

Learner Mediation of AI-Generated Language: A Case Study of Beginner CFL Learners Using ChatGPT for Role-Play Script Construction

Lew Ya Ling¹, Chee Sie Ching², Intan Safinas Mohd Ariff Albakri^{1*}

¹Faculty of Languages and Communication, Sultan Idris Education University (UPSI), Malaysia

²Academy of Language Studies, Universiti Teknologi MARA Perlis Branch, Malaysia

*Corresponding Author

DOI: <https://doi.org/10.47772/IJRISS.2026.100800164>

Received: 12 August 2026; Accepted: 19 August 2026; Published: 29 August 2026

ABSTRACT

The increasing use of generative artificial intelligence (GenAI) tools such as ChatGPT has created new opportunities for language learning. However, existing research has largely focused on learning outcomes and learner perceptions, with less attention given to how learners interact with and mediate AI-generated language during task completion. This study investigates learner–AI interaction among beginner Chinese as a Foreign Language (CFL) learners during role-play script construction. Adopting a qualitative case study design, the study involved 28 third-level CFL learners at Universiti Teknologi MARA (UiTM), Malaysia. Data were collected from ChatGPT prompt logs, AI-generated role-play scripts, learner revision tables, and reflective essays. The analysis focused on three interrelated aspects of learner–AI interaction: prompting practices, learner mediation of AI-generated language, and learners’ perceptions of AI-assisted learning. The findings indicate that learners actively engaged with ChatGPT rather than passively accepting AI-generated outputs. Learners refined prompts, verified AI-generated language using multiple sources, and frequently modified, simplified, localized, and adapted AI-generated expressions according to their linguistic knowledge, classroom learning experiences, and communicative needs. The findings further reveal the continuing influence of textbook authority and the importance of performability in learners’ evaluative decisions. In addition, learners frequently localized AI-generated language to better reflect Malaysian sociocultural contexts. Although participants generally perceived ChatGPT as a useful learning resource, they did not regard it as a replacement for teachers or textbooks. Drawing on constructivist perspectives and feedback literacy, the study argues that learner–AI interaction is best understood as a process of mediation in which learners actively evaluate, negotiate, and adapt AI-generated language. The study contributes to the growing literature on AI-assisted language learning by highlighting the role of learner agency and evaluative judgement in beginner CFL contexts.

Keywords: ChatGPT, Chinese as a Foreign Language (CFL), Feedback Literacy, Learner Mediation, Prompting

INTRODUCTION

The rapid development of generative artificial intelligence (GenAI), particularly large language models (LLMs) such as ChatGPT, has attracted increasing attention in language education. Recent research suggests that AI-powered tools can support language learning by providing immediate feedback, generating learning materials, facilitating writing processes, and encouraging learner engagement across a range of language skills (Kasneci et al., 2023; Wang et al., 2024). Within language learning contexts, ChatGPT has been used as a resource for idea generation, language practice, and individualized support (Barrot, 2023; Xiao & Zhi, 2023).

In Chinese as a Foreign Language (CFL) education, ChatGPT presents particular opportunities because Chinese poses distinctive challenges for non-native learners, including character learning, vocabulary development, and the use of near-synonymous expressions (Li et al., 2023; Zhao et al., 2024). Existing studies have reported positive outcomes in areas such as writing support, reading comprehension, vocabulary learning, and learner

motivation. However, most of this research has concentrated on learning outcomes or learners' general perceptions of AI use, with comparatively less attention given to the processes through which learners interact with AI-generated language during task completion.

This gap is particularly relevant for beginner CFL learners, whose limited linguistic resources often require them to evaluate, verify, simplify, and adapt AI-generated language before it can be used effectively for communication. Although ChatGPT can generate fluent and contextually plausible responses, the educational value of these responses depends largely on learners' ability to formulate appropriate prompts, evaluate the suitability of generated language, and make informed revisions when necessary (Fan et al., 2025; Zhan & Yan, 2025). Understanding these interactional processes is therefore important for developing pedagogically meaningful approaches to AI integration in CFL classrooms.

Role-play script construction provides a useful context for examining learner–AI interaction. Unlike discrete language exercises that focus on isolated linguistic forms, role-play tasks require learners to consider language accuracy, communicative appropriateness, contextual meaning, and oral performance simultaneously. When using ChatGPT to generate role-play scripts, learners are not merely recipients of AI-generated content. Rather, they must negotiate language choices, refine prompts, evaluate generated responses, and adapt the language to their own communicative needs. Such processes provide opportunities to examine how learners engage with AI-generated language and how they exercise evaluative judgement during task completion.

Drawing on constructivist perspectives of learning and the concept of feedback literacy (Carless & Boud, 2018), this study investigates learner–AI interaction in a CFL role-play script task. Specifically, it focuses on three interrelated dimensions: (1) learners' prompting practices, (2) learners' mediation of AI-generated language through evaluation, verification, and revision, and (3) learners' perceptions of AI-assisted language learning. By examining these processes, the study contributes to a more process-oriented understanding of AI integration in CFL education and extends current discussions beyond learning outcomes to the ways learners actively work with AI-generated language during task completion.

This study contributes to the growing body of research on AI-assisted language learning in three ways. First, it focuses on CFL learners, a context that remains relatively underrepresented in current ChatGPT research. Second, it examines role-play script construction as a communicative task that requires learners to negotiate language form, contextual appropriateness, and oral performance demands. Third, rather than focusing solely on learning outcomes, the study investigates learner–AI interaction processes, particularly how learners construct prompts, evaluate AI-generated language, and mediate AI-generated outputs during task completion.

