



Effects and Affordances of AI Chatbot-Mediated Refusal Instruction for Korean Elementary EFL Learners

Sohyun Yoo (Seoul National University)



This is an open-access article distributed under the terms of the Creative Commons License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received: February 19, 2026

Revised: March 21, 2026

Accepted: April 9, 2026

Yoo, Sohyun
Ph.D. Student, Department of
English Language Education
Seoul National University
1 Gwanak-ro, Gwanak-gu
Seoul, Korea
Tel: +82-2-880-7670
Email: yoosh111@snu.ac.kr

ABSTRACT

Yoo, Sohyun. 2026. Effects and affordances of AI chatbot-mediated refusal instruction for Korean elementary EFL learners. *Korean Journal of English Language and Linguistics* 26, 753-771.

The development of pragmatic competence is essential yet often under-addressed in elementary English as a foreign language (EFL) classrooms. To address this gap, this study investigated the impact of brief, curriculum-aligned AI chatbot-mediated instruction on the refusal strategies and perceptions of young Korean EFL learners. Using a mixed-methods approach, this study analyzed data from 57 fourth-grade elementary school students who participated in a two-lesson intervention involving pragmatically intensive (PDR-high) situations. Quantitative findings demonstrated a significant developmental shift from unmitigated, direct refusals toward contextually appropriate, mitigated patterns. Statistical analyses showed robust gains, characterized by medium-to-large effect sizes for key pragmatic shifts, suggesting that AI-mediated instruction was positively associated with pragmalinguistic development. Qualitative data further revealed that learners predominantly perceived AI chatbot interactions as a source of enjoyment and motivation, alongside valuing the immediate feedback provided. Although technological constraints such as speech-to-text (STT) recognition difficulties were frequently reported, learners demonstrated self-regulation by transforming these challenges into opportunities for pronunciation practice. In conclusion, strategically designed generative AI chatbots offer a viable pathway for cultivating early pragmatic competence beyond the limitations of the traditional classroom, provided that their pedagogical integration carefully balances cognitive and technological load.

KEYWORDS

L2 pragmatic instruction, refusal strategies, AI chatbot, pedagogical affordance, young EFL learners

1. Introduction

Developing pragmatic competence, defined as the ability to use language appropriately in accordance with social norms and contextual demands, is a fundamental goal in second language (L2) education. Among various pragmatic acts, refusals are widely recognized as prototypical face-threatening acts (FTAs) (Brown and Levinson 1987), as they can threaten both the hearer's positive and negative face. Performing refusals appropriately necessitates not only linguistic knowledge but also sensitivity to power (P), social distance (D), and the ranking of imposition (R). However, the evaluation of these social variables and the degree to which specific speech acts are treated as FTAs vary greatly across cultures. Such cross-cultural divergence in politeness norms can easily lead to misunderstanding, making the production of pragmatically appropriate refusals particularly challenging for L2 learners.

Despite the recognized importance of pragmatic competence, pragmatic instruction in English as a foreign language (EFL) contexts remains under-addressed, particularly in elementary education. At the research level, empirical studies examining the effects of pragmatic instruction have focused predominantly on adult learners, leaving the effectiveness of such intervention for child learners largely unexplored (Rose 2005). Beyond this research gap, instructional conditions in the classroom further constrain the teaching of pragmatics. Analyses of English as a second language (ESL) and EFL textbooks indicate that pragmatic input is often sparse and largely implicit, with minimal metapragmatic explanation, which limits teachers' capacity to provide explicit instruction and restricts learners' opportunities to develop pragmatic awareness (Vellenga 2004). These structural limitations are especially consequential in elementary classrooms, where instructional time, curricular priorities, and teacher preparation often preclude systematic attention to pragmatics. Consequently, even though young learners can develop pragmatic competence when provided with appropriate input and scaffolding, existing instructional practices often fail to meet these needs (Schauer 2022). This gap frequently results in learners' reliance on linguistically simple or overly direct strategies, hindering their ability to navigate complex social interactions.

Recent advances in artificial intelligence (AI) have enabled the integration of chatbots into language education, offering new avenues for immersive and interactive pragmatic instruction (Erdogan and Kitson 2025). AI chatbots can deliver individualized practice, real-time feedback, and authentic conversational scenarios, supporting the development of pragmatic competence. However, empirical evidence regarding the pragmatic effectiveness of AI chatbots remains limited, particularly for young EFL learners. Existing studies have documented the affordances of AI chatbot-mediated instruction, but few have examined its impact in elementary school classrooms or its role in shaping young learners' pragmatic performance (Ajabshir 2025, Jeon 2024, Rahman et al. 2025, Sung and Kang 2024).

In response to these gaps, the present study investigates the effects of an AI chatbot based on a generative large language model (LLM) for pragmatic instruction on Korean elementary school students' refusal strategies in PDR-high contexts. Adopting an affordance-oriented framework (Gibson 1979, van Lier 2004), this research examines how young learners' refusal performance changes following brief, curriculum-aligned chatbot-mediated instruction, while also exploring learners' perceptions of the pedagogical, technological, and social affordances and constraints of incorporating an AI chatbot in the English classroom.

2. Literature Review

To situate the present study, this section reviews the literature across three intersecting domains. First, it

examines foundational work on L2 refusals, highlighting the sociocultural and cognitive challenges of producing contextually appropriate strategies. Second, it explores the emerging role of AI chatbots in pragmatic instruction. Finally, it discusses the concept of affordances as a theoretical framework for understanding AI chatbot-mediated language learning.

2.1 L2 Refusals

The act of refusal is widely recognized as a prototypical face-threatening act in interpersonal communication (Brown and Levinson 1987). This is because a refusal, by its very nature, often contradicts the wishes, needs, or expectations of the interlocutor. Declining a request, invitation, suggestion, or offer risks damaging the hearer's positive face by signaling disapproval, disagreement, or a lack of support. Simultaneously, such an act threatens negative face as it infringes upon the hearer's fundamental desire for autonomy and self-determination. Consequently, refusals can disrupt social harmony, trigger feelings of embarrassment or disappointment, and potentially jeopardize the relationship between interlocutors. For these reasons, refusals require careful management of politeness and mitigation strategies to attenuate the threat to face.

In determining the requisite level of politeness or mitigation, three key sociological variables—power (P), social distance (D), and the ranking of imposition (R)—play a crucial role. This PDR framework, as outlined in Brown and Levinson's (1987) politeness theory, offers a systematic means of analyzing the contextual factors that influence the choice of linguistic strategies. Power (P) refers to the relative social status or authority of the hearer over the speaker. When the hearer possesses greater power, the speaker is typically compelled to employ enhanced politeness to avoid threatening the hearer's face. Distance (D) captures the degree of familiarity or solidarity between interlocutors. A greater social distance, which implies that the relationship is more formal or less intimate, generally necessitates more polite or indirect refusals. Ranking of imposition (R) pertains to the perceived burden or seriousness of the refusal act in a particular cultural or situational context. Acts that carry a higher level of imposition (e.g., declining a significant request) require more elaborate politeness strategies. Thus, the level of politeness in refusals is not arbitrary but systematically determined by the interplay of PDR.

