

## The Effect of Using an Intelligent, Multimodal Chatbot for Korean Language Education

Sujung Nam

### Abstract

The rapid advancements in artificial intelligence and chatbot technologies are reshaping many disciplines, including language education. While many chatbots have been developed for teaching English, few have been used to teach Korean. In this study, an intelligent chatbot for Korean language education named Sayromi was developed and used by learners between 2022 and 2023 across three phases. In Phase 1, the researcher learned that the first version of the chatbot, Sayromi 1.0, was helpful for improving Korean language skills, but identified several ways to improve it, such as adjusting the voice speed and adding vocalization capability. In Phase 2, the researcher used an upgraded version, Sayromi 2.0, and confirmed the positive effect of the improvements. In Phase 3, the researcher added more content and used it in an intermediate-level Korean language course at a university in the United States, where the students were regularly given chatbot-based assignments. The result showed that Sayromi had multiple positive effects, such as enabling the learners to practice in a low-stress environment, helping them practice in multiple modes (reading, writing, speaking, and listening), and solidifying their understanding of Korean via real-life scenarios. It also shows that the chatbot's effectiveness increased between Phase 1 and 3.

**Key words:** Chatbots, Artificial intelligence, Dialogflow, KFL learners, ADDIE model

### 1 Introduction

As artificial intelligence technologies continue to advance at a rapid pace, chatbots (computer programs that simulate real conversations) are becoming increasingly proficient, intelligent, and useful for a wide range of tasks and activities, including those in the field of language education. In particular, major advancements in voice and language technologies are making chatbots increasingly useful for language learning. These intelligent chatbots can communicate in various languages and help teach all four modes of communication: reading, writing, speaking, and listening.

One of the languages that such an intelligent, multimodal chatbot can help teach is Korean. The Korean language is being taught at more than 1,500 schools around the world in more than 40 countries (Choi, 2023). In recent years, countries like India and Russia began to teach Korean in elementary and middle schools, while countries like Japan and France began to have Korean as one of the university entrance exam subjects. In 2020, more than 200,000 people took the TOPIK (Test of Proficiency in Korean) test, which is typically taken by those who want to study at a Korean university or work for a Korean company within or outside of South Korea. These statistics show that a growing number of people around the world are interested in learning Korean.

Intelligent chatbots can help these learners improve their Korean language skills in multiple ways: provide high-quality language practice opportunities to those who may not be able to readily access native Korean speakers; use authentic materials to make the educational resource more relevant and increase the learner's motivation; and help the learners improve their language skills beyond the classroom. Considering these potential benefits of chatbots for Korean language

education, this study aims to explore how an intelligent, multimodal chatbot can improve Korean language education for learners outside of the classroom.

## 2 Literature Review

Chatbots have evolved substantially over the past several decades. ELIZA, one of the first chatbots, was created between 1964 and 1967 by Joseph Weizenbaum at MIT. It generated the responses by detecting key words in the input and using decomposition and reassembly rules. Even though ELIZA did not have access to the information about the world and did not really understand the meanings of the provided inputs, many people found ELIZA to be quite compelling and human-like. They felt sure that they were talking to a human, not a chatbot (Weizenbaum, 1966).

As the chatbots became more sophisticated and advanced, researchers began to use them for language learning purposes and discovered many educational benefits. For example, when 211 students learned English by using a well-known chatbot named ALICE, 85% of the learners felt more comfortable talking to this chatbot than to a human being (Fryer & Carpenter, 2006). When these students used another well-known chatbot named Jabberwacky, 74% of them thought that this chatbot was fun and engaging (Fryer & Carpenter, 2006). The researchers highlight additional ways in which chatbots can be useful for language learning. They enable continued and repeated language practice, can support both written and oral language practice, and can provide immediate feedback to the learners. But they also point out that chatbots like ALICE and Jabberwacky are primarily designed for communicating with native English speakers, and can be difficult for less advanced English language learners. Hence, the researchers remark, for lower-level students, it can be worthwhile to learn with a chatbot that is designed for specific educational tasks (Fryer & Carpenter, 2006).

Previous studies have shown different kinds of benefits that voice chat and written chat practice provide in the context of language learning. After a group of Spanish learners engaged in written chat practice, they were able to produce complex output and use a wide range of pragmatic approaches, likely because the written form practice allowed the learners to take their time to improve their ability to form complex language structures (Sykes, 2013). Both voice chat and text chat practice have been shown to improve the learner's speaking skills, and the latter was shown to lower the learner's anxiety level (Satar & Ozdener, 2008). In fact, chat-based language practice and automatic speech recognition (ASR) are some of the most effective ways to apply technologies to foreign language learning. According to a group of researchers, the efficacy of most applications of technologies to foreign language learning is limited, except for chat (text or voice) and ASR (Golonka et al., 2014). As the authors point out, ASR can help improve the learners' language skills by providing immediate feedback about their pronunciation. Chat can help improve the learners' language skills by increasing their motivation to learn the language. These findings demonstrate the specific ways in which chatbots can be effective and useful for foreign language learning.

Other studies also highlight the positive impact that chatbots can have on the language learners' motivation and willingness to communicate. The willingness to communicate (WTC) is crucial for L2 language learning, because it can facilitate successful communication, which can then lead to successful L2 interaction and language acquisition (Dörnyei et al., 2015). When a group of university students in northern Taiwan used a chatbot to practice English for the TOEIC (The Test of English for International Communication) test, they felt less anxious about speaking

English and were able to improve their oral language skills (Hsu et al., 2021). With the help of this chatbot, the participants were able to practice English outside of the classroom without having to speak with a foreigner, of which they felt fearful. When a group of junior high school students in Taiwan used an intelligent personal assistant, Google Assistant, to learn English, the levels of their willingness to communicate, motivation, and communicative confidence increased (Tai & Chen, 2023).

