Language Learning & Technology
ISSN 1094-3501 CC BY-NC-ND

2024, Volume 28, Issue 1
pp. 1–19

ARTICLE

EFL students' self-directed learning of conversation skills with AI chatbots

Sooyeon Kang, Sinyeon Middle School

Min-Chang Sung, Gyeongin National University of Education*

Abstract

This study investigates the potential of AI chatbots as practice partners in the context of EFL secondary school students' self-directed learning of conversation skills. Using the Dialogflow CX Console at Google Cloud Platform, we designed two sets of AI chatbots based on two target dialogues. A total of 112 eighth-graders at a public middle school in South Korea were divided into two experimental groups: a chatbot group and a worksheet group. The two groups practiced the target dialogues in self-directed activities with chatbots or worksheets and evaluated the quality of the chatbot- or worksheet-based self-directed activities in a questionnaire. In addition, their improvements in conversation skills and self-directedness were measured by a discourse completion task and another questionnaire in pre- and posttests. Results showed that the chatbot activities were evaluated to be better than the worksheet activities, especially in terms of entertainment. We also found that the chatbot group showed greater improvements both in conversation skills and self-directedness than the worksheet group. Based on these findings and qualitative analyses of dialogue records, implications for using AI chatbots in the self-directed learning of English are discussed.

Keywords: *Chatbots, Self-directed Learning, Conversation Skills, Korean Middle School English*

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

APA Citation: Kang, S., & Sung, M. (2024). EFL students' self-directed learning of conversation skills with AI chatbots. *Language Learning & Technology*, *28*(1), 1–19. https://hdl.handle.net/10125/73600

Introduction

It has always been intriguing that English can be learned in self-directed ways (Holec, 1996; Shawar & Atwell, 2007). Self-directed learning involves individuals proactively identifying their learning needs, setting educational objectives, seeking out resources, and assessing the results of their learning efforts, all of which can be performed with or without assistance from others (Knowles, 1975). Evidence suggests that self-directed learners often achieve favorable outcomes in language learning (Jeon, 2022; Ryu, 2017).

The growing interest in self-directed learning has provoked many meaningful innovations in the use of artificial intelligence (AI) in English language learning and teaching (Fryer et al., 2020; Hsu et al., 2023; Jeon, 2023; Sung & Kang, 2024; Wang et al., 2017; Yang et al., 2022). Indeed, AI technology is of high relevance to English language education because it offers many language-related functions such as natural language processing, virtual reality, and speech-to-text transformation (Pap, 2021; Wang et al., 2017). Among a variety of AI technologies for English language learning, the present study focuses on AI chatbots as a novel tool for enhancing conversation skills in self-directed ways.

AI chatbots have been increasingly used in English language education (Huang et al., 2021). AI chatbots are distinguished from other language learning technologies (for example, interactive websites) as they simulate intelligent conversations with human users through text or audio interactions (Kim et al., 2019). In early studies, English learners were asked to talk with publicly available chatbots (for example, ALICE, Jabberwacky) and find information about various topics, but the learning outcomes were quite limited since these chatbots were developed not for pedagogical purposes but for technological purposes, namely human-

*** Corresponding Author:** Min-Chang Sung, mcsung@ginue.ac.kr

like conversational agents (Coniam, 2008; Fryer & Carpenter, 2006). Recently, researchers began to develop their own AI chatbots for pedagogical purposes and use them with English learners (Höhn, 2019; Hsu et al., 2023; Kim, 2024; Lin & Mubarok, 2021; Shawar, 2017; Sung, 2021). These pedagogically customized chatbots have provided new ways of self-directed learning for a variety of English learners with different goals such as college students' learning of English grammar (Wang et al., 2017) and elementary students' learning of speaking skills (Yang et al., 2019).

In line with the increasing importance of chatbots in English language education, we develop and use AI chatbots for secondary school students who are learning English as a foreign language (EFL). Considering that the core feature of chatbots is the conversation itself, we aim to promote the students' self-directed learning of English conversation skills via chatbot activities. In addition, we address a notable research gap by evaluating the use of chatbots in comparison to other tools, an aspect that has been sparsely explored in previous studies (Huang et al., 2021). To this end, we design another type of self-directed activity with worksheets and compare two types of activities, chatbots versus worksheets, in terms of general evaluation, conversation skills, and self-directedness. Thus, the present study asks the following research questions:

1. Are the chatbot activities evaluated with higher scores than the worksheet activities, in terms of experience and willingness?
2. Do the chatbot activities improve EFL learners' conversation skills more than the worksheet activities?
3. Do the chatbot activities improve EFL learners' self-directedness in learning English more than the worksheet activities?

Previous Studies

Chatbots as Language Learning Tools

A chatbot is a software program that can interact with users using natural languages through audio or text. The first chatbot *Eliza* was developed in 1964 by MIT professor Joseph Weizenbaum for psycho-therapeutic purposes. Since then, many chatbots have been developed in diverse disciplines such as psychology, marketing, and education (Lee et al., 2020). Now, chatbots are a ubiquitous and vital part of people's daily routines.

Chatbots have been reported to have many advantages as conversational partners for language learners (Fryer et al., 2017; Jia & Ruan, 2008; Shawar, 2017). First of all, talking with chatbots is less expensive and more convenient than talking with native speakers. Apart from price and convenience, Fryer and Carpenter (2006, pp. 9–10) pointed out six important advantages of chatbots as language learning tools:

(a) learners feel more relaxed and comfortable when talking to a chatbot;
(b) learners can use a variety of language structures and vocabulary in chatbots;
(c) chatbots are new and interesting to students;
(d) chatbots are willing to repeat the same material endlessly;
(e) many chatbots provide both text and synthesized speech, allowing learners to practice both listening and reading skills;
(f) chatbots can provide quick and effective feedback for students' errors.

In a similar vein, Okonkwo and Ade-Ibijola (2021) highlighted that chatbots have pedagogical benefits such as the integration of diverse contents, motivation, and engagement, multiple users' access, and immediate assistance. The potential of chatbots in language learning has been identified in diverse domains such as vocabulary enhancement (Jeon, 2023), corrective feedback (Kim, 2024), and the development of speaking skills (Dizon, 2020; Yang et al., 2022).

Chatbots for Learners' Conversation Skills

Chatbots have been used to help English learners improve conversation skills (for a review of chatbot

studies, see Huang et al., 2021). Coniam (2008) investigated how undergraduate trainee teachers, all advanced learners of L2 English, evaluated six English-speaking chatbots in three linguistic levels (that is, word, sentence, and text) and proposed that chatbots as 'conversation practice machines' (Atwell, 1999) should have useful pedagogical features such as motivating learners and correcting errors. Relevant innovations were made in a recent study by Hsu et al. (2023). They developed a task-oriented chatbot system for EFL university students' speaking practice which was able to ask a topic-related question, analyze the learner's voice message, and give corrective feedback to an incorrect message. The chatbot's oral training led to significant improvements in the learners' spoken English. Lin and Mubarok (2021) highlighted the role of chatbots in remote learning conditions. They employed mind-map-guided AI chatbots in a university flipped class on English speaking and found that the chatbots helped the students organize information and perform speaking tasks.