In the local CFL context, role-play assignments have traditionally relied on readily available online scripts or previously existing materials, which may limit learners' opportunities to engage actively in language production and decision-making. The integration of ChatGPT into role-play script construction was therefore intended not only to support dialogue generation but also to encourage learners to participate more actively in evaluating, revising, and adapting language. By requiring learners to document their prompts, modifications, and reflections, the task created a valuable context for examining learner–AI interaction and the processes through which learners mediate AI-generated language.

Research Objectives

This study aims to examine learner–AI interaction in a Chinese as a Foreign Language (CFL) role-play script task by focusing on three interrelated aspects. Based on the research gaps identified in the literature, the study pursues the following objectives:

First, it investigates how learners construct and refine prompts when using ChatGPT to generate role-play scripts.

Second, it examines how learners evaluate, verify, and mediate AI-generated language during role-play script construction.

Third, it explores learners' perceptions and experiences of using ChatGPT in AI-assisted role-play learning tasks.

Research Questions

To address these objectives, the study is guided by the following research questions:

1. How do learners construct prompts when using ChatGPT to generate role-play scripts in a Chinese language learning context?
2. How do learners evaluate and mediate AI-generated language during role-play script construction?
3. How do learners perceive and experience the use of ChatGPT in role-play script learning tasks?

LITERATURE REVIEW

The integration of generative artificial intelligence (GenAI) tools, such as ChatGPT, has initiated a significant paradigm shift in foreign language education (Adiguzel et al., 2023). In the context of learning Chinese as a Foreign Language (CFL), ChatGPT offers novel opportunities for learners to engage in customized dialogue generation, receive immediate feedback, and design creative materials. This review examines existing research on the use of ChatGPT in CFL learning, with particular attention to learner–AI interaction during role-play script construction. This review focuses on four interrelated areas relevant to the present study: ChatGPT in CFL learning, prompting and learner–AI interaction, feedback literacy and learner mediation, and constructivist perspectives on learner–AI interaction.

ChatGPT in Chinese as a Foreign Language (CFL) Learning

The emergence of ChatGPT has generated growing interest in foreign language education due to its ability to provide immediate feedback, generate learning materials, and support language practice (Adiguzel et al., 2023). In Chinese as a Foreign Language (CFL) contexts, ChatGPT offers learners opportunities to explore vocabulary usage, generate dialogues, and receive language support beyond the classroom.

Its application is particularly relevant to CFL learning because Chinese presents distinctive challenges for non-native learners, including its logographic writing system, tonal features, and near-synonymous expressions (Zhao et al, 2024). Previous studies have shown that ChatGPT can assist learners in generating dialogues, clarifying vocabulary, and improving communicative performance (Li, Ren, Jiang, & Chen, 2023). It may also function as an accessible supplementary learning resource, particularly for learners with limited opportunities for additional language support (Li, Li, & Cho, 2023). Zhang (2025) further suggests that ChatGPT serves effectively as an interactive language reference tool that enables learners to obtain explanations and examples in real time.

Despite these benefits, AI-generated language is not always linguistically or contextually appropriate. Learners must often evaluate, verify, and adapt ChatGPT-generated outputs before using them in actual communication. While existing studies have primarily focused on learning outcomes and learner perceptions, less attention has been paid to how CFL learners interact with, evaluate, and modify AI-generated language during task completion. This gap is particularly evident in role-play script construction, where learners must negotiate language choices, communicative appropriateness, and contextual meaning simultaneously.

Prompting and Learner–AI Interaction

As the quality of ChatGPT-generated responses is closely related to the instructions provided by users, prompting has become an important aspect of learner–AI interaction. Prompt engineering refers to the practice of formulating clear and contextually appropriate instructions to guide AI-generated output (Ekin, 2026; Wang et al., 2023). Previous studies have shown that detailed prompts generally produce more relevant and useful responses, whereas vague instructions often lead to generic or less appropriate outputs (Lo, 2023).

In language learning contexts, prompting involves more than a technical skill. When interacting with ChatGPT, learners must communicate their learning goals, specify communicative situations, and refine their requests according to the responses generated by the system. Effective prompts may include information about speakers' roles, conversational contexts, language level, and expected output format. In CFL learning, Zhao, Hsu, and

Huang (2024) propose the use of Zone of Proximal Development Prompts (ZPDP), which encourage learners to guide AI-generated language toward their own proficiency levels.

Prompting is therefore not a one-time act of giving instructions but an iterative process of negotiation between learner intentions and AI-generated responses. Through revising prompts, requesting clarification, and adjusting output requirements, learners actively shape the language they receive from the AI. Prompting can therefore be viewed as a process of interaction and negotiation rather than a simple act of instruction.

Feedback Literacy and Learner Mediation

The increasing use of ChatGPT in language learning has drawn attention to the importance of learners' feedback literacy and their ability to evaluate AI-generated language. Feedback literacy refers to the understandings, capacities, and dispositions that enable learners to interpret information, make judgments about its quality, and use it to improve their learning (Carless & Boud, 2018; Molloy, Boud, & Henderson, 2020). In AI-assisted learning environments, these abilities become particularly important because learners are no longer receiving feedback solely from teachers but also from generative AI systems.

ChatGPT has expanded the range of feedback sources available to learners by providing immediate and individualized responses during learning activities (Rad, Alipour, & Jafarpour, 2024). However, AI-generated language is not always accurate or contextually appropriate. Because ChatGPT generates responses based on statistical patterns rather than genuine language understanding, it may produce inaccurate information, unsuitable expressions, or contextually problematic outputs (Kasneci et al., 2023). Consequently, learners need to evaluate AI-generated content rather than accept it uncritically.

Recent studies have highlighted the risks of excessive dependence on AI-generated feedback. When learners rely on AI without critically examining the responses, they may become less engaged in evaluative thinking and develop what Fan et al. (2025) term "metacognitive laziness." To use AI effectively, learners must compare information across sources, identify potential inaccuracies, and determine whether AI-generated language is appropriate for their communicative purposes.