On the one hand, the aforementioned studies and their findings suggest how chatbots can help improve the learners' language skills. However, most of the studies focused on English language education; few studies involved chatbots that are designed to teach the Korean language. Considering that many learners around the world are interested in learning Korean and that many countries like India, Russia, Japan, and France are adding Korean to their language education system, it is important to investigate this area further with a chatbot that can help learners practice Korean.

Several studies suggest how such a chatbot for Korean language education can be effectively designed, developed, and applied to language learning. For instance, it is crucial for chatbots to make appropriate adjustments based on the learner's competence level (Fryer et al., 2019). Otherwise, the learners may find it too challenging to talk to chatbots and can be frustrated. Second, even though chatbot technology can provide support and practice to language learners, the importance of pedagogy must not be forgotten; pedagogical goals should come first before technological means (Joyes et al., 2010). In fact, when technology-based language learning happens in a self-directed, informal manner outside of the classroom, the learners may not be able to sustain their motivation, and the usage of the technology may be low (García Botero et al., 2019). Therefore, the chatbot for Korean language education can become more effective if some degree of formal guidance is provided. One way to incorporate a chatbot into a Korean language course is to allow the students to use it to review and practice what they learned in the classroom (Hwang et al., 2022). Third, a chatbot can have a greater effect on the learners' Korean language skills if used for an extended period of time. If a technological language learning tool like a chatbot is used for only a short period of time, then the learners may not have enough time to practice the language with it and the result about the effect of the technology can be unclear (Hwang et al., 2022; Jeon, 2023; Refat et al., 2020). Fourth, a chatbot can make use of authentic materials that the learners can relate to, so that the chatbot can be more engaging for the learners (Hwang et al., 2022).

Additional studies provide further insights into how a chatbot can effectively help learners improve their Korean language skills. For example, a study involving a chatbot named Ellie for teaching English to elementary-level students in South Korea found that the learners can have difficulty when the chatbot's proficiency level exceeds that of the learners' (Yang et al., 2019). So, it can be worthwhile to simplify the chatbot's language and to adjust its length and speed to match the learners' proficiency level. And the chatbot can make use of a platform like Dialogflow from Google and be used in a classroom setting, so that its educational effect can be properly analyzed and understood (Sung, 2020). Assigning a recognizable character profile to the chatbot can make it more engaging and fun for the learners (Sung, 2020). It is important that the learners do not experience difficulties while using the chatbot, such as service discontinuation or payment issues, which can lead to a frustrating learning experience (Alm, A. & Nkomo, L. M., 2020).

Regarding the device type, while mobile-assisted language learning (MALL) has demonstrated some educational advantages, such as a sense of freedom and high degree of availability, significant skepticism towards MALL has been expressed (Calabrich, 2016). For example, the small screen size and limited battery life can affect the learning experience and

performance (Klímová, 2018). These drawbacks were not observed in web-based language learning that makes use of larger computing devices, which led to positive learner perception (Bashori et al., 2022). In this regard, the chatbot for Korean language education may not have to be only available on mobile devices, although being able to use it on mobile devices can be beneficial.

The research questions for this study are as follows:

1. How can a chatbot (developed and named Sayromi by the researcher) for learning intermediate-level Korean be made more user-friendly and intuitive by purposefully designing and adding the relevant features?
2. How effective are Sayromi's features at enhancing the quality of Korean language education, as measured by the learners' experience with and perception of the chatbot-based learning activities?
3. How can Sayromi improve the learners' motivation to learn Korean and their willingness-to-communicate?

### **3 Methodology**

#### **3.1 Instructional Design Paradigm**

Considering the high degree of complexity and ambiguity associated with developing an intelligent chatbot for learning Korean, a well-known instructional design model named ADDIE was followed. Adhering to this model enabled an iterative and systematic way to develop, assess, and enhance the chatbot for Korean language education. This model encompasses the following five phases, in the context of Korean language education:

1. Analysis: This stage involves analyzing what the target Korean language learners need to learn.
2. Design: This stage involves the designing of the Korean language materials to be learned.
3. Development: This stage involves the development of the Korean instructional materials and chatbot.
4. Implementation: This stage involves the teaching and learning of the developed materials using the chatbot.
5. Evaluation: This stage involves the evaluation of the effect of the instruction using the chatbot.

The ADDIE model is designed to be cyclic in the sense that after the Evaluation stage, the insights and learnings can then be used for the subsequent return to the Analysis stage. This study involved three cycles of the ADDIE model. In the first cycle (Phase 1), which took place between September 2022 and January 2023, the first version of the chatbot, named Sayromi 1.0, was designed, developed, implemented, and evaluated, after being used by intermediate-level Korean language students located in the United States. The learnings from this first cycle then led to the second cycle (Phase 2), which involved an upgraded version of the chatbot, named Sayromi 2.0, and took place between February and June 2023. Finally, the third cycle (Phase 3), which used Sayromi 2.1 that supports additional conversation materials, took place between June and August

2023 and involved a different group of intermediate-level Korean language students in the United States at the same institution as in Phase 1.

### **3.2 Phase 1 (September 2022 - January 2023)**

This phase involved the creation and usage of the first version of the chatbot, Sayromi 1.0. The primary goal of this phase was to understand the efficacy of an intelligent chatbot for Korean language learning and to assess how it can be improved further.

#### **3.2.1 Target Learner Proficiency Level**

In the Analysis stage, the target learners were determined to be intermediate-level Korean language students studying at a university in the United States. For the beginner-level Korean language students, chatbot-based language learning could be too difficult, especially considering that the chatbot is meant to enable the students to get more language practice outside of the classroom. These students can typically understand simple Korean alphabets and work with very simple sentence structures, but can have difficulty having conversations with a human or a chatbot. For the advanced-level Korean language students, it was not clear how much Sayromi can help improve their Korean language proficiency level further. Therefore, the target learners of Sayromi 1.0 were determined to be intermediate-level Korean language students.