The use of chatbots was also found to be effective among younger learners. For example, Ruan et al. (2019) investigated five preschool Chinese-speaking children's use of *BookBuddy*, an English-speaking chatbot that can interact with a learner based on reading materials. The chatbot was able to recommend a book based on the learner's level and interests, answer the learner's questions during the reading stage, and assess the learner's understanding of the book. It was found that the children happily engaged in the conversation with the chatbot. One participant commented that "the chatbot was patient, friendly, and non-judgmental" (p. 3). In the study of Yang et al. (2019), a chatbot named *Ellie* was developed for elementary school learners of English to perform conversation tasks. The learners employed various conversation strategies during the tasks and later reported that the chatbot was very helpful in learning English. However, to our knowledge, there has been no study that examines the effects of using chatbots on secondary school students' learning of conversation skills. This limitation in previous research on chatbots for language learning is addressed in the present study.

Chatbots and Self-directed Learning

In his seminal book, Malcolm S. Knowles (1975) defined the concept of self-directed learning as follows:

> Self-directed learning is a process in which individuals take the initiative, with or without the help of others, in diagnosing their learning needs, formulating learning goals, identifying human and material resources for learning, choosing and implementing appropriate learning strategies, and evaluating learning outcomes (p. 18).

This definition is, of course, applicable to the use of chatbots for enhancing conversation skills. Shawar and Atwell (2007) argued that learner autonomy can be fostered when chatbots are adapted or retrained on specific topics (for example, computing) and languages (for example, spoken Afrikaans) and are able to provide learners with relevant questions, responses, and feedback.

In this regard, Jia and Ruan (2008) and Jia (2009) developed a chatbot named CSIEC using textual knowledge and reasoning for Chinese-speaking learners of English. Using the semantic framework of WordNet (Fellbaum, 1998) and the syntactic framework of Natural Language Markup Language, this chatbot was able to process a huge range of learner messages. For example, if a learner says, "My classmate will be an officer," the chatbot analyzes that *be an officer* is syntactically a main verb phrase and *classmate* is semantically a kind of *person*. Based on the analysis, the chatbot generates a follow-up question, "How does your classmate like being an officer?" In the 2008 study, 41% of the learners voluntarily used the user log to review their exercise history, which may relate to the self-directed evaluation of learning outcomes. In the 2009 study, the percentage of the long (10-to-50 rounds) and longer (over 50 rounds) dialogues with the chatbot CSIEC (51.1%: 912 out of 1,783 cases) was measurable, and a majority of the chatting (84.7%) showed the free pattern, not the scenario one. These results may show that the free chatting quality of CSIEC helped English learners determine to stay and continue to practice conversation skills.

Indeed, many English learners nowadays use chatbots for learning English in self-directed fashions. They diagnose their needs for learning English, identify chatbots that appear to be of high relevance to their goals, and use the chatbots anytime and anywhere (Jeon, 2022). Although chatbots help us to learn English in a

self-directed manner, the relationship between chatbots and self-directed learning has not been explicitly examined. Thus, the present study evaluates the effects of chatbots on self-directed learning of English compared with those of other learning tools.

Present Study

Participants

Participants included 112 eighth-graders from six intact classes at a public middle school in South Korea. All of the participants were 13 or 14 years old except one 17-year student, and there were 54 females and 58 males. They had completed five years of formal English education and were getting English lessons four hours a week from a Korean English teacher, one of the researchers in the present study. Their English textbook covered various topics targeting all four skills. For listening and speaking, each textbook chapter contained two target communicative functions with example dialogues, closed role-plays, and other speaking activities.

The participants' English proficiency was measured through an online English test at Cambridge Assessment English (https://www.cambridgeenglish.org/test-your-english). The test is composed of 25 multiple-choice questions such as answering conversation questions and choosing appropriate expressions for blanks in sentences. Based on the test scores, their overall proficiency turned out to be A2, a basic level according to the Common European Framework of Reference for Languages (CEFR). The six intact classes were randomly divided into two treatment groups: three classes were the chatbot group ($N = 62$), and the other three classes were the worksheet group ($N = 50$).

Procedures

The overall procedures of the present study are illustrated in Table 1. In the pretest, a discourse completion task (DCT) and a questionnaire survey about self-directedness were conducted. The DCT was designed to measure whether the participants were able to produce appropriate utterances in the target dialogue contexts, and the questionnaire survey was developed to investigate how the participants evaluated their own self-directedness.

Table 1

Procedures of the Study

Procedure	Composition
Pretest	DCT, Questionnaire about self-directedness
Lesson 1	Self-directed preview, Pair work, Self-directed review
Lesson 2	Self-directed preview, Pair work, Self-directed review
Posttest	DCT, Questionnaires about self-directedness and activity evaluation

After the pretest, there were two 45-minute lessons where the participants practiced two dialogues in their textbook. Each lesson was composed of three student-centered activities: a self-directed preview, a pair work, and a self-directed review. The two self-directed activities (that is, preview and review) were designed to be the major between-group difference as one group practiced the dialogues with AI chatbots, while the other group practiced them in a traditional worksheet-based method. In the pair work, the participants in both groups got into pairs and verbally practiced the target dialogues.

In the posttest, which was conducted four days after the pretest, the DCT and the questionnaire survey about self-directedness were administered again with another questionnaire survey about activity evaluation. Details about the procedures and materials will be discussed in the following sections.

Lessons

The two lessons focused on two information-gap dialogues. In the first dialogue, a staff and a customer had a conversation at a lost-and-found center to exchange relevant information and find the lost item. In the second dialogue, a school nurse and a student engaged in a conversation at the nurse's office to figure out the student's health problem and solutions. The two contexts entailed a variety of communicative functions such as description (for example, customer: *It has a star pattern on it*) and advice (for example, school nurse: *Make sure you take a good rest*).

The participants practiced each dialogue in three types of activities in a single lesson. First, they performed a self-directed preview either with chatbots or worksheets and submit their preview records. Second, they engaged in pair work to practice the target dialogue verbally with their partners. Third, they performed a self-directed review of the dialogue either with chatbots or worksheets. Note that the two groups used different materials (that is, chatbots versus worksheets) in the self-directed preview and review.

Self-directed Learning with Chatbots

We developed four AI chatbots so that the students in the chatbot group could practice four roles in the two dialogue contexts: customer-staff and school nurse-student (for the chatbot architectures, see Appendix A). As a development platform, we used the Dialogflow CX Console, which has been used in many recent studies to develop a variety of chatbots for language learners (Lee et al., 2020; Muhammad et al., 2020; Sung, 2021; Sung & Kang, 2024). The Dialogflow CX has many advantages such as flowchart visualization, machine learning of user responses, and self-testing systems.

The developed chatbots could function both orally and textually, but the present study employed a text-based approach. This is because, when 20 or more students in a class simultaneously engage in oral dialogues with a chatbot, the noise may disrupt the chatbot's speech recognition and dialogue management, eventually interrupting self-directed learning.

