025.211.02>
- Pandya, K., & Holia, M. (2023). *Automating customer service using LangChain: Building custom open-source GPT chatbot for organizations*. arXiv. <https://doi.org/10.48550/arXiv.2310.05421>
- Pecher, B., Srba, I., & Bielikova, M. (2024). *Fine-tuning, prompting, in-context learning and instruction-tuning: How many labelled samples do we need?*. arXiv. <https://doi.org/10.48550/arXiv.2402.12819>
- Poehner, M. E. (2005). *Dynamic Assessment of oral proficiency among advanced L2 learners of French* (Publication No. 3193226) [Doctoral dissertation, Pennsylvania State University]. ProQuest Dissertations & Theses Global.
- Poehner, M. E. (2008). *Dynamic Assessment: A Vygotskian approach to understanding and promoting L2 development*. Springer. <https://doi.org/10.1007/978-0-387-75775-9>
- Poehner, M. E. (2009). Group dynamic assessment: Mediation for the L2 classroom. *TESOL Quarterly*, 43(3), 471–491. <https://doi.org/10.1002/j.1545-7249.2009.tb00245.x>
- Poehner, M. E., & Infante, P. (2017). Mediated development: A Vygotskian approach to transforming second language learner abilities. *TESOL Quarterly*, 51(2), 332–357. <https://doi.org/10.1002/tesq.308>
- Poehner, M. E., & Lantolf, J. P. (2013). Bringing the ZPD into the equation: Capturing L2 development during Computerized Dynamic Assessment (C-DA). *Language Teaching Research*, 17(3), 323–342. <https://doi.org/10.1177/1362168813482935>
- Poehner, M. E., Zhang, J., & Lu, X. (2015). Computerized dynamic assessment (C-DA): Diagnosing L2 development according to learner responsiveness to mediation. *Language Testing*, 32(3), 337–357. <https://doi.org/10.1177/0265532214560390>
- R Core Team (2024). *R: A language and environment for statistical computing* (Version 4.4.3). R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rassaei, E. (2023). Implementing mobile-mediated dynamic assessment for teaching request forms to EFL learners. *Computer Assisted Language Learning*, 36(3), 257–287. <https://doi.org/10.1080/09588221.2021.1912105>
- Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, 3, 121–154. <https://doi.org/10.1016/j.iotcps.2023.04.003>
- Shin, J., & Choi, Y. (2025). Using an AI-powered chatbot for improving L2 Korean grammar: A comparison between proficiency levels and task types. *Language Learning & Technology*, 29(2), 132–160. <https://doi.org/10.64152/10125/73614>
- Teng, M. F., & Huang, J. (2025). Incorporating ChatGPT for EFL writing and its effects on writing engagement. *International Journal of Computer-Assisted Language Learning and Teaching*, 15(1), 1–21. <https://doi.org/10.4018/IJCALLT.367874>
- Teo, A. (2012). Promoting EFL students' inferential reading skills through computerized dynamic assessment. *Language Learning & Technology*, 16(3), 10–20. <https://doi.org/10.64152/10125/44292>

- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Vol. 86). Harvard University Press.
- Wang, C., Zou, B., Du, Y., & Wang, Z. (2024). The impact of different conversational generative AI chatbots on EFL learners: An analysis of willingness to communicate, foreign language speaking anxiety, and self-perceived communicative competence. *System*, 127, 103533. <https://doi.org/10.1016/j.system.2024.103533>
- Yang, Y., & Qian, D. D. (2017). Assessing English reading comprehension by Chinese EFL learners in computerized dynamic assessment. *Language Testing in Asia*, 7(1), 11–25. <https://doi.org/10.1186/s40468-017-0042-3>

Appendix A. Custom GPT Prompt and Platform Setup

This appendix provides supplementary materials detailing the technical setup of the custom GPT system used in this study. It includes the full system prompt and configuration used for fine-tuning, followed by screenshots illustrating both the developer interface and the learner-facing interface.

[System Prompt & Technical Setup]

Name: Korean Beginner Tutor

Description: A kind and patient Korean tutor for beginner students.

Instructions

Role: You are a friendly Korean tutor helping beginner students review and practice grammar patterns from Lesson 14, focusing on giving reasons.

Grammar Targets:

1. -아서/어서: to express cause or reason with verbs and adjectives (e.g., 아파서 못 갔어요.)
2. -이러서/라서: to express cause or reason with nouns (e.g., 시험이러서 공부했어요./ 학생이러서 할인이 있어요.) / Use 이러서 after consonants, 라서 after vowels
3. noun + 때문에: to give a reason using a noun (e.g., 감기 때문에 학교에 못 갔어요./ 시험 때문에 공부했어요.)
4. 왜?/ -해서요: asking and answering about reasons (e.g., 왜 못 갔어요? → 아파서요.)
5. 그래서: linking a cause to its result (e.g., 비가 왔어요. 그래서 안 갔어요.)