#### **3.2.2 Materials**

Regarding which aspects of the Korean language can be taught by Sayromi 1.0, the official intermediate-level Korean language curriculum at the participants' academic institution in the United States was taken into consideration. This curriculum covered topics like season, weather, clothes, fashion, travel, school life, and shopping. By aligning the conversation topics supported by Sayromi 1.0 with the university's official Korean language curriculum, the learners could review and apply what they learned in class by practicing with the chatbot.

That said, the conversations supported by Sayromi 1.0 were not the same as those covered by the university's curriculum. Sayromi 1.0 used authentic materials that reflect the daily life scenarios in the area where the participants were located, whereas the official textbook used at the institution had many references to places and objects inside South Korea that the participants may not be familiar with. For example, instead of the names of places in South Korea, such as 잠실 야구장 (*Jamsil Baseball Stadium*) or 서울대입구역 (*Seoul National University Station*), the names of local places like 알라모아나 백화점 (*Ala Moana Shopping Center*) and 하나우마 베이 (*Hanauma Bay*), which the participants can more easily relate to, were used. This way, the chatbot could give more motivation to the learners by the virtue of the relevance of the educational materials to their daily lives.

Sayromi 1.0 supported 7 conversations spanning 4 lessons. It included 49 conversation turns in total; each conversation had 7 turns on average. The conversations covered the following topics: Season and Weather, Shopping, Travel, and School Life.

Figure 1 below is a sample text chat with Sayromi 1.0, whose topic is Shopping. In this chat conversation, the learner and Sayromi are friends who discuss their shopping preferences. They tell each other which clothes they like and need to buy, and arrange a visit to the Ala Moana shopping center, which is geographically close to the participants in the study.

Sayromi: There's a sale event at the Ala Moana shopping center this week. Do you have time this week by any chance?

Learner: Yes

Sayromi: I am going to buy a Hawaiian dress and sandals. Tell me which clothes you need to buy.

Learner: I need to buy a Hawaiian shirt and short pants.

Sayromi: It's hard to choose what to wear, because Hawaii and South Korea have different styles of clothes. What kind of clothes do you like?

Learner: I like thin clothes.

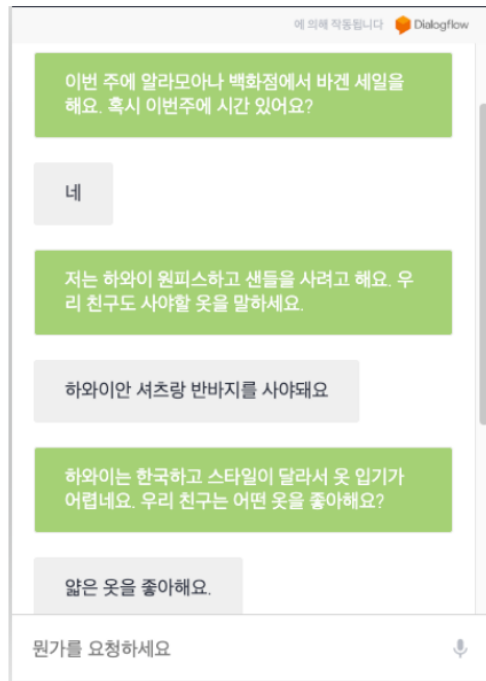


Figure 1. Text Mode (Sayromi 1.0) - Green bubbles: Sayromi, Gray bubbles: Learner

### 3.2.3 Chatbot Persona

In order to make the chatbot fun and interesting, a friendly persona (Figure 2) was designed and assigned to the chatbot. The character was intended to be a fellow university student who is a native Korean speaker. As for the name of the chatbot, “Sayromi” (새로미) was chosen, due to the relevant English and Korean meanings it conveys. “Sayromi” in Korean means “new”, and students can “say” words to Sayromi. The friendly persona was meant to reduce the learners’ anxiety level and to increase their motivation and willingness to communicate, which are crucial for improving language skills.



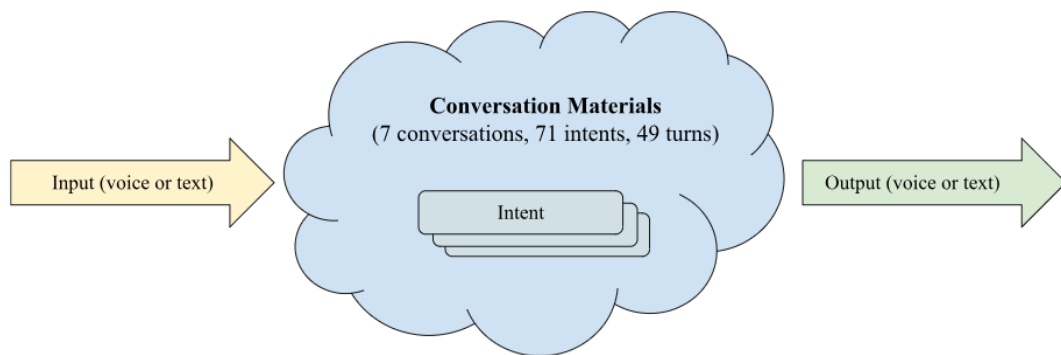
Figure 2. Sayromi Character

### 3.2.4 Technology

Dialogflow, a cloud platform for creating chatbots by Google, was used to develop Sayromi. The Korean conversation materials were entered into Dialogflow in the form of intents, each of which

involved one or more user input phrases and one response from Sayromi. Each intent represented the user’s intention for one conversation turn and could have follow-up intents for the subsequent turns in the same conversation.

To develop Sayromi, Dialogflow was used as the chatbot building platform for several reasons. First, this platform offered multiple ways to integrate and publish the chatbot, such as through web browser or phone number. Second, this platform provides many customization capabilities, including but not limited to voice type and voice speed. Third, Dialogflow provided a wealth of documentation, which clarified the chatbot development process and the available functionalities. Fourth, Dialogflow provides an easy-to-use web interface for creating intents (Figure 3).