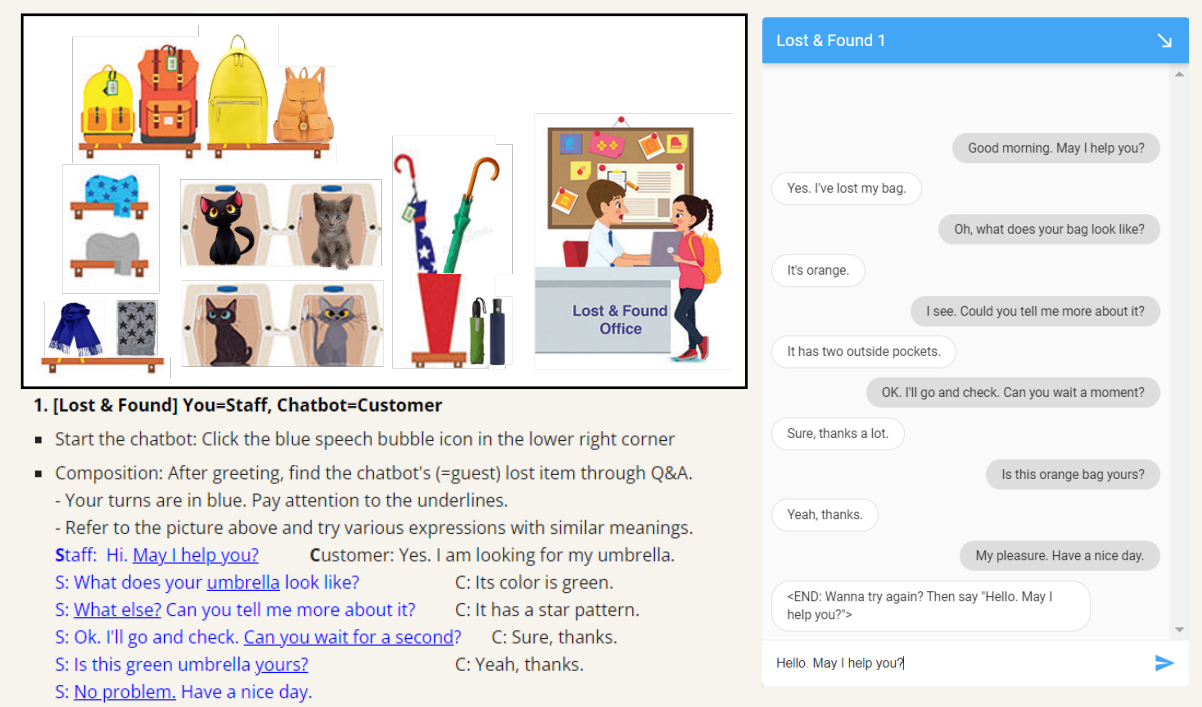
To practice the target dialogues in self-directed ways, every student in the chatbot group was provided with a personal laptop and a web address, where two chatbots representing two roles in the target dialogue had been installed. The two chatbots of each dialogue were sequenced in a way that the students would practice an easier role before the more difficult one. For example, the first chatbot of the lost-and-found dialogue played the role of the customer so that the students could practice the role of the staff with a small number of conventional expressions such as *may I help you?* and *wait a moment, please.* On the other hand, the second chatbot played the role of staff, and the students practiced the role of the customer in a variety of situations with different lost items. To promote self-directedness, the students were allowed to decide when they would move from the first to the second chatbot. For example, low-level students could practice the role of staff many times before they felt ready to practice the role of customer.

The interface of the website had three sections for each chatbot (see Figure 1). First, the left-top section showed the picture of the dialogue context to help the learners (a) understand the situation, (b) choose a conversation topic such as a green umbrella or a yellow bag, and (c) engage in dialogue using the visual information. The picture was adopted from the participants' English textbook (Yoon et al., 2018).

Second, the left-bottom section provided important information about the chatbot activity such as the roles of the user and chatbot, how to open the chat window, and a sample script. The sample script colored the students' turns in blue and underlined key expressions, but the students were encouraged to use different expressions that they thought were appropriate for the context. In fact, the chatbots were designed to understand various synonymous expressions. For example, the customer in the sample script said, "I have lost my cat," but the chatbot was able to interpret other synonymous expressions such as *I am looking for my cat* and even learner errors like *I looking for my cat* (*be*-omission).

Figure 1

User Interface of a Chatbot for the Lost-and-found Dialogue



Third, on the right-hand side was a window for Dialogflow Messenger, a text-based chat between the AI agent and a human user. In the chat window, the grey speech bubbles on the right-hand side show the user's inputs whereas the white speech bubbles on the left-hand side show the chatbot's outputs. When a dialogue was over, the chatbot announced the termination of the dialogue session and explained how to restart another session. When students found that they needed more practice, they could initiate another round of conversation.

Before the students in the chatbot group began the self-directed preview, the teacher demonstrated how to use the chatbots. After the teacher's demonstration, the students practiced the target dialogue with two chatbots for 10 minutes at their own rates and submitted their dialog records to a Google form. The self-directed review was conducted in the same way, with the only difference being the absence of the teacher's demonstration.

Self-directed Learning with Worksheets

The worksheet-based practice had similar designs to the chatbot-based practice. First, it also had a preview and a review part. Second, each part had two sets of scripts so that the students in the worksheet group could practice the two roles of each dialogue. Third, it also provided a picture describing the dialogue context. Fourth, the students used 10 minutes in a self-directed manner, so they could decide their paces and goals.

The major difference is that the students in the worksheet group were provided with a 2-page worksheet to preview or review a target dialogue in a traditional pencil-and-paper approach. Each page had three dialogue scripts based on the varying factors in a dialogue context such as lost items and symptoms, and the three scripts showed one role's utterances with every turn of the other role blanked out, as in Figure 2. The utterances to fill in the blanked turns were presented in a random order in a separate box next to each script, so the students had to find the appropriate utterance for each blanked turn and write it in.

Figure 2

An Example Script in the Worksheet

① S: Staff / C: Customer

S: _____

C: Yes. I am looking for my umbrella.

S: _____

C: Its color is green.

S: _____

C: It has a star pattern.

S: _____

C: Sure, thanks.

S: _____

C: Yeah, thanks.

S: _____

<Examples>

- Hi. May I help you?
- Can you tell me more about it?
- No problem. Have a nice day.
- What does your umbrella look like?
- Is this green umbrella yours?
- Ok, I'll go and check. Can you wait for a second?

Data Collection Tools

Discourse Completion Task

We developed a discourse completion task (DCT) to measure the participants' conversation skills. It examined whether they could produce appropriate utterances in the two target contexts, a lost-and-found center and a school nurse's office. The task design was based on Blum-Kulka et al. (1989). Each question provided an illustration of a dialogue context and a part of the dialogue. There were four to six turns between two interlocutors, and one interlocutor's turns were blanked out (see Figure 3).

Figure 3

An Example Question of the Discourse Completion Task

Look at the picture and the dialogue situation in each item and write what you would say.

(1) 

I am a staff at a lost-and-found center. A customer wants to find his scarf.

I (Staff): _____

Customer: Yes, please. I am looking for my scarf.

I (Staff): _____

Customer: It is red and long.

Note. The task instruction and the illustration of the dialogue context were originally written in Korean.

This task design was intended to promote the transfer of conversation skills (DeKeyser, 2020): the cognitive operations in the DCT resembled those in the treatment so that the transfer of conversation skills from the instruction to the DCT would occur.

There were 20 blank turns over the eight dialogue situations, but only the answers to 14 turns were scored. The rest six blanks were excluded from the analysis to avoid assigning double scores to a single communicative function or giving scores to very simple answers such as *yes* and *no*.

We followed the scoring method of Kim and Taguchi (2015) and gave two-to-zero points to each answer. Two points were given to the answer that performed the target communicative function in a grammatical form. One point was given if the answer performed the target function but was not grammatically accurate. No points were given if the answer failed to perform the target function. After specific scoring standards for each turn were developed by both researchers, all the responses were scored by one of the researchers. The reliability of the DCT was estimated using Cronbach's alpha: .829 in the pretest and .852 in the posttest.