Refusals often require sophisticated pragmatic competence, especially for L2 learners. Beebe et al. (1990) provided a foundational taxonomy of refusal strategies based on data from Japanese and American speakers. They conceptualized refusals as consisting of a sequence of “semantic formulas” (p. 57), categorizing them into direct, indirect, and adjunct strategies. The classification of refusals in their study is summarized in Table 1.

Beebe et al. (1990) found that L2 learners, particularly Japanese learners of English, tend to overuse apologetic expressions such as “I'm sorry” and frequently provide vague or less persuasive excuses. This pattern suggests a reliance on familiar, formulaic language and limited development of more complex mitigation strategies. Their findings laid the groundwork for subsequent research into the complexity and variability of refusal strategies across different linguistic and cultural backgrounds. Building on this foundation, Chang (2011) demonstrated that Chinese learners of English preferred indirect refusals to preserve social harmony, in contrast to the directness favored by American native speakers. The study further revealed that these cross-cultural differences in both strategy use and perceived appropriateness stem largely from L1 pragmatic transfer. Taguchi (2013) investigated the effect of proficiency on Japanese EFL learners' refusal performance, finding that higher-proficiency learners produced more appropriate and fluent refusals, particularly in formal situations involving higher imposition. Nevertheless, even advanced learners lacked native-like strategic variation, frequently relying on direct refusals and apologies. This body of research underscores that L2 learners, regardless of proficiency level, often fail to employ complex mitigation strategies such as conditional refusals or alternative suggestions.

Table 1. Taxonomy of Refusal Strategies (Adapted from Beebe et al. 1990)

Category	Strategy	Example
Direct Refusals	A. Performative	I refuse.
	B. Non-performative statement	No.
Indirect Refusals	A. Statement of regret	I'm sorry.
	B. Wish	I wish I could help you.
	C. Excuse, reason, explanation	I have a headache.
	D. Statement of alternative	I'd rather...
	E. Set condition for future/past acceptance	If you had asked me earlier, I would have...
	F. Promise of future acceptance	I'll do it next time.
	G. Statement of principle	I never do business with friends.
	H. Statement of philosophy	One can't be too careful.
	I. Attempt to dissuade interlocutor	That's a terrible idea!
	J. Acceptance that functions as refusal	[Unspecific or indefinite reply]
	K. Avoidance	[Silence / Topic shift]
Adjuncts to Refusal	A. Statement of positive opinion/feeling or agreement	That's a great idea...
	B. Statement of empathy	I realize you are in a difficult situation.
	C. Pause fillers	uhh; well
	D. Gratitude/appreciation	Thanks a lot, but...

Exploring cross-linguistic differences, Shishavan and Sharifian (2013) demonstrated that Persian-speaking learners maintained indirect refusal styles across both languages but utilized distinct supportive moves, with apologies and alternatives prevailing in English, and explanations, honorifics, and address terms in Persian. Al-Gahtani and Roever (2018) focused on Arabic L1 ESL learners, showing that proficiency led to improved organization of refusals and greater interactional sensitivity, with advanced learners employing more structural mitigation. However, all learner groups underused native-like pragmatic features such as pre-sequences, idiomatic expressions, and hedging, indicating ongoing challenges in achieving full pragmatic competence. In the Korean context, Park and Oh (2019) found that university students adjusted their refusal strategies according to interlocutor status, producing more indirect and deferential responses to high-power figures. Furthermore, while high-proficiency learners demonstrated interactional flexibility, low-proficiency learners often struggled to produce contextually appropriate refusals due to lexical retrieval challenges.

Collectively, these studies underscore the cognitive, sociolinguistic, and intercultural challenges inherent in L2 refusal production. They show that learners' refusal strategies are shaped by the interaction of L1 transfer, proficiency, and sociocultural context, and that even advanced learners often rely on limited and formulaic mitigation devices. Despite variation across linguistic and cultural backgrounds, a recurring finding is learners' difficulty in producing elaborated and contextually appropriate refusals in situations involving greater social demands. While these foundational studies provide critical insights, they have predominantly focused on adult or university-level learners. Consequently, there remains a notable gap in understanding how young EFL learners navigate these pragmatic complexities.

2.2 AI Chatbots in Pragmatic Instruction

For young learners acquiring L2, aspects of pragmatic competence such as contextual awareness, the

interpretation of implicature, and the use of indirectness are particularly critical. Previous research suggests that noteworthy development in pragmatic comprehension occurs between the ages of 7 and 9, a period during which significant gains have been observed (Lee 2010). Nevertheless, instructional practices and materials in many EFL classrooms often fail to provide sufficient and explicit support for this domain. As a result, young EFL learners frequently lack opportunities to engage in context-sensitive language use, indirect speech acts, and politeness strategies, despite evidence that children are capable of developing sophisticated pragmatic competence when provided with appropriate instructional input and scaffolding (Schauer 2022, Vellenga 2004).

Recent advances in artificial intelligence have facilitated the integration of AI chatbots as interactive and adaptive tools in L2 classrooms. AI-driven systems, such as generative chatbots, have the potential to supplement conventional instruction by affording learners opportunities to engage in conversational practice with simulated dialogue partners. These platforms allow learners to practice pragmatic skills, such as navigating complex speech acts, while receiving immediate, individualized feedback (Erdogan and Kitson 2025). Crucially, AI chatbots provide a low-stakes environment that encourages experimentation with language without the fear of social face-threat (Ajabshir 2025).

Despite the widely recognized potential of AI chatbots for pragmatic practice, empirical studies investigating their actual effects in L2 classrooms have only recently begun to emerge. Ajabshir (2025) compared generative chatbot feedback with teacher feedback for Iranian EFL learners practicing English requests and found greater gains in pragmalinguistic accuracy in the chatbot group, though no significant differences were observed in sociopragmatic outcomes. Sung and Kang (2024) developed a rule-based chatbot for Korean middle school learners to practice requests in PDR-high contexts and reported improvements in learners' pragmalinguistic repertoire and sensitivity to interlocutor status. Similarly, Rahman et al. (2025) examined the effects of an AI-powered chatbot on ESL university learners' pragmatic performance and found significant gains in overall pragmatic appropriateness across requests, apologies, and refusals.

Together, these studies provide emerging evidence for the pedagogical potential of AI chatbot-based pragmatic instruction across different learner populations and speech acts. At the same time, the findings suggest that instructional effectiveness may vary depending on factors such as system type, learner characteristics, and instructional focus. Notably, while positive effects have been reported for adolescent and adult learners, empirical research remains scarce for younger EFL learners, particularly in classroom-based settings. These gaps underscore the importance of designing contextually appropriate AI chatbot interactions that align with instructional goals when extending such technology to elementary-level learners.