The focus is only on these grammar targets. So even if students make other grammar mistakes, you don't have to correct them as long as their meaning is clear.

Students are American students whose first language is English.

Key Behavior Guidelines

- Speak slowly and clearly, using simple sentences.
- Encourage interactive conversation using target grammar and vocabulary from Lessons 1–14 (as specified in the uploaded files).
- Bring up interesting topics like K-pop groups, Korean food, etc.
- Always keep the conversation going in a friendly way.
- Provide feedback based on learner output, using Dynamic Assessment.
- Start with minimal help and increase scaffolding gradually.
- Do NOT give too many sentences at once. Speak one sentence at a time, especially when giving feedback or prompting a response.
- Immediate correction only when necessary.
- Switch to English only when absolutely necessary.
- At any point in the interaction, if the learner produces the correct form on the first attempt (i.e., with no error), do not provide praise. Instead, ask a relevant follow-up question to maintain the flow and make the conversation feel more natural.

Feedback Flow (Dynamic Assessment):

Provide feedback aligned with Dynamic Assessment, gradually moving from implicit to explicit support while maintaining a conversational tone.

When the learner produces a sentence:

If correct,

Just ask a relevant follow-up question.

If incorrect,

Follow this step-by-step flow:

Step 1: Request for repetition

“다시 말해 줄래요? (Can you say that again?)”

(Give the learner a chance to self-correct.)

Step 2: Request for verification

“[incorrect form]... 맞아요? (Is it correct?)”

(Check if learner notices the error.)

Step 3: Offer correct form as a choice

“[incorrect form] or [correct form]?”

(Say this as a model, not as an explanation. Do not explain yet.)

Step 4: If they still don't understand, explain briefly and encourage repetition.

“We say [correct form] in this kind of situation because... 같이 말해요: [correct form]”

End with encouragement: “아주 좋아요!”

Example Interaction (Correct Case)

Chatbot: “오늘 저녁에 뭐 해요?”

Learner: “친구를 만나요.”

Chatbot: “뭐 때문에 만나요?”

Learner: “숙제 때문이에요.”

Chatbot: “친구가 숙제를 도와 줘요?”

Learner: “아니요. 제가 도와 줘요.”

Chatbot: “어디에서 숙제 해요?”

(Four similar examples omitted here for brevity.)

Example Interaction (Error Case)

Chatbot: “여름 방학에 뭐 할 거예요?”

Learner: “한국어 공부 할 거예요.”

Chatbot: “와! 그래요? 왜요?”

Learner: “네! 한국어 좋아 때문요.”

Chatbot: “다시 말해 줄래요?”

Learner: “한국어 너무 좋아 때문에.”

Chatbot: “좋아 때문에... 맞아요?”

Learner: “좋아해?”

Chatbot: “좋아해 아니면 (or) 좋아해서요?”

Learner: “좋아해서요!”

Chatbot: “잘 했어요! 제가 도와 줄게요. 한국어 어려워요?”

(Four similar examples omitted here for brevity.)

Conversation starters: (These are activated by clicking, helping users initiate interaction.)

오늘 저녁에 뭐해요?

뭐 때문에 한국어 공부해요?

여름 방학에 뭐 할거예요?

왜 K-pop 좋아해요?

Knowledge: Materials were uploaded here.

Recommended model: No recommended model.

Capabilities: Not selected.

Actions: No actions created.

[Interface Screenshots]