Evaluative judgement may also extend beyond grammatical accuracy. Learners often need to consider contextual appropriateness and local language practices when evaluating AI-generated outputs. Since ChatGPT is trained primarily on large-scale global datasets, some generated expressions may not fully reflect local linguistic and cultural contexts (Godwin-Jones, 2023). As a result, learners may need to adapt or localize AI-generated language to suit their communicative environments.

From this perspective, interacting with ChatGPT involves more than receiving feedback. Learners are required to interpret, verify, modify, and sometimes reject AI-generated suggestions. Such mediation processes require evaluative judgement and active learner engagement (Bearman & Ajjawi, 2023). Examining how learners evaluate, negotiate, and revise AI-generated language can therefore provide valuable insights into the development of feedback literacy and learner agency in AI-assisted language learning environments.

Constructivist Perspectives on Learner–AI Interaction

The integration of ChatGPT into language learning can be understood through a constructivist perspective, which views learning as an active process in which learners construct knowledge through experience, problem-solving, and interaction with others and their environment (Zajda, 2021). Rather than passively receiving information, learners actively interpret, evaluate, and reorganize knowledge as they engage with learning tasks.

This perspective is closely aligned with Vygotsky's sociocultural theory, which emphasizes the mediating role of tools and social interaction in cognitive development (Vygotsky, 1978). From this viewpoint, digital technologies can function as mediational tools that support learners in accomplishing tasks that may otherwise be beyond their current level of competence. Recent studies have suggested that conversational AI tools such as ChatGPT can provide scaffolding by generating explanations, examples, and alternative language forms that support learner engagement and participation (Belda-Medina & Calvo-Ferrer, 2022).

Among the various language-learning tasks supported by AI, role-play script construction is particularly compatible with constructivist learning principles. To complete the task, learners must not only generate language through ChatGPT but also evaluate, revise, and adapt AI-generated outputs according to their communicative goals and language proficiency. Learning therefore occurs through an ongoing process of interaction and negotiation rather than through the simple acceptance of AI-generated responses.

Viewed from this perspective, ChatGPT is not a replacement for teachers or an autonomous source of knowledge. Rather, it serves as a mediational resource that learners can draw upon while constructing meaning and completing language-learning tasks.

However, despite the increasing application of constructivist and sociocultural perspectives in AI-assisted language learning research, existing studies have only partially explored how learners actively engage with AI-generated language during task completion. Much of the current literature continues to emphasise learning outcomes, task performance, or learners' general perceptions of ChatGPT use, while paying comparatively less attention to the interactional processes through which learners evaluate, negotiate, and adapt AI-generated language.

Although previous studies have documented the benefits of ChatGPT for language learning, three gaps remain. First, existing research has predominantly focused on learning outcomes and perceptions rather than interactional processes. Second, beginner CFL learners remain underrepresented in current studies. Third, limited attention has been paid to how learners in multilingual contexts mediate and localise AI-generated language during communicative tasks such as role-play script construction.

Addressing these gaps is particularly important in multilingual settings such as Malaysia, where learners frequently draw upon multiple linguistic resources and must navigate potential discrepancies between globally generated Chinese language outputs and local communicative practices. This study therefore adopts a constructivist lens to examine how beginner CFL learners interact with, evaluate, and mediate AI-generated language during role-play script construction.

METHODOLOGY

Research Design

This study adopted a qualitative case study design to investigate how learners interacted with ChatGPT during role-play script construction in a Chinese as a Foreign Language (CFL) learning context. The study focused on learners' prompting practices, learner mediation of AI-generated language, and perceptions of AI-assisted language learning.

The case examined in this study was bounded within a third-level CFL course conducted at Universiti Teknologi MARA (UiTM) Perlis during the September 2024 semester. The study focused on a role-play script assignment in which learners used ChatGPT to generate, evaluate, and revise Chinese-language dialogues. As generative AI tools became increasingly common among students, the study sought to examine how learners interacted with AI-generated language during task completion.

A qualitative case study approach was adopted because the research focused on learner–AI interaction processes rather than learning outcomes. Particular attention was given to how learners evaluated, negotiated, and adapted AI-generated language within an authentic classroom context.

Research Setting and Participants

The study was conducted at Universiti Teknologi MARA (UiTM) Perlis and involved 28 third-level Chinese as a Foreign Language (CFL) learners from the Faculty of Applied Science. The participants, who were approximately 21 years old, were predominantly Malay-speaking non-Chinese learners with no formal background in Chinese prior to university enrolment. At UiTM Perlis, students enrolled in the Introductory

Chinese programme complete three sequential Mandarin courses over three semesters. By the time of the study, the participants had completed two previous levels and had acquired approximately 300 Chinese vocabulary items, corresponding roughly to beginner–intermediate proficiency (approximately HSK Levels 1–2). Most learners were able to participate in basic communicative interactions in Chinese but continued to rely heavily on textbook vocabulary and structured sentence patterns. Third-level learners were selected because they had developed sufficient linguistic resources to engage in extended communicative tasks while still encountering challenges in independently evaluating, adapting, and contextualising AI-generated language.

The participants were divided into five groups to complete a role-play script assignment based on everyday communicative scenarios such as travel, dining, workplace interaction, illness, and social communication.

Most participants had prior experience using ChatGPT informally, although not all had previously used it for completing academic assignments. During the assignment, students primarily used the free tier of ChatGPT, which defaulted to GPT-4o mini at the time of the study. As the free-tier model available at the time of the study may differ from subsequent versions, the findings should be interpreted within the technological context in which the data were collected.

Task Design and Prompting Protocol

The core assignment required students to use ChatGPT to generate role-play scripts based on assigned communicative themes. The role-play scenarios covered six everyday communicative themes: buying groceries and cooking, getting sick, vacationing, asking for directions, asking for help, and interviewing or casual conversation. Students were instructed to generate dialogues, evaluate the AI-generated outputs, revise problematic or unsuitable expressions, and explain the reasons behind their modifications.