Questionnaires

Two questionnaires were used in the present study, and their overall designs are provided in Table 2 (for the complete list of statements, see Appendix B). The first questionnaire was developed with reference to Pintrich and de Groot (1990) and Ryu (2017). It examined the participants' self-directed learning of English in terms of three core domains, namely metacognition, cognition, and behavior. Each domain included three items which addressed three different themes. For example, the three items in metacognition examined whether one plans, monitors, and regulates the learning of English.

Every item was composed of a statement written in Korean and a 5-point Likert scale, from 1 (*totally disagree*) to 5 (*totally agree*). This questionnaire was administered both in the pre- and posttests to examine the change in the students' self-directedness after the two lessons.

Table 2

Compositions of Questionnaires about Self-directed Learning and Activity Evaluation

Type	Domain	Theme
Self-directed Learning	Meta-cognition	planning, monitoring, regulating
	Cognition	rehearsal, elaboration, organization
	Behavior	time environment, effort increase, help-seeking
Activity Evaluation	User experience	entertainment, convenience
	Learner experience	authenticity, improvement, motivation, participation
	Willingness	future use, recommendation

The other questionnaire was developed to investigate the students' evaluation of the chatbot or worksheet activities. We extracted three categories and eight themes for the evaluation of dialogue activities from recent studies on English-speaking chatbots (Kim, 2016; Yang et al., 2019) and composed eight evaluation statements on the 5-point Likert scale. We also included two open-ended questions: one asked about the advantages of the self-directed activities the participants performed, and the other asked about how to improve the activities. This questionnaire was administered only in the posttest because it investigated the evaluation of the instructional activities. The reliability of the questionnaires was estimated using Cronbach's alpha, .846 in the pretest and .898 in the posttest.

Results

Activity Evaluation

The students' evaluation of the self-directed activities was measured by three categories covering eight themes: user experience (entertainment, convenience), learner experience (authenticity, improvement, interest, participation), and willingness (future use, recommendation to others). The total mean score of the chatbot group (32.24) was greater than that of the worksheet group (31.24), but the total mean difference was not statistically significant, $t = 0.992$, $p = .324$ (see Table 3).

Table 3

Between-group Comparison of Students' Evaluation of Self-directed Activities

Category	Theme	Chatbot (N = 62)	Worksheet (N = 50)	Mean difference	t	p
User Experience	Entertainment	4.16 (0.79)	3.86 (0.81)	0.303	1.981	.050*
	Convenience	4.08 (0.84)	4.08 (0.75)	0.001	0.004	.997
Learner Experience	Authenticity	3.87 (0.93)	3.74 (0.83)	0.131	0.777	.439
	Improvement	4.26 (0.75)	4.18 (0.69)	0.078	0.569	.570
	Motivation	3.61 (1.03)	3.60 (0.78)	0.013	0.073	.942
	Participation	4.15 (0.85)	4.02 (0.87)	0.128	0.777	.439
Willingness	Future use	3.97 (0.96)	3.84 (0.91)	0.128	0.717	.475
	Recommendation	4.21 (0.70)	3.92 (0.97)	0.290	1.834	.069†
Total		32.24 (5.40)	31.24 (5.21)	1.002	0.992	.324

Note. The figures in the parentheses are standard deviations.

Notably, the chatbot group showed higher mean scores in every theme than the worksheet group, as indicated by the positive integers in the column of mean difference. In particular, the mean difference in *entertainment* was found to be significant (p = .050). In other words, the chatbot group had greater fun in the self-directed activities that the worksheet group. Indeed, in the open-ended question about the advantages, 11 students in the chatbot group (19.3%) said that the chatbot activities were fun.

Other students in the chatbot group reported a variety of advantages of the chatbot activities such as the repeated practice of dialogues, error feedback, vocabulary learning, and authenticity. These answers show that the students in the chatbot group had self-directed learning: they repeatedly practiced the target dialogues with the chatbots; they responded to error feedback by self-correcting their previous turns; they learned new expressions from the chatbots and used the expressions in meaningful ways; they immersed themselves into a virtual context and had a conversation with the chatbots as real.

On the other hand, the worksheet group reported other types of advantages. Many students in the worksheet group explicitly answered that the worksheet activities had fostered self-directed learning. In their answers, there were several relevant expressions such as *jagijudojeog hagseub* 'self-directed learning' and *jagi seuseulo* 'by oneself'. They discussed the effects of self-directed learning in terms of problem-solving skills, long-term memories, and diagnostic awareness. The worksheet group also had a small number of students (N = 5) who answered that the worksheet activities were fun, but the reasons seemed to be different from the chatbot group: the chatbot group focused on the experience of having a conversation with the chatbot, whereas the worksheet group focused on the feelings of accomplishment and improvement. For example, a student in the worksheet group reported, 'it was refreshing and fun to fill in the worksheet without the teacher's help [author translation]'.

The other open-ended question asked about how to improve the self-directed activities. The most frequent response from the chatbot group (36 students: 58.1%) was that the chatbot should be designed to understand various utterances. The chatbots were able to play four roles in two contexts, so utterances irrelevant to the roles and contexts were incomprehensible to the chatbots. It was also suggested that the chatbots should be equipped with technical support for those who are slow to type in English. On the contrary, the most frequent response from the worksheet group (8 students: 16%) was that the assigned time was not enough to fill out all the blanks. Another frequent response was that the worksheet activities were repetitive and boring.

Conversation Skills

Table 4 summarizes the means and standard deviations of the two groups' DCT scores in the pre- and posttests. In the pretest data, which met the assumption of normality (W = .978, p = .052), the mean of the chatbot group (13.9) was lower than that of the worksheet group (14.9), but the mean difference was found to be statistically insignificant by an independent samples t-test, $t(110) = 0.838$, p = .404. Thus, the present study assumed that the two groups had the same level of conversation skills before the lessons. In the posttest, the mean scores of both groups increased, and the mean increase was greater in the chatbot group. Consequently, the chatbot group achieved a higher mean score in the posttest than the worksheet group.

Table 4

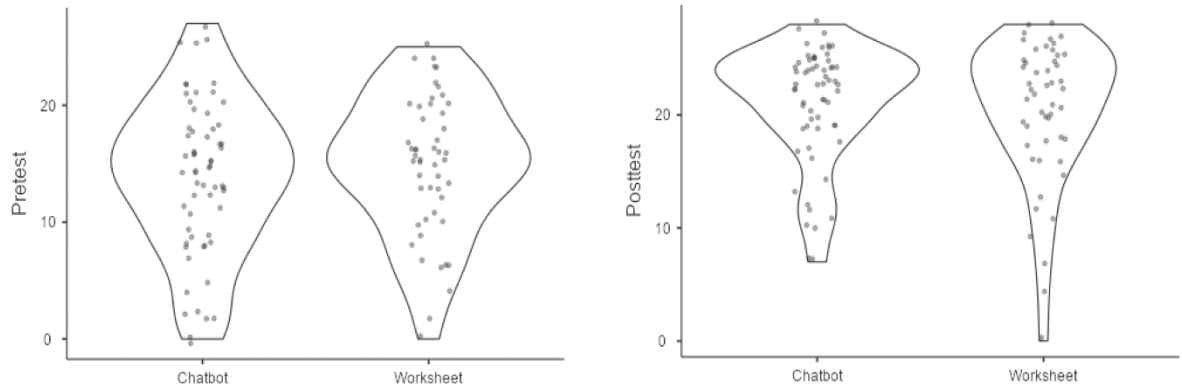
Means and Standard Deviations of DCT in the Pretest and Posttest

Group	Pretest		Posttest	
	Mean	***SD***	**Mean**	***SD***
Chatbot (N = 62)	13.9	6.58	21.0	5.13
Worksheet (N = 50)	14.9	5.96	20.4	6.29
Sum	14.3	6.30	20.7	5.66

Another noteworthy result is that the standard deviation of the chatbot group decreased between the pretest and the posttest from 6.58 to 5.13. This result contrasted with the increase in the standard deviation in the worksheet group. The two groups' changes are effectively illustrated by the violin plots in Figure 4, which shows the spread and variation within the data.