2.3 Affordances of AI Chatbots in Language Education

The concept of 'affordance' was first introduced by Gibson (1979), who defined it as the possibilities for action that the environment offers to an organism, stating that "the affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill" (p. 127). This ecological perspective has since been adapted in language education. Van Lier (2004), for instance, argued for a shift from the traditional concept of 'input' to 'affordance' in L2 acquisition, emphasizing that affordances are the actionable opportunities available to the learner in a particular context. These opportunities may be positive or negative, effective or ineffective, depending on the learner's perception and engagement. Kordt (2018) further extended this framework, highlighting that affordances arise from the dynamic interaction between individuals and their environment, and are not simply reducible to input or output. Importantly, affordances may exist without being perceived or may be misperceived, which can lead to learning failures or errors.

To operationalize this complex concept for digital learning, researchers have developed frameworks to categorize the different types of opportunities these environments provide. Kirschner et al. (2004) introduced a comprehensive affordance framework for the design of electronic collaborative learning environments, distinguishing between technological, social, and educational (pedagogical) affordances. Technological affordances refer to system characteristics that determine usability and user-friendliness; social affordances facilitate interaction and communication among learners; and educational affordances support particular types of learning behaviors and instructional goals. In the context of AI chatbots in language learning, Huang et al. (2022) organized chatbot affordances into three main categories. First, technological affordances such as timeliness, ease of use, and personalization enhance the accessibility of language learning and allow for customized educational experiences. Second, pedagogical affordances emerge as chatbots fulfill diverse instructional functions, operating as interlocutors, simulations, information transmitters, helplines, and recommenders to facilitate various learning tasks. Third, social affordances are manifested through the chatbots' capacity to build a sense of social presence, which is achieved by fostering affective, open, and cohesive communication with the learners. While these affordances support increased accessibility and engagement, Huang et al. (2022) also noted that challenges remained, including technological limitations, novelty effects, and cognitive load, and they called for improved chatbot design and further research on long-term impacts.

Empirical studies have further explored these affordances in practice. Jeon (2024) examined Korean elementary EFL learners' 16-week interaction with custom AI chatbots, illustrating how pedagogical, technological, and social affordances function as either opportunities or constraints. While chatbots facilitated oral practice and reduced anxiety, their effectiveness was often mediated by learners' proficiency and technical issues like speech recognition errors. Similarly, Ajabshir (2025) found that generative chatbots provided accessible, non-judgmental environments for practicing pragmatic request strategies. However, despite gains in pragmalinguistic skills, the chatbots' limited grasp of sociocultural context and lack of authentic human interaction constrained deeper sociopragmatic development.

In sum, the literature points to the multifaceted affordances of AI chatbots in language education. These tools can provide meaningful opportunities for language practice, immediate feedback, and learner engagement, but also present distinct challenges. Building on this foundation, the present study seeks to identify and analyze the specific affordances that a pragmatics-focused chatbot offers to elementary EFL learners, with the aim of providing insights into how such tools can be effectively integrated into elementary language education contexts.

2.4 The Present Study

The primary aim of this research is to examine the efficacy of generative AI chatbot-mediated pragmatic instruction for Korean elementary EFL learners, with a specific focus on the development of refusal strategies. In addition, this study adopts an affordance-oriented framework to investigate how AI chatbot-mediated learning can support elementary students' pragmatic development.

This study poses the following research questions:

1. How does generative AI chatbot-mediated instruction influence Korean elementary EFL learners' performance of PDR-high refusal strategies?
2. What affordances do learners perceive while using an AI chatbot for practicing English refusals?

3. Methodology

3.1 Participants

The study initially recruited 71 fourth-grade Korean elementary school students (approximately 10 years old) from three intact classes at a public school located in Incheon. All participants were native speakers of Korean receiving formal English education in accordance with the national English curriculum. At the time of data collection, the students had completed approximately 1.5 years of formal English instruction. Prior to this instructional unit, none of the participants had independent experience using AI chatbot tools for language learning, thereby ensuring a relatively homogeneous baseline in both pragmatic competence and technological familiarity. Of the original cohort, 14 students were excluded based on pre-established criteria: one for not meeting the L1 (Korean) requirement, eight for absences during the intervention or testing, and five for invalid response records. This resulted in a final analytic sample of 57 students.

Through a government initiative, all fourth-grade students at the school were equipped with individual laptops, which were used to complete the pre- and post-tests and to participate in the AI chatbot-based activities. Prior to the study, participants had interacted with the AI chatbot Mizou once in a previous lesson as a conversational partner for asking and answering questions about the location of objects. This session was intended solely to build technological familiarity with the platform, and did not involve any instruction or practice related to L2 pragmatics.

Written parental consent and student assent were obtained for all participating students prior to the study, confirming their voluntary agreement to allow their interaction logs to be used for research purposes. It was explicitly communicated that non-participation in the research would carry no academic disadvantages, and no monetary or material rewards were provided for participation. To ensure participant safety and ethical compliance, all student data were fully anonymized, and the chatbot interactions were conducted within a secure digital platform complying with data privacy laws.

3.2 Procedure

The study employed a design consisting of a pre-test, instructional intervention, and post-test, followed by an additional post-task perception survey. The entire procedure was conducted over a two-week period, with two- to three-day intervals between each phase (pre-test, Lesson 1, Lesson 2, and post-test/survey) to accommodate the regular school schedule. Data collection was embedded within four regular English lessons facilitated by the researcher, who was also the students' regular English teacher. The procedure consisted of three components: oral Discourse Completion Tests (DCTs) administered before and after the instructional intervention, a two-lesson AI chatbot-mediated instructional sequence, and a brief post-task survey eliciting learners' perceptions of chatbot use.

3.2.1 Pre- and post-test

Data were collected through oral DCTs administered preceding and following the two-lesson instructional intervention utilizing the AI chatbot. For these tests, learners completed two oral DCTs that simulated PDR-high scenarios. Given the participants' age and proficiency level, all DCT prompts were presented in Korean to ensure accurate comprehension of the situational context.

The pre-test included two refusal scenarios in which a teacher suggested participating in leisure activities during break time and after class. The post-test employed parallel scenarios that maintained the same PDR configuration

while varying the specific activities to mitigate practice effects. The full wording of all DCT prompts is provided in Appendix 1. To reduce test predictability and elicit more natural responses, the two target DCTs were interleaved with three filler items prompting other speech acts (apologies, requests, and expressions of gratitude).

All testing was conducted in the regular classroom setting. Students were seated individually and wore headsets equipped with microphones. Learners interacted with the AI chatbot on their personal laptops, listened to scenario prompts delivered orally by the chatbot, and responded verbally. The AI chatbot platform displayed the speech-to-text (STT) transcription on the screen, allowing students to verify it before submission. When recognition errors occurred, students manually edited the text or re-recorded their speech. The AI chatbot securely stored all finalized transcripts for subsequent coding and analysis.

3.2.2 Instructional intervention

This study utilized Mizou (<https://mizou.com>), an education-focused generative AI platform, to ensure a safe and age-appropriate digital environment. The platform enabled the creation of customized chatbots that comply with privacy regulations, filter content according to curricular goals, and provide real-time feedback. Its capacity for one-on-one interaction, real-time monitoring, and adaptive pragmatic practice was utilized to facilitate the refusal instruction among the elementary EFL learners.