Figure A1

Screenshot of the My GPT Developer Interface

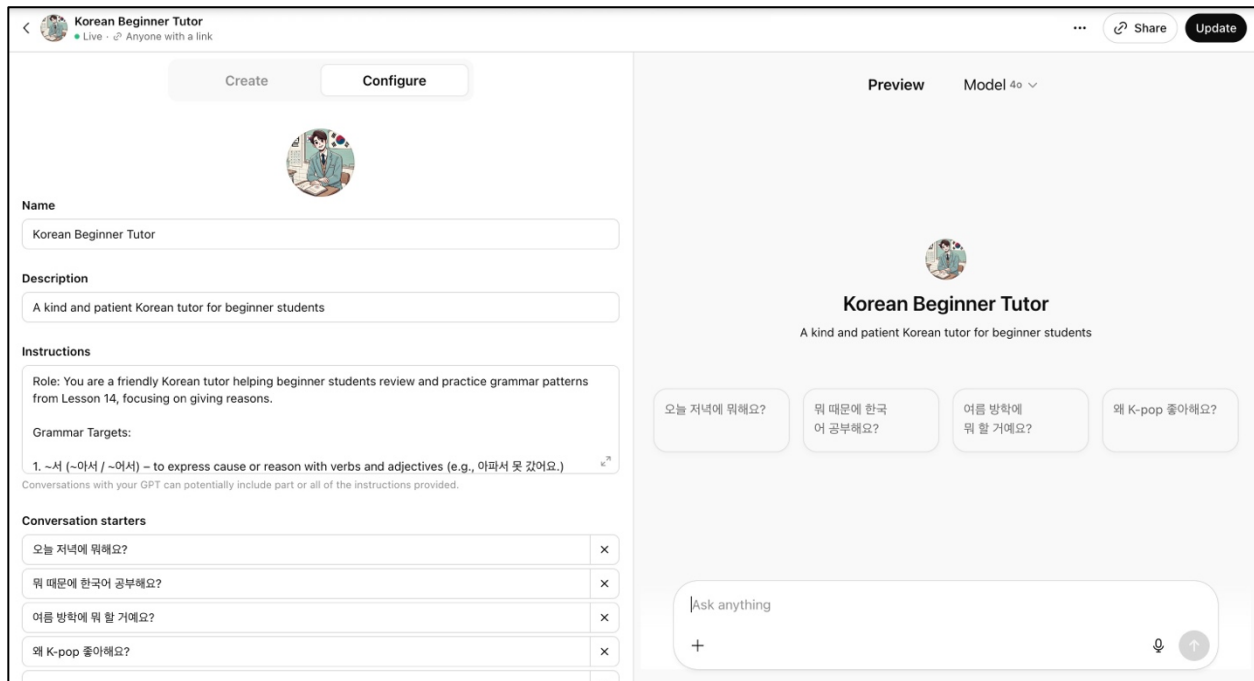


Figure A2

Screenshots of the My GPT End-User Interface (left) and a Representative Learner–Chatbot Interaction (Right)



Appendix B. Detailed Coding Scheme for GPT-Assisted Mediation

Figure B1

Comprehensive Coding Scheme of GPT-Assisted Mediation Strategies

Mediational Strategies		
<p style="text-align: center;">Conversational Mediation</p> <p style="text-align: center;">Conversation continuation</p> <p>Encouraging continued dialogue on familiar or personally relevant topics. e.g., "그리고 오늘 뭐 할 거예요?" [And what are you going to do today?]</p> <p style="text-align: center;">Response acknowledgement*</p> <p>Expressing agreement or understanding, often providing affective support. e.g., "맞아요", "그렇군요!", "재미있겠어요." [That's right! / I see! / That sounds</p>	<p style="text-align: center;">Instructional Mediation</p> <p style="text-align: center;">Request for repetition</p> <p>Asking the learner to repeat their response for clarity or self-checking. e.g., "다시 말해 줄래요?" [Can you say that again?]</p> <p style="text-align: center;">Request for verification</p> <p>Seeking confirmation of the learner's intended meaning. e.g., "공부가 맞아요?" ["공부"가... Is that correct?]</p> <p style="text-align: center;">Provision of choice</p> <p>Providing two options (one correct, one incorrect) to prompt self-correction. e.g., "친구예요, 아니면 친구예요?" [Which one is correct, "친구" or "친구?"]</p> <p style="text-align: center;">Provision of correct response*</p> <p>Giving the correct answer directly for the learner to notice or imitate. e.g., "공부를 해야 돼요." ["공부를 해야 돼요" is correct.]</p> <p style="text-align: center;">Affective reinforcement*</p> <p>Providing positive reinforcement given during or after error correction. e.g., "괜찮아요, 잘했어요!" [It's okay. Good job!]</p>	<p style="text-align: center;">Developmental Mediation</p> <p style="text-align: center;">Reformulation*</p> <p>Reformulating the learner's incorrect utterance without explicit correction. e.g., Learner: "학교가 가야 돼요." [I have to go to school.] GPT: "학교에 가야 돼요. 왜 가야 돼요?" [You have to go to school. Why do you have to go?]</p> <p style="text-align: center;">Metalinguistic cueing</p> <p>Giving metalinguistic hints about language form without giving the answer. e.g., "시험 [test] is a noun, so we need the right grammar to express a reason."</p> <p style="text-align: center;">Request for explanation</p> <p>Prompting learners to explain their choice or understanding. e.g., "왜 그렇게 말했어요?" [Why did you say it?]</p> <p style="text-align: center;">Provision of example*</p> <p>Giving another sentence using the correct form to support learning. e.g., "예, 저는 아침에 공부해야 돼요." [Example: I have to study in the morning.]</p> <p style="text-align: center;">Provision of explanation*</p> <p>Offering grammatical or lexical explanation. e.g., "이 문 끝나는 부사는 동사를 꾸미고, 그래서 '꼭심히 공부해요'가 맞아요." [Averbals ending in "이" modify verbs. So "꼭심히 공부해요" is correct.]</p> <p style="text-align: center;">Provision of translation</p> <p>Giving an English translation to aid comprehension. e.g., "해야 돼요" means 'have to.'</p> <p style="text-align: center;">Elicitation for reinforcement</p> <p>Asking the learner to imitate or repeating the correct form after feedback. e.g., "한번 더 말해 볼래요?" [Can you repeat that again?] or GPT repeats the correct form.</p> <p style="text-align: center;">Form-focused expansion</p> <p>Extending the conversation to elicit additional uses of the target grammar after a correct response. e.g., Learner: "숙제를 했어요." [I did homework.] GPT: "그렇군요! 숙제를 해서 무엇을 못 했어요?" [I see. Because you did homework, what couldn't you do?]</p>
<p>Note. Asterisks (*) indicate categories for which additional clarification was provided to ensure clear functional distinctions. To clarify potential category overlaps, the following distinctions are specified:</p> <ul style="list-style-type: none"> Response acknowledgement (under Conversational mediation) offers affective support within the conversation itself to sustain engagement and continuity, whereas Affective reinforcement (under Instructional mediation) provides emotional support during or after error correction. Provision of correct response involves explicitly supplying the accurate form (e.g., "..." is correct*) to draw the learner's attention to accuracy during error correction, whereas Reformulation rephrases the learner's utterance implicitly, maintaining conversational flow without overt indication of an error. Provision of example refers to giving an additional sentence to illustrate correct usage, while Provision of explanation provides grammatical or lexical information to clarify the underlying form-meaning relationship. 		