Figure 4

Violin Plots of Individual DCT Scores in the Pretest and Posttest



In the pretest, the two groups were similar in the distribution of individual scores, as indicated by the two similar violin plots. However, in the posttest, the two groups showed very different distributions of individual scores. The violin plot of the chatbot group showed a higher density around the mean score (21.0) than that of the worksheet group. In addition, the vertical range of the chatbot group was narrower than that of the worksheet group, as all the eight students in the chatbot group who scored five or lower points in the pretest increased their scores to seven or higher points in the posttest.

To examine these between-group differences, we conducted a two-by-two mixed-design ANOVA with *Group* (chatbot and worksheet) as a between-group variable and *Time* (pretest and posttest) as a within-group variable. The assumptions of sphericity and homogeneity were found to be met by a Greenhouse-Geisser and a Levene's test (*p*s > .05). We found a main effect of *Time*, F (1, 110) = 183.23, $p < .001$, $\eta_p^2 = .630$ (See Table 5), which indicates that there was a significant increase of the total mean score between the pretest ($M = 14.3$) and posttest ($M = 20.7$). A post hoc analysis of paired t-tests found that both groups had significant mean increases ($p < .001$).

Table 5

Mixed ANOVA of DCT Scores

Factor	Sum of Squares	*df*	Mean Square	*F*	*p*	η_p^2
Time	2224.2	1	2224.2	183.23	<.001*	.630
Time x Group	38.9	1	38.9	3.28	.073†	.029
Residual	1306.8	110	1306.8			

More importantly, there was a marginally significant interaction between *Time* and *Group*, $F(1, 110) = 3.28$, $p = .073$, $\eta_p^2 = .029$, indicating that the mean increase of the chatbot group (= 7.1; from 13.9 to 21.0) was more substantial than that of the worksheet group (= 5.5; from 14.9 to 20.4). However, post hoc analyses of independent t-test reported that there was no significant mean difference between the two groups in the posttest, which is probably due to the chatbot group's lower mean score in the pretest.

Self-directed Learning

The students' self-directedness in learning English was measured by nine items on the 5-point Likert scale (1: totally disagree ~ 5: totally agree). Table 6 summarizes the means and standard deviations of the two groups in the pre- and posttests.

Table 6

Means and Standard Deviations of Self-directed Learning in the Pretest and Posttest

Group	Pretest		Posttest	
	Mean	*SD*	Mean	*SD*
Chatbot ($N = 62$)	29.7	5.69	32.0	6.26
Worksheet ($N = 50$)	30.0	6.79	31.7	5.81
Sum	29.8	6.18	31.9	6.04

In the pretest, the chatbot group had a lower mean score in self-directed learning than the worksheet group. After the instruction, the mean scores of both groups significantly increased, and the mean increase of the chatbot group was greater than that of the worksheet group. In the posttest, the chatbot group showed a higher mean score than the worksheet group. These results show that there was an interaction between Time and Group, but a mixed ANOVA found that the interaction is not statistically significant, $F(1, 110) = 0.791$, $p = .376$, $\eta_p^2 = .007$.

The two groups' mean increases were further examined item by item as each item represented one of the nine themes in three categories. Both groups showed mean increases in every theme, but statistically significant increases were observed in different themes for each group (see Table 7).

Table 7

By-item Mean Comparisons between the Pretest and Posttest

Category	Theme	Chatbot (N = 62)				Worksheet (N = 50)			
		Pre	**Post**	***t***	***p***	**Pre**	**Post**	***t***	***p***
Meta-cognition	Planning	3.26	3.61	2.84	.006*	3.54	3.56	0.17	.868
	Monitoring	3.87	4.06	1.69	.096	3.80	4.02	1.47	.147
	Regulating	2.77	3.15	3.07	.004*	2.86	3.02	1.48	.146
Cognition	Rehearsal	3.94	4.05	1.04	.300	3.78	3.92	1.19	.241
	Elaboration	3.24	3.45	1.63	.107	3.02	3.44	3.92	<.001*
	Organization	3.13	3.55	3.30	.002*	3.14	3.42	1.85	.070†
Behavior	Time environment	2.40	2.79	3.06	.003*	2.66	2.76	0.71	.481
	Effort increase	3.61	3.75	1.59	.118	3.64	3.82	2.14	.038*
	Help seeking	3.56	3.65	0.53	.596	3.60	3.70	0.80	.429
Total		29.68	32.03	4.10	<.001*	30.04	31.66	2.78	.008*

In the category of metacognition, only the chatbot group showed significant mean increases in the themes of *planning* (p = .006) and *regulating* (p = .004). In the category of cognition, the chatbot and worksheet groups showed a significant (p = .002) or a marginally significant increase (p = .070) respectively in *organization*, whereas the worksheet group showed a significant mean increase in *elaboration* (p < .001). In the category of behavior, the chatbot and worksheet groups showed a significant increase in *time environment* (p = .003) and *effort increase* (p = .038) respectively. In sum, the two groups showed significant increases in self-directed learning of English, but they showed different gains for different themes of self-directedness.

Discussion

The first research question asks whether the chatbot activities were evaluated with higher scores than the worksheet activities. To answer the question, we analyzed the results of the activity-evaluation questionnaire in the posttest and found that the total mean score of the chatbot group (= 32.24) was greater than that of the worksheet group (= 31.24). Although the difference between the two groups' total mean scores was not statistically significant, we found that there was a significant between-group difference in the theme of *entertainment*. This finding supports the previous claim that chatbot activities can be more fun and interesting than traditional activities (Fryer & Carpenter, 2006; Jia & Chen, 2008).

We also examined the students' responses to the open-ended questions and found two features that made the chatbot activities fun. One feature was newness. Students appeared to evaluate the chatbot activities as a novel alternative to old-fashioned approaches such as rote memorization and listening to lectures. For example, a student answered that the chatbot activities were easier and more fun than *class*. Moreover, we observed that the chatbots in the present study went beyond novelty and led to enduring effects (compare Fryer et al., 2017): the students' participation and gain did not decrease between the first and second sessions. This outcome is probably because the second chatbot introduced a new situation and new expressions. Therefore, it is suggested that the use of chatbots should promote new learning opportunities to serve as novel alternatives to traditional English learning activities.