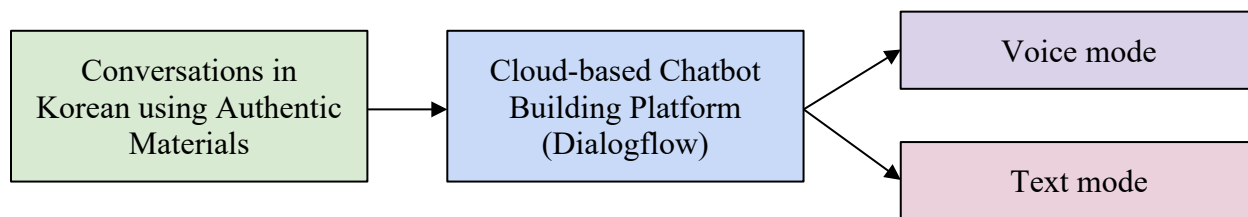


**Figure 3. Sayromi 1.0’s Usage of Intents**

In terms of the supported devices, Sayromi 1.0 was intended to be functional on both mobile phones and web browsers. Learners could dial a designated phone number to talk to Sayromi vocally, or they could access the web application and communicate with Sayromi via text. This way, the students could use Sayromi to learn Korean on large computer screens as well as on their smartphones, and did not have to be hindered by productivity issues or battery problems that can potentially happen on mobile devices with small screen size or low battery power (Klímová, 2017).

### 3.2.5 Multimodality (voice & text modes support)

In terms of the modes of communication that Sayromi can support and help the learners with, both written (reading and writing) and oral (speaking and listening) forms were chosen (Figure 4). The advantages of multimodality were demonstrated in existing literature, which found, for example, that the written format can help the students compose more complex language structures (Golonka et al., 2014) and that automatic speech recognition (ASR) can enable effective speaking practice (de Vries et al., 2015).



**Figure 4. Supporting Chatbot Multimodality**

In Phase 1, I designed Sayromi 1.0 for both the text-based and voice-based language functions by first entering the dialogical text into the chatbot building platform, Dialogflow. After entering the conversation data, I developed a web application where the learners can practice Korean for different lessons by 1) sending and receiving text messages, 2) using the speech-to-text feature to speak to Sayromi vocally. I paid close attention to how well Sayromi responds to different variants of user response in each conversation turn in each mode. Initially, during testing, it was discovered that Sayromi didn't have enough conversations data to be able to handle these different variants. After adding more conversation data, Sayromi could handle a wider range of inputs in both voice and text modes.

Figure 5 below shows how Sayromi 1.0 looked like on the web browser:

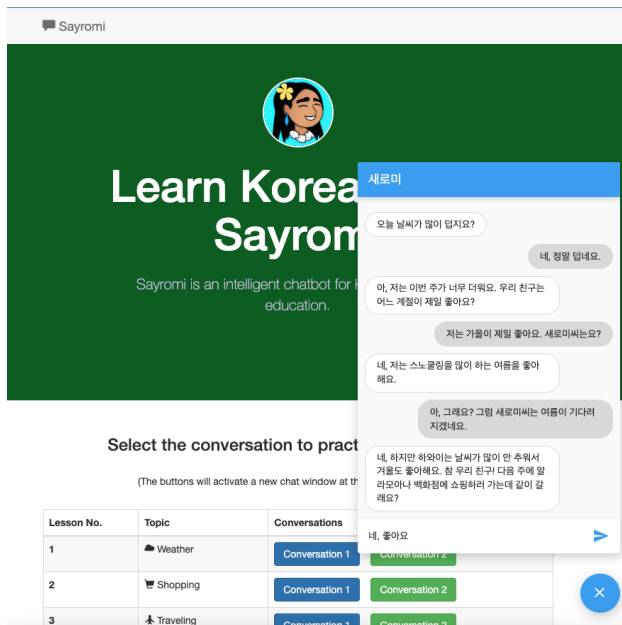


Figure 5. Sayromi 1.0 Web Interface

### 3.2.6 Participants

A total of 21 university students in the United States participated in Phase 1. The participants were intermediate-level Korean language learners, who took at least one Korean language course at the university. English was the native language for 95% of the participants.

### 3.2.7 Procedure

The procedure included the following three parts.

1. The participants completed the pre-study questionnaire, which mainly involved questions about demographic information.
2. The participants used Sayromi 1.0 to practice Korean, in voice and text modes.
3. The participants completed the post-study questionnaire, which involved questions about the effect of the chatbot on the learners' Korean language skills.

The first two parts of the procedure took about 45 minutes. The researcher provided assistance in person, such as preparatory voice practice before the participants used Sayromi 1.0 for Korean language practice, and helped resolve technical difficulties when needed (e.g., by suggesting that the participants reload the conversation if necessary). Since Phase 1 was designed and intended to be a pilot study for understanding the aspects of Sayromi that worked well and those that could be improved, each participant used Sayromi once in the semester.

The post-study questionnaire included 31 questions about the participants' perception of and experience with the chatbot. The questionnaire covered the voice and text modes separately and included both numeric and open-ended questions. The numeric questions were based on the 6-point Likert scale, where 1 meant Strongly Disagree and 6 meant Strongly Agree.

The numeric questions covered various aspects of Sayromi, such as the fluency of Sayromi's usage of the Korean language, how well it understood the user's speech and writing, how well it could improve the learner's speaking/listening/reading/writing skills, how it could increase their confidence in their Korean language skills, how fun it was to talk to Sayromi, its voice speed, and the relevance of the conversation topics on real life situations.