The instructional intervention was integrated into two regular English lessons, focusing on enhancing learners' pragmatic competence in refusing suggestions. The approach combined brief explicit instruction with extended implicit practice using an AI chatbot, tailored to the developmental and linguistic needs of fourth-grade elementary students. To ensure pedagogical continuity, all intervention activities were modeled on the suggestion-refusal content present in the school's English textbook, thereby aligning the study with the official curriculum.

In Lesson 1, instruction began with an explicit teaching segment, lasting 15 minutes, during which students were introduced to a carefully selected set of refusal strategies appropriate for their age and proficiency level. Using a worksheet (see Appendix 2), the lesson focused on both direct refusals (e.g., "No," "I can't") and simple indirect strategies such as giving reasons ("I'm busy," "I have to study") or postponing ("Maybe next time"), as well as adjuncts to refusal. The instruction deliberately prioritized a curated selection of accessible forms to avoid cognitive overload. The majority of the lesson (approximately 25 minutes) was then dedicated to implicit learning activities. During this phase, students assumed the role of the teacher and made suggestions to the chatbot, which functioned as a student. The chatbot responded by employing polite refusal strategies, providing learners with contextualized exposure to the target language.

In Lesson 2, the focus shifted to productive practice. The AI chatbot assumed the role of a teacher and initiated suggestions, requiring students to formulate their own refusals. The chatbot was programmed to deliver scaffolded feedback in Korean, particularly when a student provided only a direct refusal, thus encouraging the use of indirect strategies to enhance pragmatic appropriateness. The lesson included a 5-minute review and a 5-minute wrap-up, with the remaining 30 minutes dedicated to AI chatbot-based interaction. Feedback in Korean was utilized to ensure clarity given the learners' limited English proficiency and their familiarity with simple classroom discourse.

During these sessions, the AI chatbot initiated various distinct suggestion scenarios based on the textbook unit, such as inviting students to play sports. Because the interactions were self-paced, the total volume of practice varied among individual learners. Based on manual counts of the final submitted turns within the interaction logs stored on the Mizou platform, students produced an average of 7.75 utterances ($SD = 2.60$) in Lesson 1 and 12.60 utterances ($SD = 5.15$) in Lesson 2. The higher mean and individual variance in Lesson 2 reflect the differences in task design. Lesson 1 primarily involved implicit exposure with students initiating brief suggestions. In contrast,

Lesson 2 required active productive practice embedded within a scaffolded feedback loop, naturally prompting multiple conversational turns to refine each refusal response. Below is a representative excerpt illustrating the scaffolded feedback loop during Lesson 2. For the Korean utterances, English translations are provided in brackets on the following line:

AI Chatbot: 지금부터 거절하는 영어 말하기를 연습해 봅시다. Let's play soccer.

[Now, let's practice speaking English refusals. Let's play soccer.]

Student: I can't.

AI Chatbot: 좋은 표현이야! 이제 이유를 추가해 볼까요? 예를 들어 "I'm busy"나 "I'm tired"라고 말해볼 수 있어. 어떤 이유를 추가할 수 있을까?

[Good expression! Now, shall we add a reason? For example, you could say "I'm busy" or "I'm tired". What reason could you add?]

Student: I'm tired.

AI Chatbot: 잘했어요! "I'm tired"는 아주 좋은 이유예요. 이제 전체 문장을 이렇게 말해보면 좋을 것 같아요: "Sorry, I can't. I'm tired." 다시 해볼래요?

[Well done! "I'm tired" is a very good reason. Now, I think it would be good to say the whole sentence like this: "Sorry, I can't. I'm tired." Would you like to try it again?]

Student: Sorry, I can't. I'm tired.

AI Chatbot: 아주 잘했어요! 그 표현이 정말 정중해요. 이제 다른 상황으로 넘어가 볼까요?

[Very well done! That expression is really polite. Now, shall we move on to another situation?]

Across the two lessons, the combination of concise explicit instruction, extended practice with the AI chatbot, and immediate tailored feedback was intended to support both awareness and practical ability in refusal strategy use among young Korean EFL learners.

3.2.3 Post-task survey

Immediately following the post-test, students completed a brief perception survey to evaluate their experiences with the AI chatbot-based activities (approximately 5 minutes). The survey adopted the identical five open-ended questions used in Jeon (2024), shifting the administration format from an oral interview to a written survey. The instrument consisted of the following five open-ended questions:

1. Can you describe how you used the AI chatbots?
2. Did you find using AI chatbots helpful when studying English? Why?
3. What difficulties did you have when studying with the AI chatbots?
4. How did you overcome these difficulties?
5. Do you want to continue to use AI chatbots in English class? Why?

These questions provided qualitative insights into the students' perceived affordances. To ensure clarity and elicit authentic responses from the fourth-grade participants, the survey was administered in their L1 (Korean), and their responses were subsequently translated into English for analysis. Prior to administering the instrument, the researcher provided a detailed explanation of each question to ensure complete comprehension, explicitly assuring students that their responses would not affect their English grades in any way.

3.3 Data Analysis

A mixed-methods approach was employed to analyze both the oral DCT responses and the data collected from the post-task perception survey. Quantitative analysis focused on the transcribed spoken refusal responses generated via the Mizou chatbot. All responses were coded according to the refusal strategy taxonomy proposed by Beebe et al. (1990). To establish coding reliability, all data were double-coded by the author and an independent rater (an elementary school teacher with an M.A. in English education). Inter-coder reliability was established using Cohen's kappa, yielding a coefficient of .927 across the four DCT items, which indicates an almost perfect level of agreement. Any discrepancies between the raters were resolved through discussion until a consensus was reached.

The quantitative analysis was conducted at three levels to comprehensively track changes in learners' pragmatic performance. First, a response-level analysis examined the overall distribution of refusal categories across the pre- and post-tests. Second, a learner-level analysis was performed by calculating refusal indices for each participant. Since the data did not meet the normality assumption, as indicated by Shapiro-Wilk tests, the statistical significance of the pre-to-post changes was assessed using Wilcoxon signed-rank tests. To gauge the magnitude of the instructional impact, effect sizes (r) were calculated for each index, with values interpreted according to Cohen's (1988) benchmarks (small = .1, medium = .3, large = .5). Finally, a token-level analysis was conducted to examine the specific distribution of individual refusal strategies and their subtypes, providing a detailed account of the learners' expanded linguistic repertoire.

Qualitative analysis was performed on the open-ended survey responses. Thematic coding was used to classify learners' reflections and experiences into three categories: technological, pedagogical, and social affordances. This process allowed for a nuanced understanding of how students perceived the affordances of using an AI chatbot to practice English refusals in the classroom context.