The other feature was the conversational process. Many students enjoyed talking with the chatbots. They liked the way that the chatbots immediately responded to their input in a contextually appropriate manner, which is one of the major advantages of chatbots (Fryer & Carpenter, 2006; Petrović & Jovanović, 2021;

Yang et al., 2022). It was also highly appreciated that they could try various conversation plots in a single context such as different lost items at a lost-and-found center. In addition, they were surprised to discover the powerful effects of conversation on implicit learning. For example, one student answered, 'I was happy to find that I naturally learned expressions as I had a conversation with the chatbots.' These responses indicate that chatbots for English learning should be designed to provide contextually appropriate responses, embrace various conversation plots, and facilitate vocabulary learning.

The second research question relates to the effects of self-directed activities on English conversation skills. Between the pre- and posttests of the DCT, the chatbot group showed a greater mean increase than the worksheet group, and the between-group difference of the mean increases was marginally significant. Therefore, it may be presumable that the chatbot activities are more effective in improving conversation skills than the worksheet activities (Muhammad et al., 2020).

To investigate what features of the chatbot activities led to the improvements in conversation skills, we first examined the students' answers to the activity-evaluation questionnaire. Many students answered that talking with chatbots was very similar to talking with real people. In other words, the chatbots offered authentic conversation experiences (Jia, 2009). This feature appeared to increase students' motivation and memory. For example, one student answered, 'It was like talking to a real person, so it was interesting and memorable.' The authenticity of the chatbots also facilitated meaningful learning (Ausubel, 1963). The students did not mechanically follow the sample dialogues; they considered multiple components of the conversation, such as contexts, roles, and previous turns, to manage the conversation flows. For example, one student answered, 'As I kept thinking before replying, I felt like my English skills improved a little.' In sum, the authenticity of the chatbots fostered the students' motivation, memory, and meaningful learning and led to improvements in conversation skills (Coniam, 2008; Okonkwo & Ade-Ibijola, 2021).

We also examined the dialog records between the chatbots and students to identify meaningful experiences related to the improvements in conversation skills. Interestingly, there were many cases of negotiation of meaning, a process that speakers go through to address a breakdown in communication and reach a clear understanding of each other. During negotiation of meaning, speakers should understand what the conversation is about and how to make the conversation unproblematic, and this experience can improve their conversation skills (Varonis & Gass, 1985; Yeh & Lai, 2019).

Negotiation of meaning is usually composed of the following four moves: (a) trigger, for example, a student's incorrect speech, (b) indicator, for example, a chatbot's sign of trouble, (c) response, for example, a student's corrected speech, and d) continuation, for example, a chatbot's acceptance and follow-up speech (Varonis & Gass, 1985). For example, see (1):

(1) Negotiation of meaning in a dialogue at the lost-and-found center

Turn	Interlocutor	Speech	
1	Student	*Hi. May I help you?*	
2	Chatbot	*Yes. I am looking for my scarf.*	
3	Student	*What does your umbrella look like?*	
		What else? Can you tell me more about it?	[Trigger]
4	Chatbot	*One more time?*	[Indicator]
5	Student	*What does your scarf look loke?*	[Response]
6	Chatbot	*Its color is grey.*	[Continuation]

This dialog record was collected in the first lesson, so the student was less familiar with the use of chatbots. In turn 3, the student requested the description of an 'umbrella' even though the chatbot was looking for a 'scarf', as shown in turn 2. In addition, the student performed two functions in a single turn, that is, request

of description (*What does your umbrella look like?*) and request of additional information (*What else? Can you tell me more about it?*), even though the two functions should have been performed in two separate turns. This problem happened probably because the student mechanically copied and pasted the expressions in the sample script. The problem triggered negotiation of meaning: in the following turn 4, the chatbot failed to move on to the next step and indicated its trouble. In turn 5, the student corrected his previous utterance. Even though the corrected response had a misspelling of *loke* instead of *like*, the chatbot accepted it and continued the conversation. Between turn 3 and turn 5, there were notable improvements in the student's conversation skills such as consideration of the opponent's situation and management of communicative functions.

We also found many cases of negotiation of meaning in the dialogues at the school nurse's office. For example, excerpt (2) showed how the student learned to combine two relevant communicative functions in a single turn.

(2) Negotiation of meaning in a dialogue at the school nurse's office

Turn	Interlocutor	Speech	
1	Student	*What's wrong?*	
2	Chatbot	*I have a sore throat. I have a fever, too*	
3	Student	*That's too bad.*	[Trigger]
4	Chatbot	*Sorry, can you say that again?*	[Indicator]
5	Student	*That's too bad. Do you have cough?*	[Response]
6	Chatbot	*Yes.*	[Continuation]

In (2), the student played the role of a school nurse, and the chatbot played the role of a sick student. In turn 2, the chatbot reported a health problem, and in turn 3, the student performed the function of consolation in the response. Given that the student was playing the role of a school nurse and needed to diagnose the patient's symptoms, the consolation alone appeared to be inappropriate. In the following turn 4, the chatbot indicated that the student's consolation speech triggered a problem. This feedback was possible because the chatbot was designed to expect the student to ask a diagnostic question (for example, *Do you cough?*). After the chatbot's indicator in turn 4, the student appeared to understand what the problem was and successfully provided a corrected response in turn 5 by combining two communicative functions, a consolation (*That's too bad*) and a diagnostic question (*Do you have cough?*). These observations could potentially indicate that chatbots can enhance learners' conversational skills via the negotiation of meaning (Qinghua & Satar, 2020; Yang et al., 2022).

The last research question asks whether the chatbot activities improved EFL learners' self-directedness in learning English more than the worksheet activities. Although the chatbot group showed a greater mean increase in the self-directedness questionnaire than the worksheet group (2.3 versus 1.7), the difference was not significant. This indicates that the chatbot activities might be slightly more effective in fostering self-directed learning than the worksheet activities, but there will be much room for improvement. For example, many students in the chatbot group suggested that the chatbots should be upgraded to understand more various expressions. This suggestion may show that they wanted to try more expressions with the chatbot in a self-directed manner. It was also suggested that other types of chatbots should be developed for those who are slow to type in English. One possible technical improvement is to provide a small number of example speeches for each turn and ask a student to choose the most appropriate one. If the student chooses the correct one, the chatbot will continue the conversation; if not, the chatbot will indicate the problem and show the list of example speeches again. Such improvements in the development of chatbots may increase the positive effects of the chatbot activity on self-directed learning of English.

It is also noteworthy that both the chatbot and worksheet groups had significant mean increases in the

questionnaire about self-directedness. In other words, both types of self-directed activities in the present study were effective in improving the students' self-directed learning of English. However, we have observed notable group differences in the item-based thematic analysis of self-directedness. The questionnaire had nine themes, and only the chatbot group showed significant increases in the three themes of *planning, regulating,* and *time environment.* This result appeared to be related to the thematic statements that the students evaluated. The statement about *planning* was 'I try to imagine a relevant situation to figure out new contents,' and this may relate to how the students in the chatbot group situated themselves in the dialogue context. The statement about *regulating* was 'I change the way I study to fit into the teaching style of the English teacher,' which appeared to resonate with the way the students learned the usage of chatbots and practiced English dialogues in novel ways. The statement about *time environment* was 'I make a study schedule to study English efficiently,' and this may relate to the way they managed the given activity time to practice both roles of each context with two conversational agents.