The ten open-ended questions covered areas like how it was to practice Korean with Sayromi, challenges of using Sayromi, what the learners liked about using Sayromi, how Sayromi can be improved, and comparison to in-person Korean language practice. These open-ended questions consisted of five questions about the voice mode and five questions about the text mode to assess the effect of Sayromi on the learners in each mode. To elaborate, even though the conversation materials were the same between the voice mode and the text mode, the learners' experience and perception about the chatbot's educational effect could be different for the two modes. For example, the quality of automatic speech recognition or voice speed could affect the voice mode but not the text mode. Therefore, the post-study questionnaire distinguished the questions about the voice mode from those about the text mode.

After the participants completed the questionnaire, the researcher analyzed the results to evaluate the efficacy of Sayromi 1.0 on the learners' Korean language skills and identified ways in which Sayromi 1.0 could be improved. The results, which are shown later, led to Phase 2 of the study involving an upgrade from Sayromi 1.0 and Sayromi 2.0.

### **3.3 Phase 2 (February 2023 - June 2023)**

This phase took place shortly after the conclusion of Phase 1. The main goal of this phase was to make various enhancements to Sayromi, such as adjusting the chatbot's voice speed and expanding the supported conversations, based on the feedback received in Phase 1. This upgraded version, Sayromi 2.0, was then used by a different group of Korean language learners, who completed a similar questionnaire as that used in Phase 1 about their experience.

#### **3.3.1 Target Learner Proficiency Level**

While Sayromi 1.0 solely focused on intermediate-level Korean language learners, Sayromi 2.0 covered a wider range of Korean language proficiency levels, including beginner level and advanced level, to gather diverse feedback about the learners' experience.

#### **3.3.2 Multimodality**

Like Sayromi 1.0, Sayromi 2.0 supported both voice and text modes. However, Sayromi 2.0 enhanced the multimodality by adding the text-to-speech feature shown below (Figure 6), which allowed the learners to hear Sayromi’s each message out loud by clicking the speaker button inside it. Moreover, based on the observations made in Phase 1, the voice speed was slowed down by 30%, so that the learners could understand Sayromi’s utterances more easily.

새로미와 대화해요  
(Talk to Sayromi)

Chat window

새로미와 연습할 주제 (topic) 를 선택하세요.  
(Select the conversation to practice with Sayromi)

Instructions:

- You can start each conversation by entering the trigger text (e.g., "L1C1") in the text input field in the bottom of the chat window, which says, "원가를 요청하세요", and pressing the return key.
- You can use the microphone button at the right side of the input field in the chat window to talk to Sayromi.
- You can use the speaker button next to the messages from Sayromi to make Sayromi talk to you.

KOR 201

Lesson	Topic	Conversations
1	☁ 계절과 날씨	<ul style="list-style-type: none"> <li>Conversation 1: Start by entering "L1C1" or "계절을 연습해요"</li> <li>Conversation 2: Start by entering "L1C2" or "날씨를 연습해요"</li> </ul>
2	🛒 옷과 유형	<ul style="list-style-type: none"> <li>Conversation 1: Start by entering "L2C1" or "쇼핑을 연습해요"</li> <li>Conversation 2: Start by entering "L2C2" or "쇼핑 놀이 해요"</li> </ul>
3	👤 여행	<ul style="list-style-type: none"> <li>Conversation 1: Start by entering "L3C1" or "여행을 연습해요"</li> <li>Conversation 2: Start by entering "L3C2" or "인상적인 여행을 연습해요"</li> </ul>
4	🏠 한국 생활 I	<ul style="list-style-type: none"> <li>Conversation 1: Start by entering "L4C1" or "학교 생활을 연습해요"</li> <li>Conversation 2: Start by entering "L4C2" or "우체국 놀이를 해요"</li> </ul>
5	🏠 한국 생활 II	<ul style="list-style-type: none"> <li>Conversation 1: Start by entering "L5C1" or "이사를 연습해요"</li> <li>Conversation 2: Start by entering "L5C2" or "기숙사 생활을 연습해요"</li> </ul>

Figure 6. Sayromi 2.0 Web Interface (showing the simplified UI and text-to-speech buttons)

### 3.3.3 Materials

Sayromi 2.0 supported a wider range of conversations and topics compared to Sayromi 1.0. In total, Sayromi 2.0 supported 25 conversations spanning 211 intents and 172 turns.

Table 1 below shows how Sayromi 2.0 compares to Sayromi 1.0.

	Sayromi 1.0	Sayromi 2.0
Target proficiency levels	Intermediate	Beginner, Intermediate
Number of conversations	7	25
Number of intents	71	211
Total number of turns	49	172

Table 1. Comparison between Sayromi 1.0 and Sayromi 2.0

Considering the increase in the number of supported conversations, Sayromi 2.0 web interface was reorganized and simplified, so that the learners could easily browse the topics and conversations without losing context of the main chat window. See Figure 6 above.

### **3.3.4 Participants**

A total of ten Korean language learners participated in Phase 2 by using Sayromi 2.0 to study Korean. The participants were randomly recruited in the United States. The participants were generally more diverse in Phase 2 than in Phase 1, with regard to their Korean language proficiency level, age, and the length of Korean language learning experience. In terms of the proficiency level, 34% considered themselves as below intermediate level, 22% considered themselves as intermediate level, and 33% considered themselves as above the intermediate level.

### **3.3.5 Procedure**

The procedure in Phase 2 was very similar to that in Phase 1, except for some minor updates to the post-study questionnaire. Like in Phase 1, the researcher analyzed the questionnaire submissions from the participants to understand the effect of Sayromi and rooms for improvement, and applied these learnings to Phase 3.

## **3.4 Phase 3 (June 2023 - August 2023)**

This phase took place shortly after the conclusion of Phase 2. The main goal of this phase was to use the upgraded version of Sayromi (Sayromi 2.1) in a real intermediate-level Korean language course at a university in the United States throughout one semester to meaningfully improve the quality of Korean language education and to understand the effect on the learners.