4. Results

4.1 Refusal Strategy Use in PDR-high Contexts

To evaluate the impact of the AI chatbot-mediated intervention, the learners' refusal responses were analyzed at three levels: response-level, learner-level, and token-level. First, the overall pragmatic quality of refusal responses was examined by classifying each production into one of four mutually exclusive categories. This procedure involved the categorization of all 228 responses (114 pre-test and 114 post-test items) according to a specific hierarchy of strategic complexity based on the degree of indirectness and mitigation.

Specifically, Direct-only responses were defined as those containing only direct strategies without any accompanying indirect strategies or supportive adjuncts. Indirect-only responses featured at least one indirect strategy but lacked adjuncts. A more sophisticated level of pragmatic performance was captured in the Indirect+Adjunct category. These target-like responses were characterized by the combination of at least one indirect strategy and at least one supportive adjunct. Any utterances that did not meet the functional requirements of the task were labeled as Non-target. This systematic classification ensured that every response was uniquely assigned to a single category. The resulting distribution of these categories across the pre- and post-test phases is summarized in Table 2.

Table 2. Frequencies and Percentages of Mutually Exclusive Refusal Categories

Category	Pre-test <i>n</i> (%)	Post-test <i>n</i> (%)
Direct-only	34 (29.8%)	3 (2.6%)
Indirect-only	62 (54.4%)	89 (78.1%)
Indirect+Adjunct	0 (0.0%)	21 (18.4%)
Non-target	18 (15.8%)	1 (0.9%)
Total	114 (100%)	114 (100%)

The distribution analysis reveals a clear transition from direct, unmitigated refusals to more complex, mitigated forms. In the pre-test, nearly 30% of the responses were Direct-only, often consisting of a simple “No” or “I can’t,” while another 15.8% were Non-target, indicating difficulties in formulating an appropriate response. By the post-test, both of these pragmatically inadequate categories were almost eliminated (2.6% and 0.9%, respectively). Concurrently, the use of Indirect-only strategies surged from 54.4% to 78.1%, making it the predominant response type. A further notable development is the emergence of the Indirect+Adjunct category. While no learner produced this complex pattern in the pre-test, 18.4% of the post-test responses successfully combined indirect strategies with supportive adjuncts, as exemplified by responses such as, “Thank you, but I’m tired. Maybe next time.” This pattern indicates progression from mere task completion toward pragmatically appropriate performance in PDR-high situations.

While these findings illustrate a general shift in the distribution of refusal types, further analysis examined whether these changes were statistically significant at the individual learner level. For this purpose, four learner-level indices were calculated for each participant as proportions (0, 0.50, or 1.00) based on their two responses per test. Notably, the Indirect-included rate was computed to capture any instance of indirectness, while the Indirect+Adjunct rate measured the consistent use of complex mitigation. Due to the violation of normality assumptions (confirmed by Shapiro-Wilk tests), the statistical significance of the pre-to-post changes was assessed using Wilcoxon signed-rank tests (see Table 3).

Table 3. Wilcoxon Signed-Rank Test Results for Learner-Level Refusal Indices

Index	Pre-test <i>M</i> (<i>SD</i>)	Post-test <i>M</i> (<i>SD</i>)	<i>Mdn</i> (Pre/Post)	<i>Z</i>	<i>p</i>	<i>r</i>
Direct-only rate	0.30 (0.43)	0.03 (0.15)	0.0 / 0.0	-4.34	< .001***	.41
Indirect-included rate	0.54 (0.48)	0.97 (0.16)	0.5 / 1.0	5.05	< .001***	.47
Indirect+Adjunct rate	0.00 (0.00)	0.18 (0.28)	0.0 / 0.0	4.35	< .001***	.41
Non-target rate	0.16 (0.34)	0.01 (0.07)	0.0 / 0.0	-2.92	.004**	.27

Note. *M* = mean; *SD* = standard deviation; *Mdn* = median; *r* = effect size.

*** $p < .001$, ** $p < .01$

As shown in Table 3, there were significant changes in all categories. The Indirect-included rate showed the most substantial increase ($Z = 5.05$, $p < .001$, $r = .47$), followed by a significant rise in the Indirect+Adjunct rate ($Z = 4.35$, $p < .001$, $r = .41$). Notably, because the pre-test score for the Indirect+Adjunct rate was zero across all participants ($SD = 0.00$), the Wilcoxon test operated on a degenerate baseline distribution, meaning the statistical significance reflects the absolute emergence of this strategy rather than a comparison of pre-existing variances. Conversely, the Direct-only rate ($Z = -4.34$, $p < .001$, $r = .41$) and Non-target rate ($Z = -2.92$, $p = .004$, $r = .27$) significantly decreased after the intervention. These results suggest that the AI chatbot-mediated instruction was associated with positive changes in learners’ pragmatic competence by promoting more sophisticated and

contextually appropriate refusal strategies.

To identify the specific linguistic resources driving the observed pedagogical shifts, a granular analysis of individual refusal strategy tokens was conducted. Table 4 presents the frequencies and percentages of tokens for each macro-category and their respective subtypes according to Beebe et al.'s (1990) taxonomy.

Table 4. Frequencies and Percentages of Refusal Strategy Tokens by Category

Category / Strategy	Pre-test <i>n</i> (%)	Post-test <i>n</i> (%)
Direct	37 (25.9%)	3 (1.4%)
Direct B (non-performative “No”)	37 (25.9%)	3 (1.4%)
Indirect	88 (61.5%)	194 (88.6%)
Indirect A (statement of regret)	38 (26.6%)	72 (32.9%)
Indirect C (excuse, reason, explanation)	50 (35.0%)	92 (42.0%)
Indirect F (promise of future acceptance)	0 (0.0%)	30 (13.7%)
Adjunct	0 (0.0%)	21 (9.6%)
Adjunct C (pause fillers)	0 (0.0%)	1 (0.5%)
Adjunct D (gratitude/appreciation)	0 (0.0%)	20 (9.1%)
Non-target	18 (12.6%)	1 (0.5%)
Total	143 (100%)	219 (100%)

Note. Percentages are based on the total number of coded strategy tokens in each test (143 tokens in the pre-test and 219 tokens in the post-test). Because a single response could contain multiple strategies, the number of strategy tokens exceeds the number of responses.

The total number of strategy tokens increased from 143 in the pre-test to 219 in the post-test. This quantitative growth reflects an overall increase in the elaboration and linguistic richness of learners' refusal responses following the intervention. In the pre-test, learners relied heavily on Direct-only strategies (25.9%, primarily the non-performative “No”) and simple Indirect strategies, such as providing reasons (35.0%). Notably, no instances of supportive adjuncts (e.g., gratitude) or more complex indirect strategies (e.g., promises of future acceptance) were observed. Furthermore, Non-target tokens, which failed to appropriately realize the intended refusal, accounted for 12.6% of the pre-test data.