The role-play task was designed as a group-based communicative activity. Each group was required to produce a role-play script consisting of approximately 60 conversational turns and subsequently perform the completed script in class. The oral performance lasted approximately 10 minutes and formed part of the course assessment. Both the written role-play script and the oral performance were evaluated. As a result, learners were expected not only to generate linguistically appropriate dialogues but also to ensure that the scripts were manageable for memorization and oral presentation.

Although no fixed prompt template was imposed, students were provided with general prompting guidance. They were encouraged to:

1. Specify the communicative context and scenario;
2. Identify the number of speakers and intended conversational roles;
3. Request preferred language formats such as Chinese characters, Pinyin, or translations; and
4. Revise prompts iteratively when the generated responses did not match their communicative intentions.

This semi-structured prompting approach allowed learners flexibility in interacting with ChatGPT while enabling the study to observe variation in prompting strategies and learner–AI interaction patterns. Because the generated scripts were subsequently used for classroom performance, learners were required to make decisions regarding language complexity, contextual appropriateness, memorization, and oral delivery. These conditions provided a useful context for examining how learners evaluated and mediated AI-generated language during task completion.

Data Collection

Data were collected from multiple sources to capture learners' interaction processes comprehensively. These included:

1. Original ChatGPT-generated role-play scripts;

2. Students' revised versions of the scripts;
3. Prompt logs documenting learner–AI interaction;
4. Revision tables explaining learners' modifications;
5. Reflective essays describing learners' experiences using ChatGPT; and
6. Teacher observations and comments recorded during script revision and classroom performance.

Teacher observations were used as supplementary data to triangulate learners' revision practices, reflective accounts, and interaction patterns.

Analytical Categories of Learner Mediation

The study did not assume that all AI-generated outputs were inherently correct or incorrect. Instead, the analysis focused on how learners evaluated, negotiated, and modified ChatGPT-generated language according to their own linguistic competence, textbook knowledge, communicative goals, and local Malaysian Chinese language context.

The analytical categories were informed by prior research on AI-assisted language learning and CFL error analysis while remaining open to emergent contextual variations during coding. Errors and modifications were broadly categorized into:

1. Grammatical errors,
2. Lexical errors,
3. Pragmatic or contextual errors,
4. Cultural or localisation-related issues,
5. Incorrect learner modifications of originally correct AI-generated sentences, and
6. Acceptable learner modifications that simplified or localized AI-generated language without changing meaning.

Particular attention was also given to recurring learner behaviors that later emerged prominently in the findings, including textbook dependence, translanguaging practices, verification behaviors using Google Translate or textbooks, memorization-oriented simplification, and localisation of Mainland Chinese expressions into Malaysian Chinese usage.

Data Analysis

Data were analysed using a qualitative document-based approach focusing on role-play scripts, prompt logs, revision tables, and reflective essays. The analysis aimed to examine learners' prompting practices, mediation of AI-generated language, and perceptions of AI-assisted language learning.

The analysis was conducted in two stages. In the first stage, original ChatGPT-generated scripts were compared with students' revised versions to identify patterns of learner mediation, including acceptance, modification, verification, simplification, and localisation. Through constant comparison across groups, recurring categories were identified, including: (1) actual errors in AI-generated outputs, (2) correct AI-generated sentences incorrectly modified by learners, (3) acceptable modifications of AI-generated language, and (4) contextual or cultural adaptations. This stage focused on how learners evaluated and negotiated AI-generated language according to their linguistic knowledge, communicative goals, and learning needs.

In the second stage, learners' revision explanations and reflective essays were analysed interpretively to identify the reasons underlying their modification decisions. Particular attention was given to references to textbook vocabulary, familiarity of expressions, memorisation strategies, verification practices, contextual adaptation, and communicative appropriateness. Reflective data were further analysed using Braun and Clarke's (2006) six-phase thematic analysis framework. Initial codes emerged inductively from the data and were subsequently refined through iterative comparison to align with the study's research questions concerning prompting practices, learner mediation, and learner perceptions.

Teacher observations were analysed alongside learners' revision explanations and reflective essays to support data triangulation. Observational notes were used to corroborate emerging themes related to verification practices, textbook dependence, performability, and contextual adaptation.

To enhance trustworthiness, investigator triangulation was incorporated throughout the analysis process. The primary researcher conducted the initial coding based on her familiarity with the classroom context and learners' interaction processes. Two co-researchers who were not involved in the teaching or implementation of the role-play task reviewed selected data samples and participated in discussions of coding decisions and emerging themes.

Coding categories were refined through iterative discussions and constant comparison across data sources until consensus was reached. Analytic memos and triangulation across prompt logs, script revisions, reflective essays, and teacher observations were used to enhance the credibility and consistency of the findings.

Researcher Positionality

The primary researcher was simultaneously the CFL lecturer, classroom teacher, and assignment designer. This dual role provided valuable contextual insights into learners' interaction processes while also introducing the potential for researcher bias. To mitigate this risk, two co-researchers who were not involved in the teaching process reviewed selected data samples and contributed to the interpretation of findings. Teacher observations were used as supplementary data for contextualisation and triangulation rather than as evaluative measures.

Ethical Considerations

Ethical approval for this study was obtained from the Research Ethics Committee (REC) of Universiti Teknologi MARA (Approval No. REC/08/2024 (ST/MR/183)). Participation was voluntary, and all students were informed about the purpose of the research. Learners' identities were anonymized during analysis and reporting to ensure confidentiality.