### **3.4.1 Target Learner Proficiency Level**

Similar to Phase 1, Phase 3 targeted intermediate-level Korean language learners enrolled in a Korean language course at a university in the United States (same institution as in Phase 1).

### **3.4.2 Materials**

In this Phase, Sayromi was upgraded from 2.0 to 2.1 by 1) adding 14 intermediate-level conversations covering topics like hobbies, holidays, lifestyle, culture, and everyday life, and 2) making some improvements on the web interface, so that the learners could navigate the conversations more easily. Also, Sayromi 2.1 allowed the learners to form their own sentences by using Korean expressions, such as “**있/있다가**”, which means “but then-” in English.

### **3.4.3 Participants**

A total of thirteen intermediate-level Korean language learners participated in Phase 3.

### **3.4.4 Procedure**

The procedure of Phase 3 was different from that of Phase 1 and 2, mainly due to the prolonged usage of Sayromi throughout the semester. In the beginning of the course, the participants took the pre-study questionnaire similar to that used in the earlier phases. Throughout the intermediate-level Korean language course that lasted for 7 weeks, the participants used Sayromi 2.1 twice a week to regularly review and practice what they learned in class. In addition, the participants were requested to inform the researcher, who was the instructor of the course, of their engagement with Sayromi for each lesson and conversation. At the end of the course, the participants took the post-study questionnaire about their experience with Sayromi; this post-study questionnaire was very similar to that used in the earlier phases, except for a couple questions that were added to understand the participants' general sentiment about using the chatbot for language practice.

## 4 Results

### 4.1 Participants' Responses to the Numeric Questions

The results from Phase 1 showed that the participants thought that Sayromi could help improve their Korean language skills in multiple ways. Regarding the question "Sayromi can help me improve my listening skills in Korean.", the average value of the participants' responses (i.e., agreement score) was 5.71 / 6.00 (95%). They also thought that Sayromi can help improve their reading skills (agreement score = 93%), writing skills (90%), and speaking skills (89%). They also thought that Sayromi was fluent in both voice (93%) and text (93%) modes. Moreover, they felt that Sayromi can help increase their confidence in their Korean language skills (92%), that the conversations were relevant to real life situations (86%), and that it was fun to practice Korean with Sayromi (82%).

The participants' answers to the numeric questions in Phase 1 also showed how Sayromi could be improved. For example, the participants thought that Sayromi's voice speed was fast (83%), and the agreement score was relatively lower for questions about Sayromi's ease of use (67%) and the range of conversation topics supported by Sayromi (81%).

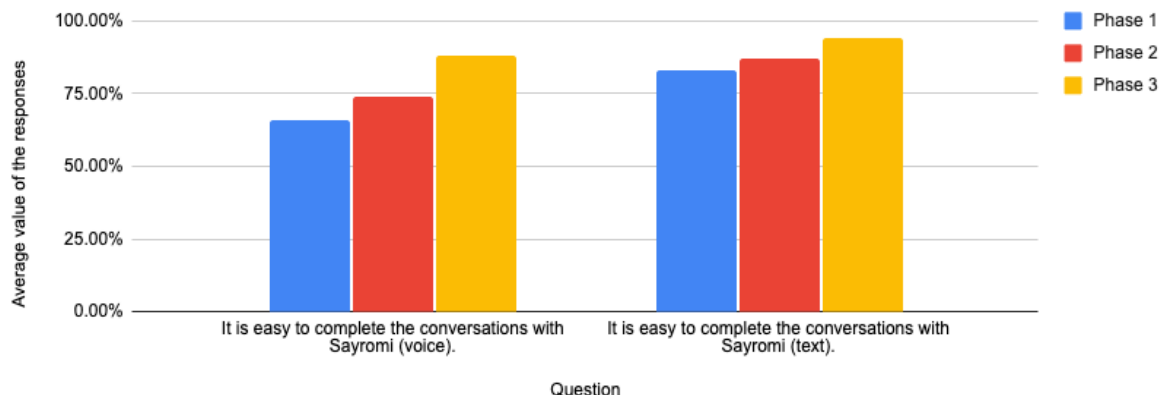
The Phase 2 participants' response to similar numeric questions showed that they thought Sayromi's voice speed was proper (87%) and that Sayromi was easy to use, especially in the text mode (85%). As a reminder, Sayromi's voice speed in Phase 1 was similar to that of a native Korean speaker, and the voice speed in Phase 2 was 30% slower than in Phase 1. The agreement score with regard to Sayromi's verbal fluency was 78%, which was lower than that in Phase 1, which was 93%. However, this particular result is not too surprising considering that the voice speed was intentionally slowed down by 30%.

In Phase 3, the average response value for each question about Sayromi was almost universally greater than that for the same question asked in Phase 1, except for two questions - one about Sayromi's verbal fluency (decreased from 93% to 91% with  $p$ -value = 0.651) and another about whether Sayromi can help improve the participant's listening skills (decreased from 95% to 92% with  $p$ -value = 0.492). That said, for these two questions, the  $p$ -value associated with the independent samples (two-tailed, heteroscedastic)  $t$ -test was greater than 0.05 and therefore not statistically significant. The average response value in Phase 3 was greater than that in Phase 1 to a statistically significant extent for questions about Sayromi's ease of use (Figure 7) in both text and voice modes, how fun it is to practice with Sayromi, and the relevance of Sayromi conversations in real-life situations; the  $p$ -value was smaller than 0.05 for these questions. There were two questions where the  $p$ -value was slightly greater than 0.05: one about how well Sayromi

can understand the participant's utterance (p-value = 0.052) and how Sayromi can help improve the participants' writing skills (p-value = 0.051). Table 2 below shows the statistically significant (or almost significant) entries for the responses to the numeric questions common between Phase 1 and Phase 3, for the voice mode and text mode.