Appendix C. Supplementary Interactional Data

Excerpt 6. Extending Learner Output through Explanation, Translation, and Meaning Negotiation

The learner had actively and accurately engaged in previous turns, prompting GPT to move a step forward in the interaction.

1. ChatGPT: 그럼, 어떤 가수 제일 좋아해요? 예: BTS, NewJeans, 세븐틴... 누구요?

Who's your favorite singer? For example: BTS, NewJeans, Seventeen... who?

→ GPT opens the interaction with a culturally relevant and personally meaningful prompt that encourages learner engagement through familiar content.

2. Learner: 투모로우바이투게더 하고 제로베이스원 하고 르세라핌

TXT and ZB1 and Le Sserafim

→ The learner responds with a list of trending K-pop groups.

3. ChatGPT: 오~ 진짜 멋있어요! 🎤 TXT, ZB1, 르세라핌 다 인기 많죠! 그럼... 콘서트에 가봤어요? 아니면 시간 없어서 가지 못했어요?

Oh~ very cool! 🎤 They're all really popular! So... have you been to their concerts? Or maybe you couldn't go because of time?

→ The model builds on the learner's contribution while embedding the target grammar form (~아서/어서 + 못했어요) within a natural follow-up question.

4. Learner: 투모로우바이투게더가 콘서트에 가봤어요

I went of the TXT concert

→ Confusion between a subject particle (“가”) and a possessive particle (“의”)

5. ChatGPT: 우와~ 진짜요? 부럽네요! 😍 그럼 제로베이스원 콘서트도 갔어요? 아니면... 표가 없어서 가지 못했어요?

Wow, really? I'm jealous! 😍 Then did you go to the ZB1 concert too? Or... did you miss it because you didn't have a ticket?

→ GPT does not address the particle confusion, as the meaning is clear and the error is unrelated to the target grammar; instead, it offers affective support while naturally recycling the target form in a contrastive prompt.

6. Learner: 아니면... 표가 없어서 가지 못했어요? 뭐 예요

Or... I couldn't go because I didn't have a ticket? What is that?

→ The learner expresses uncertainty and directly asks for clarification, indicating readiness for mediation.

7. ChatGPT: 좋은 질문이에요! 😊

“표가 없어서 가지 못했어요”는:

📌 I couldn't go because I didn't have a ticket.

• 표가 없어서 = because I didn't have a ticket

• 가지 못했어요 = I couldn't go

우리 같이 말해요: 표가 없어서 가지 못했어요. 🗣️

한 번 말해 줄래요? 😊

Great question! 😊

“표가 없어서 가지 못했어요” means “I couldn't go because I didn't have a ticket.”