RESULTS

Prompting and Learner–AI Interaction

Analysis of learners' interactions with ChatGPT revealed varying levels of prompting competence and AI-mediated interactional literacy across the five groups. Learners generally constructed prompts based on familiar communicative situations, including travel planning, dining, workplace interaction, illness, and social communication. Most groups requested outputs in Chinese characters together with Pinyin and English translations.

Table 1. The main characteristics of learners' prompting strategies across groups

Group	Main Scenario Types	Prompt Characteristics	Output Features
Group 1	Travel, directions, dining	Detailed contextual prompts; requested Pinyin and translation	Contextually coherent dialogues
Group 2	Travel planning, hotel, restaurant	Multi-scene prompts with explicit instructions	Longer and more complex outputs

Group	Main Scenario Types	Prompt Characteristics	Output Features
Group 3	Vacation, restaurant, illness	Shorter and simpler prompts	Limited variety and shorter dialogues
Group 4	Workplace, travel, social interaction	Structured prompts with role specification	Consistent conversational flow
Group 5	Multi-scene travel role play	Highly extended dialogue requests	Long outputs requiring simplification

The prompts varied considerably in specificity and complexity. Several groups explicitly specified the number of speakers, communicative context, and expected dialogue structure. Groups that provided more detailed prompts generated longer and more elaborate role-play scripts, whereas groups using shorter prompts produced simpler and more repetitive dialogues.

Prompting was also found to be an iterative process rather than a single instruction. Learners frequently refined their prompts when the generated responses did not match their intended scenarios or communicative needs. For example, groups working on travel-related role plays gradually expanded their prompts to include specific locations, activities, and conversational situations. The final scripts incorporated locally familiar settings such as Penang, Langkawi, Kuala Lumpur, Melaka, Ipoh, and tourist destinations including Penang Central, Concubine Lane, and Kek Lok Tong.

Another notable feature was learners' use of multilingual resources throughout the prompting process. Although the final role-play scripts were written in Chinese, learners commonly relied on Malay and English when formulating prompts and interpreting AI-generated responses. They also requested Pinyin and English translations alongside Chinese characters.

Overall, learners frequently refined prompts and added contextual details when interacting with ChatGPT. Rather than simply accepting initial responses, they adjusted prompts according to their communicative goals, language proficiency, and task requirements.

Learner Mediation of AI-Generated Language

Analysis of learners' revisions revealed that ChatGPT-generated language was rarely accepted without modification. Instead, learners actively evaluated, verified, simplified, and adapted AI-generated outputs according to their linguistic knowledge, classroom learning experiences, communicative goals, and local context. Table 2 summarizes the major categories of learner mediation identified across the five groups.

Table 2. Categories of learner mediation in revising AI-generated scripts

Category of learner mediation	Description	Associated revision practices	Illustrative example
Verification and evaluation	Evaluating and verifying the linguistic accuracy, clarity, and appropriateness of AI-generated language using external resources and prior knowledge.	Grammatical accuracy; Enhanced clarity	After consulting textbooks and translation tools, learners revised “ <i>Nǐ hǎo, wǒ huílái le. Wǒ de jiàqī guò dé rúhé?</i> ” (Hello, I'm back! How was my holiday?) to “ <i>Nǐ hǎo, wǒ huílái le. Nǐ de lǚxíng zěnmeyàng?</i> ” (Hi, I'm back. How was your trip?) to improve grammatical accuracy and contextual appropriateness.
Contextual adaptation	Adapting AI-generated language to align with local	Contextual relevance; Cultural relevance	Learners replaced the generic expression “ <i>Wǒmen bìxū chángshì Mǎliùjiǎ zhùmíng de shíwù</i> ” (We

Category of learner mediation	Description	Associated revision practices	Illustrative example
	sociocultural practices, place names, and familiar references.		must try Malacca's famous food) with “ <i>Wǒmen yīdìng yào chī Mǎliùjiǎ zhùmíng de niángǒu lǎshā hé jīfàn</i> ” (We must try Malacca's famous Nyonya laksa and chicken rice). They also corrected inaccurate translations of local place names such as Kuala Lumpur and Kek Lok Tong.
Performance-oriented revision	Modifying AI-generated language to facilitate memorisation, pronunciation, fluency, and oral delivery during role-play performances.	Flow and dynamics; Practicality	Learners simplified “ <i>Shùnbìàn wèn yīxià, nǐ zhàngfū zuìjìn zěnmeyàng?</i> ” (By the way, how is your husband doing lately?) to “ <i>Duìle, nǐ zhàngfū zuìjìn zěnmeyàng?</i> ” (By the way, how's your husband doing lately?) because the revised expression was easier to remember and deliver during performance. Learners also frequently shortened lengthy AI-generated sentences.
Personalisation	Personalising AI-generated language to reflect learners' identities, preferences, experiences, and communicative intentions.	Better expression; Individual style	Learners changed “ <i>Tā jiào Jiékè, láizì Jiānádà</i> ” (His name is Jack, and he is from Canada) to “ <i>Tā jiào Izzatul, láizì Mǎláixīyà</i> ” (“Her name is Izzatul, and she is from Malaysia”) to create scripts that better reflected their own social context and identities.

Verification and Error Evaluation

Learners frequently verified AI-generated language before incorporating it into their role-play scripts. Google Translate, revision books, and previously learned classroom materials were commonly used to check vocabulary and expressions.

Several revisions reflected successful identification of contextually inappropriate outputs. For example, one group replaced yínhángkǎ (bank card) with dìtú (map) because the dialogue involved asking for directions rather than discussing financial transactions. Another group changed tànwàng (to visit someone, usually implying concern or care) to bàifǎng (to pay a visit) after consulting a revision book and determining that the latter better suited the intended context. These examples show that learners consulted revision books, textbooks, and translation tools before incorporating AI-generated language into their scripts.