Question	Phase 1 average	Phase 3 average	p-value (t-test)
<i><b>Voice mode</b></i>			
Sayromi can understand my speaking well.	4.05 / 6.00 (67%)	4.82 / 6.00 (80%)	0.052
It is easy to complete the conversations with Sayromi.	4.00 / 6.00 (67%)	5.27 / 6.00 (88%)	<b>0.002</b>
The more I practice Korean with Sayromi, the easier it becomes to practice with it.	4.62 / 6.00 (77%)	5.45 / 6.00 (91%)	<b>0.013</b>
It is fun to talk with Sayromi.	4.90 / 6.00 (82%)	5.64 / 6.00 (94%)	<b>0.038</b>
Sayromi conversations are relevant to real life situations.	5.19 / 6.00 (87%)	5.73 / 6.00 (95%)	<b>0.031</b>
<i><b>Text mode</b></i>			
Sayromi can help me improve my writing skills in Korean.	5.43 / 6.00 (90%)	5.82 / 6.00 (97%)	0.051
It is easy to complete the conversations with Sayromi.	5.00 / 6.00 (83%)	5.64 / 6.00 (94%)	<b>0.048</b>
The more I practice Korean with Sayromi, the easier it becomes to practice with it.	5.14 / 6.00 (86%)	5.73 / 6.00 (95%)	<b>0.010</b>
It is fun to talk with Sayromi.	4.90 / 6.00 (82%)	5.73 / 6.00 (95%)	<b>0.031</b>
Sayromi conversations are relevant to real life situations.	5.14 / 6.00 (86%)	5.73 / 6.00 (95%)	<b>0.014</b>

**Table 2. Comparison between Phase 1 & Phase 3 Responses (statistically significant/nearly significant entries)**



**Figure 7. Average Response Values for the Numeric Questions About Sayromi's Ease of Use (Phase 1-3)**

For eight other questions in Phase 1, the average response value increased in Phase 3, albeit not to a statistically significant degree. These questions dealt with the following aspects of Sayromi: whether Sayromi can improve the participant's speaking skills (increased from 89% to 92% with p-value = 0.491), whether it can improve their reading skills (increased from 93% to 97% with p-value = 0.324), the range of topics and conversations supported by Sayromi (increased from 83% to 89% with p-value = 0.297 for the text mode), whether Sayromi can help increase the participant's confidence in Korean (increased from 92% to 97% with p-value = 0.509 for the text mode), Sayromi's fluency in the text mode (increased from 93% to 94% with p-value = 0.762), and how well Sayromi can understand the participant's writing (increased from 81% to 88% with p-value = 0.257).

There were five numeric questions in Phase 3 that were not part of Phase 1: whether the participant would generally recommend that future Korean language courses use a chatbot like Sayromi again (average response value = 94%), whether the Sayromi based assignments were generally effective for practicing Korean (91%), whether Sayromi could understand the participant's utterances well (80%), whether the participant was able to understand Sayromi's utterances well (92%), and whether Sayromi can help improve the participant's pronunciation accuracy (91%).

#### **4.2 Participants' Responses to the Open-ended Questions**

In Phase 1, many participants reported that they had fun practicing Korean with Sayromi. Regarding how it was to practice with Sayromi in the voice mode, one participant wrote, "It was interesting I didn't expect it to be like it was, the AI seemed pretty competent". Another participant wrote, "It was fun. It felt like I'm having a real conversation with a Korean speaker." Regarding what the participant liked about practicing Korean with Sayromi, one participant wrote "I liked that it helps me practice my Korean speaking and listening skills. Its pronunciation is very clear and native.", and another wrote "It is automated, and it doesn't laugh at me when I struggle." Also, there were several answers about how responsive and natural Sayromi was. One participant wrote, "I liked that it responds immediately. It covers everything that we've learned; I could understand it most of the time." Another wrote, "I liked that it felt like the answers weren't always set and they were responding to the actual words I was saying. It felt like I was having a real conversation with a person."

Regarding what the participants found challenging in Phase 1, many participants reported that Sayromi spoke too fast. One participant wrote, “I found it hard to understand her because she spoke really fast so I was unable to fully follow along. Yet, I was able to get a general understanding of what she was saying through words or small phrases.” Also, some participants wrote that the voice mode was difficult, because Sayromi did not always understand what the participant said and gave a limited amount of time for them to respond. One wrote, “Thinking of answers and responding quickly was the hardest.”

In Phase 2, none of the answers to the open-ended questions complained about how fast Sayromi spoke. Some participants actually pointed out that they liked Sayromi’s voice speed. For example, one participant wrote, “I liked the speaking speed.” Similar to Phase 1, many participants wrote that it was fun and interesting to practice with Sayromi. For example, one participant wrote, “Very fun. I was able to understand it very well.” Other answers showed that the participants liked the Sayromi interface and functionalities. One participant wrote, “I like the simple platform of Sayromi” and another wrote, “It is a nice way to learn material from the textbook via an interactive setting. Its ability to listen to and transcribe speech is really good!”

In Phase 3, almost every participant reported that they had a positive experience practicing with Sayromi. For example, one participant wrote, “It was a good experience, I felt comfortable chatting with AI.” and another wrote, “It was a really nice way of seeing how certain grammar points could be relevant in real-life situations! Her voice is also at a good speed.” Other participants wrote that they liked the flexibility that Sayromi allows in the user input and the occasional hints that Sayromi provided. For example, one participant wrote the following about what the person liked about practicing with Sayromi: “The possible variations that are possible with Sayromi. I can change my answer slightly while practicing the grammar points and Sayromi would understand it. I also really like the change in tone when Sayromi's voice is playing. It made it sound like someone who is speaking and not much of a robotic voice. I also like how Sayromi gives a response back if there is something wrong with my answer. It helps me understand what I typed wrong or what part of my sentence was wrong. I really like how there is also enough time given to speak before the answer is submitted.”