On the other hand, only the worksheet group showed significant increases in *elaboration* and *effort increase.* The statement about *elaboration* was 'when I study English, I make associations between similar learning contents,' and this may relate to the way they took into consideration multiple clues presented on each worksheet. The statement about *effort increase* was 'when I study English, I do my best,' and this reminded us that many of the students tried very hard to fill in all the blanks on the worksheet even though they were asked to decide their own pace and goal. In sum, self-directedness is a complex trait involving metacognitive, cognitive, and behavioral aspects (Knowles, 1975), and different features of self-directed activities contribute to different components of self-directed learning capacity.

Conclusion

In this study, we have compared the chatbot activities with the worksheet activities in terms of three aspects, namely activity evaluation, conversation skills, and self-directed learning. We have found that AI chatbots can be used to promote self-directed learning of English conversation skills. We also suggest several technological enhancements that may contribute to more effective use of chatbots in English language teaching and learning.

The use of AI chatbots is still in its initial phase, so there should be more empirical evidence for the wider use of AI chatbots in English language teaching. Since this study is limited to a specific context of L2 learning, we need to test the pedagogical effectiveness of chatbots in a variety of contexts including learners with various L1 backgrounds, different levels of access to smart devices, and varying degrees of familiarity with AI designs. In addition, there should be longitudinal research that goes beyond the novelty effect of technological gloss (Fryer et al., 2019) and investigates what aspects of chatbots motivate (or demotivate) language learners to use the chatbots for a substantial period.

References

Atwell, E. (1999). *The language machine: The impact of speech and language technologies on English language teaching.* British Council.

Ausubel, D. (1963). *The psychology of meaningful verbal learning.* Grune & Stratton.

Blum-Kulka, S., House, J., & Kasper, G. (1989). *Cross-cultural pragmatics: Requests and apologies*. Ablex.

Coniam, D. (2008). Evaluating the language resources of chatbots for their potential in English as a second language. *ReCALL, 20*(1), 98–116. https://doi.org/10.1017/s0958344008000815

DeKeyser, R. (2020). Skill acquisition theory. In B. VanPatten, G. Keating & S. Wulff (Eds.), *Theories in second language acquisition: An introduction* (pp. 83–104). Routledge.

Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language Learning & Technology, 24*(1), 16–26. https://doi.org/10125/44705

Fellbaum, C. (1998). *WordNet: An electronic lexical database*. MIT Press.

Fryer, L., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of chatbot and human task partners. *Computers in Human Behavior, 75*, 461–468. https://doi.org/10.1016/j.chb.2017.05.045

Fryer, L., & Carpenter, R. (2006). Bots as language learning tools. *Language Learning & Technology, 10*(3), 8–14. http://llt.msu.edu/vol10num3/emerging/

Fryer, L., Coniam, D., Carpenter, R., & Lăpușneanu, D. (2020). Bots for language learning now: Current and future directions. *Language Learning & Technology, 24*(2), 8–22. http://hdl.handle.net/10125/44719

Fryer, L., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest and competence. *Computers in Human Behavior, 93*, 279–289. https://doi.org/10.1016/j.chb.2018.12.023

Holec, H. (1996). Self-directed learning: An alternative form of training. *Language Teaching*, *29*(2), 89–93. https://doi.org/10.1017/s0261444800008387

Höhn, S. (2019). *Artificial companion for second language conversation*. Springer.

Hsu, M., Chen, P., & Yu, C. (2023). Proposing a task-oriented chatbot system for EFL learners speaking practice. *Interactive Learning Environments, 31*(7), 4297–4308. https://doi.org/10.1080/10494820.2021.1960864

Huang, W., Hew, K., & Fryer, L. (2021). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning, 38*(1), 237–257. https://doi.org/10.1111/jcal.12610

Jeon, J. (2022). Exploring a self-directed interactive app for informal EFL learning: a self-determination theory perspective. *Education and Information Technology, 27*, 5767–5787. https://doi.org/10.1007/s10639-021-10839-y

Jeon, J. (2023). Chatbot-assisted dynamic assessment (CA-DA) for L2 vocabulary learning and diagnosis. *Computer Assisted Language Learning*, *36*(7), 1338–1364. https://doi.org/10.1080/09588221.2021.1987272

Jia, J. (2009). CSIEC: A computer assisted English learning chatbot based on textual knowledge and reasoning. *Knowledge-Based Systems, 22*(4), 249–255. https://doi.org/10.1016/j.knosys.2008.09.001

Jia, J., & Chen, W. (2008). Motivate the learners to practice English through playing with chatbot CSIEC. In *International conference on technologies for e-learning and digital entertainment* (pp. 180–191). Springer.

Jia, J., & Ruan, M. (2008). Use chatbot CSIEC to facilitate the individual learning in English instruction: A case study. In *International conference on intelligent tutoring systems* (pp. 706–708). Springer.

Kim, N. (2016). Effects of voice chat on EFL learners' speaking ability according to proficiency levels. *Multimedia-Assisted Language Learning, 19*(4), 63–88. https://doi.org/10.15702/mall.2016.19.4.63

Kim, N., Cha, Y., & Kim, H. (2019). Future English learning: Chatbots and artificial intelligence. *Multimedia-Assisted Language Learning, 22*(3), 32–53. https://doi.org/10.15702/mall.2019.22.3.32

Kim, R. (2024). Effects of learner uptake following automatic corrective recast from artificial intelligence chatbots on the learning of English caused-motion construction. *Language Learning & Technology, 28*(2), 109–133. https://hdl.handle.net/10125/73574

Kim, Y., & Taguchi, N. (2015). Promoting task-based pragmatics instruction in EFL classroom contexts: The role of task complexity. *The Modern Language Journal*, *99*(4), 656–677. https://doi.org/10.1111/modl.12273

Knowles, M. S. (1975). *Self-directed learning: a guide for learners and teachers*. Association Press.

Lee, J., Yang, H., Shin, D., & Kim, H. (2020). Chatbots. *ELT Journal, 74*(3), 338–344. https://doi.org/10.1093/elt/ccaa035

Lin, C., & Mubarok, H. (2021). Learning analytics for investigating the mind map-guided AI chatbot approach in an EFL flipped speaking classroom. *Educational Technology & Society, 24*(4), 16–35.

Muhammad, A., Susanto, D., Alimudin, A., Adila, F., Assidiqi, M., & Nabhan, S. (2020) Developing English conversation chatbot using Dialogflow. In *2020 International Electronics Symposium* (pp. 468–475). IEEE.

Okonkwo, C., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, *2*, Article 100033. https://doi.org/10.1016/j.caeai.2021.100033

Pap, E. (2021). *Artificial intelligence: Theory and applications*. Springer.

Petrović, J., & Jovanović, M. (2021). The role of chatbots in foreign language learning: The present situation and the future outlook. In E. Pep (Ed.), *Artificial intelligence: Theory and applications* (pp. 313–330). Springer.