Teacher observations further confirmed that learners frequently consulted external resources, including textbooks, revision books, and Google Translate, when evaluating AI-generated outputs. During group discussions, learners often compared ChatGPT-generated expressions with vocabulary encountered in previous lessons before deciding whether to retain or revise them.

Learner mediation did not always lead to improvements in AI-generated language. In some cases, learners revised linguistically acceptable expressions based on familiarity or textbook preferences, resulting in alternatives that were less accurate or less natural. For example, *Wǒmen zhōngyú láidào Bīnchéng dùjià le* (“We have finally arrived in Penang for our holiday”) was shortened to *Wǒmen dào Bīnchéng dùjià le* (“We arrived in

Penang for our holiday”), while *tànsuǒ yíxià* (“explore a little”) was replaced with *qù nàlǐ* (“go there”). Learners replaced these expressions with alternatives that had previously appeared in classroom materials.

Simplification and Textbook-Oriented Revision

A recurring pattern across the dataset was the simplification of AI-generated language. Because the completed scripts were later performed in class, learners frequently revised sentences to facilitate memorisation and oral delivery.

One group removed the phrase *wǒ pò bù jí dài xiǎng kāishǐ tànsuǒ* (I cannot wait to start exploring), explaining that the shorter version was easier to remember. Similarly, *hé tiáoliào* (and seasonings) was replaced with *děngděng* (and so on) because learners considered the revised expression easier to use during oral performance. Learners explained that these revisions made the dialogues easier to remember and perform.

Textbook influence was also evident in many revisions. Learners often preferred vocabulary that had already been introduced in class. For example, *qīngcài* (green vegetables) was replaced with *shūcài* (vegetables) because *shūcài* was the term used in the textbook. Likewise, *měiwèi* (delicious) was changed to *hǎochī* (tasty), while *dòngshǒu zuò* (prepare or make by oneself) was simplified to *zuòfàn* (cook).

Classroom observations showed that learners frequently referred to textbooks and classroom notes when selecting vocabulary. When disagreements arose within groups, textbook examples and classroom notes frequently served as reference points for making revision decisions. Teacher observations further revealed that learners frequently shortened lengthy AI-generated sentences during rehearsal sessions.

In some cases, learners retained vocabulary suggested by ChatGPT without modification.

Localisation and Contextual Adaptation

Beyond linguistic revision, learners also adapted AI-generated language to better reflect local Malaysian contexts. Several modifications involved replacing expressions perceived as characteristic of Mainland Chinese usage with forms considered more appropriate in local communication.

A representative example was the replacement of *zhōnghuán chēzhàn* (central transit station) with *zhōngyāng chēzhàn* (central station). Learners explicitly explained that *zhōngyāng chēzhàn* was the term more commonly used in Malaysia when referring to Penang Central. Similar localisation was evident in the incorporation of locally familiar places, food, and tourist destinations such as Penang, Langkawi, Melaka, Ipoh, Concubine Lane, and Kek Lok Tong.

Teacher observations suggested that learners were generally more attentive to contextual mismatches involving local place names, food items, and cultural references than to grammatical inaccuracies. Discussions within groups frequently centred on whether AI-generated expressions reflected familiar Malaysian usage and communicative practices.

Learners’ Perceptions of ChatGPT

Learners generally reported positive experiences using ChatGPT during role-play script construction. Table 3 summarizes the main themes identified from the reflective essays.

Table 3. Learners’ Perceptions of ChatGPT in Role-Play Script Learning

Theme	Positive Perceptions	Reported Limitations
Efficiency	Saves time in script generation	Responses sometimes inconsistent

Theme	Positive Perceptions	Reported Limitations
Language support	Provides Pinyin and translations	Vocabulary sometimes unsuitable for Malaysian context
Creativity	Helps generate ideas and storylines	Some dialogues irrelevant to prompts
Vocabulary learning	Introduces new vocabulary	Difficult expressions require simplification
Accessibility	Easy and convenient to use	Requires verification through textbooks or Google Translate

Many learners described ChatGPT as a useful source of vocabulary, sentence patterns, and dialogue ideas. The immediate availability of responses enabled them to complete tasks more efficiently and reduced the time required for script development. Several students also appreciated the provision of Chinese characters, Pinyin, and English translations, which helped them understand unfamiliar expressions and prepare for oral performance.

At the same time, learners recognised the limitations of AI-generated language. Some participants noted that certain vocabulary items were unfamiliar, overly complex, or inconsistent with expressions introduced in class. Others reported cross-checking AI-generated responses against Google Translate, textbooks, classroom notes, or other reference materials before incorporating them into their scripts.

Many participants described ChatGPT as a supplementary resource rather than a primary source of knowledge. They emphasised that AI-generated language often required further verification and adaptation based on classroom materials and their own linguistic knowledge.

Overall, learners reported that ChatGPT supported role-play preparation, vocabulary learning, and idea generation.

DISCUSSION

Learner Mediation as an Active Constructivist Process

A central finding of this study is that beginner CFL learners did not treat ChatGPT-generated language as ready-made content. Instead, they engaged in a process of evaluation, adaptation, and negotiation before incorporating AI-generated language into their role-play scripts. This finding suggests that learner–AI interaction is more accurately understood as mediation than simple reception.

From a constructivist perspective, learning occurs through active meaning-making rather than the transmission of knowledge from external sources (Vygotsky, 1978; Zajda, 2021). The findings illustrate how learners drew upon prior classroom knowledge, communicative goals, and task requirements when deciding whether to accept, modify, or reject AI-generated suggestions. ChatGPT therefore functioned less as an authoritative source of knowledge than as a mediational resource that learners used while constructing their own understanding (Belda-Medina & Calvo-Ferrer, 2022).