In contrast, the post-test profile reflects a substantial shift toward more sophisticated and conventionally polite realizations. Direct tokens were nearly eliminated (1.4%), while the use of Indirect strategies surged to 88.6%. This increase was characterized by the increased use of expressions of regret (32.9%) and the emergence of promises of future acceptance (13.7%). Most significantly, Adjuncts (9.6%), particularly expressions of gratitude (9.1%), appeared for the first time in the post-test. These newly adopted strategies were limited to the specific categories introduced during the AI chatbot intervention, suggesting that the learners' expanded repertoire was a direct result of the targeted instruction and feedback.

4.2 Learners' Perceived Affordances of AI Chatbot Use

To illuminate the subjective experiences behind the quantitative improvements observed above, this section addresses the second research question by thematically analyzing the students' open-ended survey responses within an affordance-oriented framework. The analysis revealed three broad categories: pedagogical, technological, and social affordances. Each of these categories functioned as both an opportunity and a constraint depending on

learners' perceptions and levels of engagement. Table 5 provides an overview of the identified themes and their frequencies.

Table 5. Overview of Perceived AI Chatbot Affordances

Affordance	Theme	<i>n</i>
Pedagogical	Enjoyment and motivation	25
	Immediate feedback and error correction	21
	Exposure to diverse expressions	8
	Perceived realism of interaction	8
	Cognitive load from unfamiliar vocabulary	6
	Perceived artificiality of interaction	2
Technological	STT recognition difficulties	21
	Pronunciation practice via STT errors	7
	System errors, lag, and fast TTS output	6
Social	Non-judgmental, low-stakes practice space	4
	Anxiety from recognition failure	2

Note. All 57 participants completed the five-question survey. The *n* values represent the number of participants who mentioned each theme. Multiple mentions of the same theme by a single participant were counted only once.

The analysis identified pedagogical affordances as the most prominent theme. The most commonly cited opportunity was enjoyment and motivation ($n = 25$), with learners describing the chatbot activities as more engaging than textbook-only lessons and as something they looked forward to. As S15 noted, "It was fun and made me want to do more. I used to dislike English, but with the chatbot I could enjoy it." Similarly, S35 remarked, "Class used to be a little boring with just the book, but now I'm excited for English because we have the chatbot." Given that students had only minimal prior experience with conversational chatbots, these responses suggest a possible novelty effect that may have heightened perceived engagement.

A second, widely mentioned strength concerned feedback. Students emphasized that the chatbot identified errors, modeled more appropriate formulations, and supported pronunciation. The comments included:

It helps with English because it tells me when something is missing or wrong. I like that it speaks accurately and helps. (S3)

The chatbot corrected my English and showed me how to say it. It even helped with pronunciation. (S39)

When I talk with a partner at a similar level, it's hard to find what's missing. With the chatbot, it replaces my answer with a better one, so I want to keep using it. (S41)

Learners also highlighted exposure to a wider range of expressions ($n = 8$). Several perceived this adaptivity as helpful. For instance, S36 noted, "The chatbot taught me new words and even told me when another expression was better," linking this feature to perceived improvement beyond fixed textbook phrases. However, for some ($n = 6$), the same feature operated as a constraint when unfamiliar vocabulary increased cognitive load. As S33 remarked, "Sometimes it told me to use words I hadn't learned, so it was hard." Notably, some students reported self-initiated strategies to manage this difficulty. This was evidenced by S44, who stated, "There were words I hadn't heard, but I asked the AI what they meant," and S52, who explained, "When it used difficult English, I asked in Korean, 'What does that mean?'" These responses illustrate how pedagogical affordances could shift

dynamically between opportunity and constraint depending on learner engagement.

Perceptions of realism were mixed. Some learners ($n = 8$) felt the interaction was realistic and helpful.

It felt like talking with a friend in English. (S9)

It was like speaking with a real person. (S54)

Like talking to a native teacher, and I could ask ‘What does it mean?’ to learn new words. (S14)

It felt like a real conversation with a foreigner, so it helped. (S28)

Conversely, a small number of participants ($n = 2$) perceived the interaction as artificial, noting its robotic nature. This divergence suggests that simulated interaction was interpreted differently across learners, functioning as either a low-stakes practice space or a limitation in authenticity.

Beyond pedagogical considerations, learners’ responses also foregrounded a set of technological affordances, the most salient of which involved STT recognition difficulties ($n = 21$). Students reported misrecognition and unintended language switching, which often led to frustration. To address these frictions, students employed various workarounds, such as speaking more clearly, repeating their utterances, or switching to typing when necessary. For instance, S4 explained, “Maybe my pronunciation wasn’t good, but it kept turning into another word, so I kept correcting it.” Similarly, S36 noted, “Sometimes it didn’t recognize me, but if I tried again or changed my pronunciation a little, it worked.”

Importantly, several learners ($n = 7$) reframed recognition difficulties as opportunities for pronunciation practice, reporting perceived gains. S25 said, “My pronunciation got better, so I want to keep using it.” S44 also reported, “I think my pronunciation improved with the chatbot.” This pattern reflects how technological constraints can be reinterpreted as learning opportunities. Additional constraints included occasional system errors, freezing or lag, slow loading times, and fast text-to-speech (TTS) output, indicating that technical reliability is critical for sustaining affordances in elementary-level classrooms.

Social affordances were mentioned less frequently, but several students emphasized the AI chatbot’s patience and non-judgmental stance, aligning with a low-stakes, non-judgmental practice space. S26 said, “It used polite language, listened well, and felt like a friend. It corrects mistakes and praises me.” At the same time, a minority reported heightened anxiety when recognition failed or when they worried about getting it wrong. S31 stated, “I felt anxious when it didn’t understand me.” Similarly, S37 said, “I worried I might do it wrong and felt nervous.” Thus, although the AI chatbot reduced social pressure for many learners, technology-related uncertainty could still elicit anxiety for some.

5. Discussion

5.1 Refusal Strategies and the Impact of Instruction

This study examined whether brief AI chatbot-mediated instruction can influence young EFL learners’ refusal behavior in PDR-high contexts. The findings indicate a clear shift away from unmitigated Direct-only and Non-target realizations toward mitigated refusals built on indirect cores with supportive adjuncts. Following instruction, learners more frequently combined expressions of regret and reasons with additional supportive moves such as gratitude, while unmitigated “No” responses and off-target replies were substantially reduced. Specifically, the token-level enrichment, characterized by the emergence of gratitude and promises, served as the micro-level driver

that transformed the overall response-level structure from bare refusals to complex, stacked patterns.

The fact that these shifts were statistically reliable at the learner level (Wilcoxon signed-rank tests), accompanied by medium-to-large effect sizes for the positive shifts in pragmatic strategy choices (r ranging from .41 to .47) and a significant decrease in Non-target responses ($r = .27$), suggests that the AI chatbot-mediated intervention fostered consistent and robust pragmatic development within a short period. These results reflect a clear shift toward complex refusal structures, characterized by reduced directness, increased reliance on indirect strategies, and the emergence of stacked mitigation patterns appropriate for maintaining the interlocutor's face in high-imposition interactions (Brown and Levinson 1987). Such substantial gains provide empirical evidence that AI chatbots can serve as an effective scaffold for developing L2 pragmatic proficiency, which has traditionally been considered difficult to acquire through conventional classroom settings.