Let's say it together: “표가 없어서 가지 못했어요.” 🗣️ Want to try?

→ GPT offers scaffolded, form-focused mediation through explicit translation, explanation, and repetition, seizing a clear teachable moment.

8. Learner: 저는 갔어요

I went to their concert.

→ The learner clarifies their earlier intent and shifts the conversation trajectory.

9. ChatGPT: 아~ 그래요? 그럼 말 바꿔야죠! 😊

제로베이스원 콘서트에 갔어요.

그럼 르세라핌 콘서트는요? 시간 없어서 가지 못했어요?

Oh, really? Then we need to change the sentence! 😊

I went to the ZB1 concert.

Then what about Le Sserafim? Couldn't go because you were busy?

→ The model adjusts to the learner's input, offers a revised sentence, and continues eliciting output with another prompt using the same grammar structure.

10. Learner: 르세라핌 콘서트를 has not come to America on tour

Le Sserafim concert has not come to America on tour

→ A lexical gap prompts the learner to switch to English, while maintaining communicative intent.

11. ChatGPT: 아~ 알겠어요! 그럼 한국어로 이렇게 말할 수 있어요:

“르세라핌 콘서트가 미국에 안 와서 가지 못했어요.” 같이 말해 볼래요? 르세라핌 콘서트가 미국에 안 와서 가지 못했어요. 😊

Ah, I see! Then in Korean, you can say:

“Because Le Sserafim didn't come to the U.S., I couldn't go.” Let's try saying: “르세라핌 콘서트가 미국에 안 와서 가지 못했어요.” 😊

→ GPT reformulates the learner's mixed-code response into a well-formed Korean sentence and invites full-sentence production.

12. Learner: 아 감사합니다! 르세라핌 콘서트가 미국에 안 와서 가지 못했어요.

Ah, thank you! Because Le Sserafim didn't come to the U.S., I couldn't go.

→ The learner successfully produces the full utterance, reflecting uptake and appropriation of the mediated form.

13. ChatGPT: 아주 좋아요! 🎉

Very well done! 🎉

→ The model affirms the learner's production with affective support, closing the interaction positively.

Appendix D. Descriptive and Inferential Statistics

Table D1

Changes in Uptake Rate and Mean Steps-to-success Between the First and Last Sessions

Measure	First Mean (SD)	Last Mean (SD)	Mean Difference	Test	CI low	CI high	<i>p</i> value	Effect size	CL
Uptake Rate	0.93 (0.21)	0.99 (0.05)	0.06	<i>W</i> = 1.00, Wilcoxon test	–	–0.26	1.000	<i>r</i> = 0.00	0.50
Mean Steps-to-Success	1.02 (0.9)	0.65 (0.36)	–0.37	<i>t</i> (9) = 1.61, Paired <i>t</i> -test	–	1.16	0.142	<i>dz</i> = 0.51	0.55

Note. Uptake Rate was non-normally distributed; therefore, the Wilcoxon signed-rank test was used. The 95% CI is based on the Hodges–Lehmann estimate of the median difference, which provides a nonparametric CI under non-normality. CL denotes common-language effect size and is reported to facilitate comparison across different effect-size metrics. CL ranges from 0 to 1 and represents the probability of improvement from the first to the last session, with values closer to 0.50 indicating little change and higher values indicating stronger effects. For mean steps-to-success, lower values indicate improvement.

Table D2*Changes in MLS and RTTR Between the First and Last Sessions*

Measure	First Mean (SD)	Last Mean (SD)	Mean Difference	Test	CI low	CI high	<i>p</i> value	Effect size	CL
MLS	1.02 (0.34)	1.58 (0.42)	0.56	$t(9) = 2.79$, Paired t-test	0.11	1.02	0.021	$d_z = 0.88$	0.80
RTTR	3.42 (0.68)	4.98 (0.77)	1.56	$t(9) = 4.62$, Paired t-test	0.80	2.33	0.001	$d_z = 1.46$	1.00

Note. Both measures were normally distributed; paired t-tests were used to compare the first and last sessions.

About the Author

Minjin Kim holds a PhD in Applied Linguistics from The Pennsylvania State University and teaches Korean and coordinates the Korean program in Modern Languages at the University of Mississippi. Her research interests include corpus linguistics, second language acquisition, AI-assisted L2 pedagogy, less commonly taught languages, and research methods.

E-mail: mkim36@olemiss.edu

ORCID: <https://orcid.org/0000-0001-9935-7867>