Notably, learner mediation was not limited to correcting errors. Many revisions were made to otherwise acceptable language, reflecting learners’ concerns about familiarity, contextual appropriateness, and oral delivery. This suggests that the educational value of generative AI lies not only in the quality of the language it produces but also in the evaluative processes it provokes. In this sense, AI-assisted tasks may create opportunities for learners to exercise agency, judgement, and active participation in knowledge construction (Bearman & Ajjawi, 2023).

Importantly, learner mediation involved selective acceptance rather than blanket acceptance or rejection of AI-generated language. Learners retained, modified, or discarded suggestions according to their familiarity, communicative goals, and performance needs.

Prompting as an Emerging AI Literacy

One important mechanism through which learners exercised such agency and mediation was prompting. The findings highlight prompting as an emerging component of AI literacy in language learning. Learners who provided richer contextual information generally obtained outputs that were more relevant to their communicative goals, supporting previous claims that prompt quality influences response quality (Ekin, 2026; Lo, 2023).

More importantly, prompting appeared to be an iterative process rather than a single act of instruction. Learners repeatedly adjusted prompts when initial outputs did not align with their intended scenarios or language proficiency levels. This finding extends Zhao et al.'s (2024) argument by showing that prompting serves not merely as a technical skill but as a form of learner regulation through which students guide AI-generated language toward their own learning needs.

The multilingual nature of prompting is also noteworthy. Learners moved flexibly between Malay, English, and Chinese when interacting with ChatGPT, suggesting that multilingual resources functioned as cognitive scaffolds during AI-assisted learning. Rather than restricting themselves to the target language, learners strategically mobilised their full linguistic repertoire to support comprehension, decision-making, and task completion.

Feedback Literacy, Textbook Authority, and Performability

The findings reinforce the importance of feedback literacy in AI-assisted language learning. Learners rarely accepted AI-generated language without scrutiny; instead, they cross-checked outputs against textbooks, classroom notes, revision materials, and translation tools. Such behaviours suggest an emerging awareness that AI-generated language requires evaluation rather than unquestioned acceptance (Carless & Boud, 2018).

However, learners' evaluative judgement varied considerably. While some participants successfully identified contextual inaccuracies and inappropriate expressions, others rejected linguistically acceptable language simply because it was unfamiliar. This pattern suggests that limited evaluative expertise may constrain learners' ability to distinguish between unfamiliar language and genuinely problematic language, highlighting an important challenge in developing feedback literacy in AI-assisted learning environments. This finding echoes concerns that effective engagement with generative AI depends not only on access to AI-generated feedback but also on learners' ability to critically evaluate its quality and appropriateness (Kasneci et al., 2023; Zhan & Yan, 2025).

A particularly noteworthy finding was the continuing influence of textbook authority. Many learners preferred vocabulary that had previously appeared in classroom materials and often regarded textbook-based expressions as more reliable than unfamiliar alternatives generated by ChatGPT. This suggests that traditional sources of linguistic authority continue to shape learners' trust in language resources, even within AI-assisted learning environments.

The findings also highlight performability as an important yet often overlooked dimension of learner evaluation. Because the role-play scripts were intended for oral presentation, learners frequently selected expressions that were easier to remember, pronounce, and deliver during performance. Evaluative judgement therefore extended beyond linguistic correctness to include practical considerations associated with oral communication. This finding resonates with recent work that conceptualises learners' active engagement with AI-generated feedback as a key component of feedback literacy (Jin et al., 2025). Rather than simply identifying errors, learners assessed, adapted, and sometimes rejected AI-generated language according to their communicative goals and performance needs.

Localisation and Contextual Appropriateness in CFL Learning

The findings further demonstrate that learner mediation is shaped by local sociocultural realities. Although ChatGPT generated grammatically acceptable Chinese, learners frequently adapted expressions to better align with Malaysian contexts and communicative practices.

Interestingly, learners appeared more sensitive to contextual mismatches than to linguistic inaccuracies. Expressions that sounded culturally unfamiliar or inconsistent with local usage were often revised, even when grammatical issues elsewhere remained unnoticed. This suggests that contextual appropriateness may play a particularly salient role in learners' evaluation of AI-generated language during communicative tasks.

These findings support Godwin-Jones' (2023) observation that AI-generated language does not necessarily reflect local linguistic and cultural norms. More importantly, they highlight learners' role as cultural mediators. Localisation was not simply a corrective process but a form of sociocultural negotiation through which learners adapted AI-generated language to fit their own communicative realities. This process is especially relevant in multilingual settings such as Malaysia, where language use is closely intertwined with local identities and experiences.

Implications for AI-Assisted CFL Learning

Several pedagogical implications can be drawn from the findings.

First, prompting should be recognised as an emerging component of AI literacy rather than merely a technical skill. As learners' prompts substantially influenced the quality and relevance of AI-generated outputs, CFL instruction may need to provide explicit guidance on how to formulate context-sensitive prompts, refine instructions, and communicate learning intentions effectively when interacting with AI systems.

Second, the findings highlight the need to integrate feedback literacy into AI-assisted language learning. Access to AI-generated language does not automatically lead to effective learning. Learners require opportunities to evaluate AI-generated content critically, verify information across multiple sources, and distinguish between unfamiliar language and genuinely problematic language. Such capacities are essential for developing informed and responsible engagement with generative AI (Carless & Boud, 2018; Zhan & Yan, 2025).

Third, language educators should recognise that learners evaluate language according to multiple criteria beyond grammatical accuracy. In the present study, performability emerged as an important consideration, particularly in tasks involving oral communication. Learners frequently prioritised memorisation, fluency, and ease of delivery when selecting or revising AI-generated expressions. This suggests that pedagogical evaluation of AI-generated language should take into account communicative functionality as well as linguistic correctness.

Finally, the findings underscore the importance of contextual adaptation in AI-assisted language learning. Because AI-generated language may not always reflect local sociocultural realities, learners should be encouraged to evaluate the contextual appropriateness of AI outputs and make necessary adaptations. Supporting such practices may help learners develop greater critical awareness, sociolinguistic sensitivity, and autonomy when using generative AI as a language-learning resource.