While foundational research provides extensive insights into L2 refusals, these studies have predominantly focused on adult or university-level learners (Al-Gahtani and Roever 2018, Beebe et al. 1990, Chang 2011, Park and Oh 2019, Taguchi 2013). Consequently, a distinctive feature of the present data is the developmental trajectory observed among young beginners. Although the learners initially relied on single-move formulas (regret and reason), the post-test emergence of target-like combinations (Indirect+Adjunct) reflects a shift toward more elaborated mitigation. Taguchi (2013) observed that such contextually appropriate refusals in formal settings are typically associated with increasing proficiency. What is particularly notable in this study is that initial indications of these sophisticated shifts were observed in a 10-year-old elementary-school cohort after only two lessons. This finding challenges the assumption that complex structural mitigation is exclusively the domain of older learners, suggesting instead that even young beginners may possess the capacity to employ strategic variation when supported by interactive AI scaffolding. At the same time, the absence of fine-grained markers confirms Al-Gahtani and Roever's (2018) note on the enduring difficulty of acquiring micro-pragmatic features, demonstrating that while the overarching structure can be scaffolded, internal lexical richness likely requires extended exposure.

The findings also resonate with a growing body of technology-mediated pragmatics research. Recent studies have demonstrated that AI chatbot-mediated instruction can effectively promote pragmatic competence when interaction is pedagogically structured and supported by feedback (Ajabshir 2025, Rahman et al. 2025, Sung and Kang 2024). Across diverse contexts, AI chatbot use has been linked to improvements in pragmatic appropriateness, heightened awareness of context-sensitive language use, and positive learner engagement. The present study aligns with this work by showing that even brief AI chatbot-mediated practice, when embedded in textbook-aligned content, can lead to observable changes in refusal behavior. In particular, the largest gains were found in behaviors that the chatbot explicitly modeled and reinforced, such as replacing bare refusals with regret- or reason-based responses and appending supportive moves like gratitude.

At the same time, the scope of pragmatic development observed here appears closely tied to the instructional input provided. This instructional dependency is evidenced by the token-level data, where gains were strictly limited to the specific formulas (e.g., 'I'm tired', 'Maybe next time') modeled by the chatbot. The absence of non-modeled micro-devices, such as hedges or fillers, suggests that under brief instructional conditions, young learners' pragmatic production tends to be primed by the forms made salient through explicit teaching and repeated practice, rather than reflecting broader generalization across the full range of refusal strategies.

5.2 Perceived Affordances and Constraints

Learners' perceptions of the AI chatbot further illuminate how these gains were achieved and where constraints emerged. Students consistently highlighted pedagogical affordances related to noticing and feedback, valuing

immediate, individualized explanations delivered in their L1 that clarified why a particular refusal was appropriate or incomplete in context. They also reported heightened engagement and motivation, attributing this to the opportunity for repeatable interaction without social pressure, a pattern consistent with earlier work on AI chatbot use among L2 learners (Ajabshir 2025, Jeon 2024, Sung and Kang 2024). While enjoyment was often mentioned, it was typically linked to perceived learning utility rather than entertainment alone, though the cohort's limited prior exposure to AI chatbots suggests that a novelty effect may have contributed to initial enthusiasm. Access to expressions beyond fixed textbook frames was also valued, although for some learners this adaptivity introduced lexical overload, requiring clarification or simplified alternatives. These findings align with Huang et al. (2022), who identified the novelty effect and cognitive load as persistent challenges in AI chatbot-mediated language learning.

Technological affordances were closely intertwined. Speech-to-text recognition problems were the most frequently reported source of frustration. Such frustration was a byproduct of the pedagogical decision to prioritize oral interaction over text-based input, as the latter was deemed less developmentally appropriate for young learners. However, many learners described adapting their pronunciation or repeating utterances until recognition succeeded, and some even perceived improvements in their pronunciation as a result. This illustrates how the same technological feature can function as either a barrier or a learning opportunity, depending on how learners engage with it. Other constraints, including system latency, fast TTS output, and inconsistent response validation, did not overshadow the overall perceived value of the AI chatbot but shaped the conditions under which its affordances were most effectively realized.

Social affordances were mentioned less frequently but still played a role. Many learners described the AI chatbot as friendly, patient, and non-judgmental, which reduced anxiety and supported experimentation with refusals. Specifically, the AI chatbot's non-judgmental stance provided a 'safe space' for learners to experiment with the socio-pragmatic nuances of refusals without the fear of social face-threat. For a smaller group, however, recognition failures or uncertainty about being understood heightened anxiety, underscoring the need to balance low-stakes practice with sufficient technological reliability. Divergent perceptions of realism further suggest that simulated interaction can provide an effective practice space for some learners, while others may require complementary human interaction to achieve a stronger sense of authenticity and social attunement.

5.3 Pedagogical Implications and Limitations

Taken together, these findings suggest that the integration of AI chatbots into elementary EFL classrooms shows promise in facilitating pragmalinguistic development, even within a brief instructional period. The robust quantitative gains observed in this study appear to be closely linked to the synergistic effect of high learner engagement and the pedagogical affordances of noticing and immediate feedback. Specifically, the AI chatbot's ability to provide individualized, low-stakes practice allowed learners to produce complex refusal structures, such as combining indirect cores with supportive adjuncts, at a pace that might be unattainable in traditional teacher-fronted settings. Furthermore, learners' agentic behaviors in resolving lexical or recognition difficulties indicate the emergence of self-regulation in a technology-mediated environment, suggesting that AI chatbots can foster adaptive learning strategies alongside linguistic proficiency. At the same time, these findings underscore that instructional impact depends on the careful management of cognitive and technological load through calibrated prompts, paced input, alternative response modes, and explicit preparation for voice-based interaction.

Several constraints should be considered when interpreting these results. First, the study employed a one-group pretest-posttest design without a control group, which precludes definitive causal attribution. The observed

changes may reflect the combined effects of explicit instruction, AI chatbot practice, test-retest familiarity, or their interaction. Second, the sample was relatively small and drawn from a single school, limiting the generalizability of the findings. Furthermore, the intervention was brief, leaving questions of long-term durability and transfer to unpracticed contexts unresolved. Additionally, the exclusive focus on PDR-high contexts limits generalizability, leaving it unclear whether learners could appropriately adjust their pragmatic strategies in PDR-low scenarios, such as peer interactions. Reliance on speech-to-text technology, while efficient, introduces potential transcription inaccuracies. To mitigate this, learners were permitted to manually edit the text or re-record their speech. Nevertheless, the inability to trace whether a final response was finalized through unedited STT, oral re-recording, or manual editing introduces a potential hybridity to the data, representing a limitation in strictly isolating pure oral pragmatic competence.