Pintrich, P., & de Groot, E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33–40. https://doi.org/10.1037//0022-0663.82.1.33

Qinghua, Y., & Satar, M. (2020). English as a foreign language learner interaction with chatbots: Negotiation for meaning. *International Online Journal of Education and Teaching, 7*(2), 390–410. http://iojet.org/index.php/IOJET/article/view/707

Ruan, S., Willis, A., Xu, Q., Davis, G. M., Jiang, L., Brunskill, E., & Landay, J. A. (2019). Bookbuddy: Turning digital materials into interactive foreign language lessons through a voice chatbot. In *Proceedings of the sixth (2019) ACM conference on learning@ scale* (pp. 1–4). https://doi.org/10.1145/3330430.3333643

Ryu, J. (2017). The effects of digital textbooks on college EFL learners' self-regulated learning. *Multimedia-Assisted Language Learning, 20*(4), 99–126. https://doi.org/10.15702/mall.2017.20.4.99

Shawar, B. (2017). Integrating CALL systems with chatbots as conversational partners. *Computación y Sistemas, 21*(4), 615–626. https://doi.org/10.13053/cys-21-4-2868

Shawar, B., & Atwell, E. (2007). Fostering language learner autonomy through adaptive conversation tutors. In *Proceedings of the fourth Corpus Linguistics conference* (Vol. 3, pp. 186–193).

Sung, M. (2021). Development of negotiation chatbots for primary English learners. *Primary English Education, 27*(1), 55–78. https://doi.org/10.25231/pee.2021.27.1.55

Sung, M., & Kang, S. (2024). Developing AI chatbots for pragmatic instruction of Korean secondary L2 English learners. *Korean Journal of English Language and Linguistics, 24*, 441–459. https://doi.org/10.15738/kjell.24..202405.441

Varonis, E., & Gass, S. (1985). Non-native/non-native conversations: A model for negotiation of meaning. *Applied Linguistics, 6*(1), 71–90. https://doi.org/10.1093/applin/6.1.71

Wang, Y., Petrina, S., & Feng, F. (2017). VILLAGE—Virtual immersive language learning and gaming environment: Immersion and presence. *British Journal of Educational Technology*, *48*(2), 431–450. https://doi.org/10.1111/bjet.12388

Yang, H., Kim, H., Lee, J., & Shin, D. (2022). Implementation of AI chatbot as an English conversation partner in EFL speaking classes. *ReCALL, 34*(3), 327–343. https://doi.org/10.1017/s0958344022000039

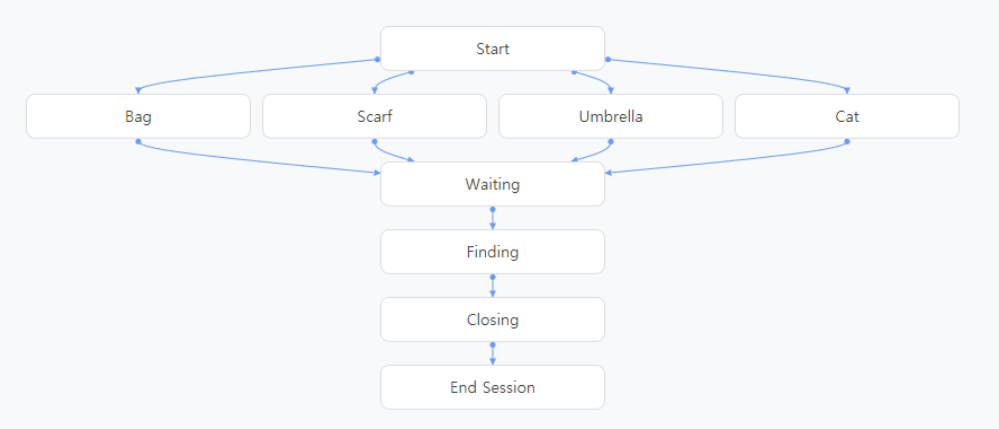
Yang, H., Kim, H., Shin, D., & Lee, J. (2019). A study on adopting AI-based chatbot in elementary English-speaking classes. *Multimedia-Assisted Language Learning, 22*(4), 184–205. https://doi.org/10.15702/mall.2019.22.4.184

Yeh, H-D., & Lai, W-Y. (2019). Speaking progress and meaning negotiation processes in synchronous online tutoring. *System, 81*, 179–191. https://doi.org/10.1016/j.system.2019.01.001

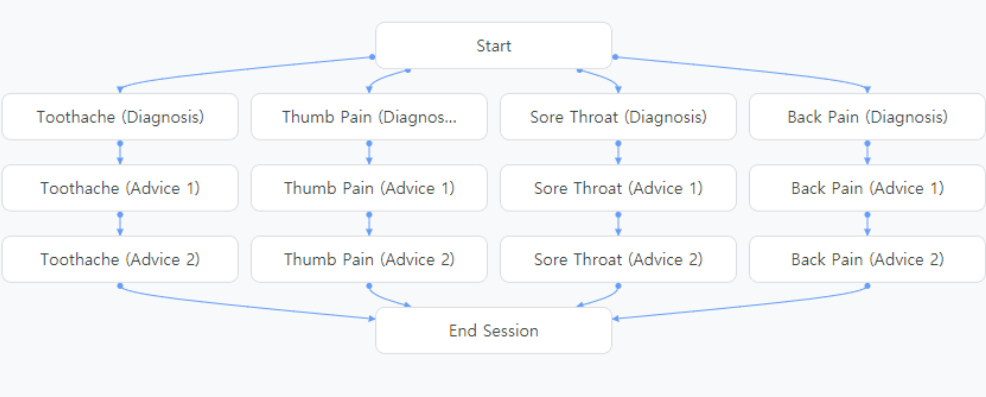
Yoon, J. M., Lee, H. G., Song, H. H., Yeom, M. S., Jin, S. I., Jang, S. W., Son, J. S., & Sundeen, G. (2018). *Middle school English 2*. Dong-A Publishing Co.

Appendix A. Structures of Chatbots for Two Dialogue Situations: Lost-and-found and Nurse's Office

1. Lost-and-found



2. Nurse's Office



Appendix B. Questionnaire Statements on Self-directed Learning and Activity Evaluation (Translated from Korean)

1. Statements on Self-directed Learning

Domain	Theme	Statement
Meta-cognition	planning	I try to imagine a relevant situation to figure out new contents.
	monitoring	When I get confused about contents, I go back and read or listen again.
	regulating	I change the way I study to fit into the teaching style of the English teacher.
Cognition	rehearsal	I mark the expressions emphasized by the English teacher.
	elaboration	When I study English, I make associations between similar learning contents.
	organization	When I study English, I go through the materials and try to find the most important contents.
Behavior	time environment	I make a study schedule to study English efficiently.
	effort increase	When I study English, I do my best.
	help seeking	When I have difficulty studying English, I look for related resources or online contents.

2. Statements on Activity Evaluation

Domain	Theme	Statement
User experience	entertainment	The chatbot/worksheet activity was fun.
	convenience	The activity was easy and convenient.
Learner Experience	authenticity	The activity was similar to a real English conversation.
	improvement	The activity helped me develop my English conversation skills.
	motivation	The activity made me more interested in speaking English.
	participation	The activity made me participate more actively in English classes.
Willingness	future use	I want to do the activity in the future.
	recommendation	I recommend the activity to my friends.

About the Authors

Sooyeon Kang received her doctoral degree in English Language Education at Seoul National University. She has been teaching English as a foreign language at a public middle school in Korea. Her research focuses on TBLT, pragmatics, corpus linguistics, AI chatbots, and instructed SLA.

E-mail: englishsk329@naver.com

ORCiD: https://orcid.org/0000-0002-9701-4155

Min-Chang Sung is an Associate Professor of English Education at Gyeongin National University of Education in South Korea. He has published on AI chatbots, corpus linguistics, and instructed SLA. He is the corresponding author.

E-mail: mcsung@ginue.ac.kr

ORCiD: https://orcid.org/0000-0001-9561-6374