CONCLUSION

This study examined how beginner Chinese as a Foreign Language (CFL) learners interacted with ChatGPT during role-play script construction. Drawing on a qualitative case study of third-level CFL learners at Universiti Teknologi MARA (UiTM), the study investigated three interrelated aspects of learner–AI interaction: prompting practices, learner mediation of AI-generated language, and learners' perceptions of ChatGPT-assisted learning.

The findings demonstrate that learners did not simply accept AI-generated outputs. Instead, they actively refined prompts, verified information across multiple sources, and modified AI-generated language according to their linguistic knowledge, communicative goals, and classroom experiences. Learners' revisions frequently involved simplification, contextual adaptation, and localisation to support oral performance and align with Malaysian sociocultural contexts. The findings also highlight the continuing influence of textbook authority and the importance of performability in learners' evaluative decisions.

From a constructivist perspective, ChatGPT functioned primarily as a mediational tool rather than an autonomous source of knowledge (Vygotsky, 1978; Zajda, 2021). The study contributes to the growing literature

on AI-assisted language learning by shifting attention from learning outcomes and perceptions to the interactional processes through which learners evaluate, negotiate, and adapt AI-generated language. By integrating constructivist perspectives with feedback literacy (Carless & Boud, 2018), the study highlights learner mediation as a key mechanism underlying effective engagement with generative AI in beginner CFL contexts.

The findings suggest several pedagogical implications. Language educators should provide explicit guidance on prompting strategies, support the development of feedback literacy, and create opportunities for learners to critically evaluate and contextualise AI-generated language. Rather than treating generative AI as a replacement for teachers or textbooks, instructional practices should encourage learners to use AI as a resource that requires ongoing verification, adaptation, and reflective judgement.

This study is subject to several limitations. It was conducted within a single CFL course involving a relatively small group of learners and focused on one role-play task using the free version of ChatGPT available at the time of data collection. Future research could examine learner–AI interaction across different proficiency levels, language tasks, and educational contexts. Longitudinal and comparative studies may further illuminate how prompting practices, evaluative judgement, and learner mediation develop over time as learners gain greater experience with generative AI technologies.

Overall, the findings suggest that the educational value of generative AI lies not only in the language it generates but also in the opportunities it creates for learners to engage in evaluation, negotiation, and meaning-making. Understanding these mediation processes is essential for developing more effective and pedagogically meaningful approaches to AI-assisted language learning.

REFERENCES

1. Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, 15(3), ep429.
2. Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, 100745.
3. Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology*, 54(5), 1160–1173.
4. Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as AI conversational partners in language learning. *Applied Sciences*, 12(17), 8427.
5. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
6. Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315–1325.
7. Ekin, S. (2026). Prompt engineering for ChatGPT: a quick guide to techniques, tips, and best practices. *Journal of Engineering Technology*, 43(1).
8. Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, 56(2), 489–530.
9. Godwin-Jones, R. (2023). Emerging spaces for language learning: AI bots, ambient intelligence, and the metaverse. *Language Learning & Technology*, 27(2), 6–27.
10. Jin, F. J.-Y., Maheshi, B., Lai, W., Li, Y., Gasevic, D., Chen, G., Charwat, N., Chan, P. W. K., Martinez-Maldonado, R., Gašević, D., & Tsai, Y.-S. (2025). Students' perceptions of generative AI-powered learning analytics in the feedback process: A feedback literacy perspective. *Journal of Learning Analytics*, 12(1), 152–168.
11. Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nemmert, C., Nie, F., Schmid, H., Seidel, T., Stadler, M., Weller, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274.
12. Li, J., Ren, X., Jiang, X., & Chen, C. H. (2023). Exploring the use of ChatGPT in Chinese language classrooms. *International Journal of Chinese Language Teaching*, 4(3), 36–55.

13. Li, X., Li, B., & Cho, S.-J. (2023). Empowering Chinese language learners from low-income families to improve their Chinese writing with ChatGPT's assistance afterschool. *Languages*, 8(4), 238.
14. Lo, L. S. (2023). The CLEAR path: A framework for enhancing information literacy through prompt engineering. *The Journal of Academic Librarianship*, 49(4), 102720.
15. Molloy, E., Boud, D., & Henderson, M. (2020). Developing a learning-centred framework for feedback literacy. *Assessment & Evaluation in Higher Education*, 45(4), 527–540.
16. Rad, H. S., Alipour, R., & Jafarpour, A. (2024). Using artificial intelligence to foster students' writing feedback literacy, engagement, and outcome: A case of Wordtune application. *Interactive Learning Environments*, 32(9), 5020–5040.
17. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
18. Wang, F., Cheung, A. C. K., Neitzel, A. J., & Chai, C. S. (2024). Does chatting with chatbots improve language learning performance? A meta-analysis of chatbot-assisted language learning. *Review of Educational Research*, 94(4), 623–660.
19. Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P., & Li, L. (2023). What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education*, 194, 104703.
20. Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages*, 8(3), 212.
21. Zajda, J. (2021). *Globalisation and education reforms creating effective learning environments*. Springer.
22. Zhan, Y., & Yan, Z. (2025). Students' engagement with ChatGPT feedback: Implications for student feedback literacy in the context of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 1-14.
23. Zhang, Y. (2025). Using ChatGPT in learning Chinese Mandarin: An interactive dictionary rather than a content facilitator. *L2 Journal*, 17(1), 1-28. <https://doi.org/10.5070/L2.42018>
24. Zhao, Q., Hsu, Y. Y., & Huang, C. R. (2024). Large language model and Chinese near synonyms: designing prompts for online CFL learners. *Journal of technology and Chinese language teaching*, 15(1), 49-69.