In addition, the qualitative data are based on self-reports that may be subject to social desirability bias, particularly among young learners. Furthermore, the written survey format may underrepresent young learners' deeper perceptions due to their limited writing and expressive abilities at this developmental stage. Another constraint is the reliance on Beebe et al.'s (1990) adult-oriented taxonomy, which may not fully capture the highly simplified, developmental pragmatic behaviors unique to children. Outcomes may also be sensitive to specific system configurations, as the chatbot was optimized as a friendly assistant for fourth-grade learners with a deliberately simplified proficiency setting. Finally, although the sports activities used in the pre- and post-test DCTs were selected from the exact same unit of the official English textbook to ensure curricular and cognitive equivalence, their empirical equivalence in task difficulty was not formally verified through a statistical pilot study, representing an instrumentation limitation.

Given these constraints, future research must utilize controlled designs and larger samples to definitively establish the causal impact of AI chatbot interventions. It is also critical to examine the long-term durability of pragmatic skills and their transfer to unpracticed, naturalistic settings through delayed post-tests and observational methods. Furthermore, systematically manipulating sociolinguistic contexts alongside pedagogical variables and chatbot configurations will help identify the most effective conditions for L2 pragmatic development, moving beyond self-reports to include objective analyses of learner-system interactions.

6. Conclusion

This study investigated the impact of curriculum-aligned AI chatbot-mediated instruction on young Korean EFL learners' refusal strategies and their perceptions of this technology. The findings demonstrate a significant developmental shift from unmitigated, direct refusals toward contextually appropriate, mitigated patterns characterized by indirect cores and supportive adjuncts. Crucially, the rapid emergence of these complex structures, even after a brief intervention, suggests that generative AI chatbots can serve as a supportive tool for pragmatic development at an earlier stage of L2 development.

The qualitative data further elucidate these gains, revealing that the AI chatbot's pedagogical affordances, particularly its immediate feedback and non-judgmental practice environment, played a central role in lowering psychological barriers and fostering learner engagement. While technological constraints such as speech recognition errors persisted, learners' agentic attempts to overcome these hurdles indicate that AI chatbot environments can promote both linguistic proficiency and self-regulated learning behaviors. Ultimately, this research provides empirical evidence that pedagogically integrated chatbot instruction can serve as a promising scaffold for early pragmatic development. By extending communicative practice opportunities beyond the

limitations of the traditional classroom, such tools offer a viable pathway for cultivating nuanced pragmatic competence in elementary EFL contexts.

References

- Ajabshir, Z. F. 2025. A comparative study of teacher feedback and chatbot feedback on second language learners' pragmalinguistic and sociopragmatic competences. *International Journal of Human-Computer Interaction* 41(16), 9881-9892.
- Al-Gahtani, S. and C. Roever. 2018. Proficiency and preference organization in second language refusals. *Journal of Pragmatics* 129, 140-153.
- Beebe, L. M., T. Takahashi and R. Uliss-Weltz. 1990. Pragmatic transfer in ESL refusals. In R. C. Scarcella, E. S. Andersen and S. D. Krashen, eds., *Developing Communicative Competence in a Second Language*, 55-73. New York: Newbury House.
- Brown, P. and S. C. Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge, UK: Cambridge University Press.
- Chang, Y. F. 2011. Refusing in a foreign language: An investigation of problems encountered by Chinese learners of English. *Multilingua* 30(1), 71-98.
- Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Erdogan, N. and C. Kitson. 2025. Integrating AI in language learning: Boosting pragmatic competence for young English learners. *LatIA* 3, 115.
- Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.
- Huang, W., K. F. Hew and L. K. Fryer. 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.
- Jeon, J. 2024. Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. *Computer Assisted Language Learning* 37(1-2), 1-26.
- Kirschner, P., J. W. Strijbos, K. Kreijns and P. J. Beers. 2004. Designing electronic collaborative learning environments. *Educational Technology Research and Development* 52(3), 47-66.
- Kordt, B. 2018. Affordance theory and multiple language learning and teaching. *International Journal of Multilingualism* 15(2), 135-148.
- Lee, C. 2010. An exploratory study of the interlanguage pragmatic comprehension of young learners of English. *Pragmatics* 20(3), 343-373.
- Park, Y. L. and S. Y. Oh. 2019. Korean EFL learners' refusals to requests and their perceptions. *English Teaching* 74(4), 75-103.
- Rahman, G., B. A. Mudhsh, M. Almutairi and M. Kouki. 2025. Optimizing ESL learners' speech act performance: The role of AI-powered chatbots in pragmatic competence development. *Theory and Practice in Language Studies* 15(10), 3187-3198.
- Rose, K. R. 2005. On the effects of instruction in second language pragmatics. *System* 33(3), 385-399.
- Schauer, G. A. 2022. Teaching L2 pragmatics to young learners: A review study. *Applied Pragmatics* 4(2), 137-158.
- Shishavan, H. B. and F. Sharifian. 2013. Refusal strategies in L1 and L2: A study of Persian-speaking learners of English. *Multilingua* 32(6), 801-836.

- Sung, M. and S. Kang. 2024. Developing AI chatbots for pragmatic instruction of Korean secondary L2 English learners. *Korean Journal of English Language and Linguistics* 24, 441-459.
- Taguchi, N. 2013. Refusals in L2 English: Proficiency effects on appropriateness and fluency. *Utrecht Studies in Language and Communication* 25, 101-119.
- van Lier, L. 2004. *The Ecology and Semiotics of Language Learning: A Sociocultural Perspective*. Dordrecht, Netherlands: Springer.
- Vellenga, H. 2004. Learning pragmatics from ESL & EFL textbooks: How likely? *TESL-EJ* 8(2), 1-18.

Examples in: English

Applicable Languages: English

Applicable Level: Primary

Appendix 1. Pre- and Post-test Oral DCT Prompts (Originally written in Korean)

Test Phase	Prompt
Pre-test	During break time, your teacher suggests, “Let’s play badminton with our class.” Everyone else agrees, but you cannot participate at the moment. Refuse the teacher’s suggestion in English. After class, your teacher says, “Let’s play soccer.” Everyone else agrees, but you cannot participate at the moment. Refuse the teacher’s suggestion in English.
Post-test	During break time, your teacher suggests, “Let’s play basketball with our class.” Everyone else agrees, but you cannot participate at the moment. Refuse the teacher’s suggestion in English. After class, your teacher says, “Let’s play baseball.” Everyone else agrees, but you cannot participate at the moment. Refuse the teacher’s suggestion in English.

Appendix 2. Instructional Intervention Worksheet (Originally written in Korean)

Category	Content
Title	Learning Refusal Expressions
Section 1	Direct Refusal - No. - I can’t.
Section 2	Indirect Refusal 2.1 Providing Reasons - I’m busy. - I’m tired. - I have to study. 2.2 Suggesting Future Interaction - Maybe next time.
Section 3	Using Adjuncts 3.1 Expressing Gratitude - Thank you, but ... 3.2 Apologizing - I’m sorry, but ... 3.3 Hesitating - Uhh...
Instruction	In situations requiring a polite refusal, do not refuse directly. Consider the listener’s feelings by utilizing adjuncts and indirect strategies.