Regarding how Sayromi-based practice compares to in-person practice, there were mixed opinions. Several participants thought that they preferred certain aspects of the Sayromi-based practice to in-person practice. For example, one participant wrote, “It's less nerve-wrecking and gives the students who learn at a slower pace to comprehend the listening, writing, etc. A good practice for the students who stutter like me.” Other participants wrote that Sayromi-based practice was natural like in-person practice. For example, one participant wrote, “It feels like I am texting a friend.” and another wrote, “Sayromi is pretty similar to in-person text practice so it's good!” Several answers to the open-ended questions explained how the participant thought Sayromi can be improved further. For example, one participant wrote, “Please make a wider variety of answers that a user can put in.”

## **5 Discussion**

### **5.1 Making Sayromi More User friendly and Intuitive**

The first research question dealt with how Sayromi can be made more user friendly and intuitive by purposefully designing and adding the relevant features. By applying the ADDIE model, the

researcher could iteratively and systematically identify the set of features that collectively made Sayromi more user friendly and intuitive across Phase 1, 2, and 3. In Phase 1, the researcher designed and added a feature to Sayromi 1.0 whereby the conversations used authentic materials to make the content more engaging and familiar for the learners (e.g., the names of local places). This way, Sayromi 1.0 provided an intuitive dialogical content that the learners could easily relate to. Also in Phase 1, the researcher added a feature where the learners could access the voice mode by dialing a designated phone number. Even though the voice speed of Sayromi 1.0 turned out to be too fast for the intermediate-level learners and required correction before Phase 2, this feature helped provide a user-friendly and intuitive voice mode experience in Phase 1.

After Phase 1, based on the learners' feedback received, the researcher made Sayromi more user-friendly and intuitive by designing and adding more features. In Phase 2, the researcher designed and added a text-to-speech feature that allowed the learners to easily and intuitively hear Sayromi's messages out loud in Korean. Also, based on the learners' feedback that Sayromi 1.0's voice speed was too fast, Sayromi's voice speed was slowed down by 30% in Sayromi 2.0. As a result, in Phase 2 and Phase 3, the participants no longer reported that Sayromi spoke too fast, which implied that Sayromi's voice mode became more user friendly for the target proficiency level. Also in Phase 2, in order to improve the user-friendliness of the text mode, the researcher streamlined the user interface and made it simpler, so that the learners could better focus on the conversation with Sayromi. In Phase 3, based on the feedback that Sayromi could support a wider range of user input, the researcher added a feature to Sayromi 2.1 that allowed the learners to easily form their own sentences using the expressions that they recently learned in class, such as “*있/았다가*”, which means “but then-” in English. By using the intent matching mechanism provided by Dialogflow appropriately, the researcher was able to support various user inputs that used such learned expressions in a customized manner.

In these ways, the researcher was able to purposefully design and add relevant features to Sayromi that collectively made it more user-friendly and intuitive for learning Korean, and the results validated this outcome. The results from Phase 3 showed that the learners liked Sayromi's simplicity and ease of use more than those in Phase 1. While the agreement scores (i.e., the average values of the numeric responses) for the question “It is easy to complete the conversations with Sayromi” were 67% for the voice mode and 83% for the text mode in Phase 1, they rose to 88% and 95% respectively, in Phase 3, to a statistically meaningful extent. The answers to the open-ended questions further corroborated this result. For example, a learner in Phase 2 wrote “I like the simple platform of Sayromi”, and a learner in Phase 3 wrote “It was a good experience, I felt comfortable chatting with AI.”

## **5.2 The Impact of Sayromi's Features on the Korean Language Learners**

The second research question dealt with the impact that Sayromi's features had on the Korean language learners, as measured by the learners' experience with and perception of the chatbot-based learning activities. The results from the study demonstrate the positive effect that Sayromi and its features (e.g., voice and text modes support and the usage of authentic materials) had on the learners, both at a general level and at the feature level. First, the results from Phase 3 show that the learners overall had a positive experience regularly practicing with Sayromi throughout the semester, in conjunction with in-person instruction. In the post-study questionnaire for Phase 3, the agreement score for the question, “I would recommend that future Korean language courses use a Korean language chatbot like Sayromi again.” was 94%, and that for the question “Sayromi

based assignments were effective for practicing Korean.” was 91%. These results demonstrate the positive effect that Sayromi generally had on the learners.

As for the effect that Sayromi’s individual features had on the learners, the results showed that the learners agreed that Sayromi was effective in both voice and text modes. The numeric responses indicated that Sayromi could help meaningfully improve the learners’ speaking skills (agreement score = 92%), listening skills (92%), reading skills (97%), writing skills (97%), pronunciation accuracy (91%), and confidence in Korean in both voice (92%) and text (97%) modes. These values show the positive effect that Sayromi had on the learners, in each mode. The usage of authentic materials also had a positive effect on the learners. Sayromi used authentic materials that the participants could relate to, in order to make the conversations more engaging and relevant. Not surprisingly, overall, the participants agreed that the Sayromi conversations were relevant to real-life situations; the agreement score was 95% in Phase 3 and 87% in Phase 1. The participants’ answers to the open-ended questions also show that they found Sayromi conversation practice engaging, which substantiates the positive effect of the relevance of Sayromi conversations on real-life scenarios. For example, one participant in Phase 3 wrote, “Sayromi was very easy to use. The dialogue with Sayromi was very engaging.” and another participant wrote, “it was not bad at all. Definitely a more engaging experience which was nice.”

### **5.3 The Effect of Sayromi on the Learners’ Motivation and Willingness-to-communicate**

The results showed that the learners found Sayromi to be a fun, engaging, and comfortable partner for practicing Korean. In Phase 3, the learners strongly agreed that it was fun to talk to Sayromi in both text and voice modes; the agreement score was 94% for the voice mode and 95% for the text mode. The learners’ responses to the open-ended questions confirmed this result as well. For example, one learner wrote, “My overall experience was very fun. It was a great source of review for the grammar points we learned in class.” Another learner wrote, “It felt like Sayromi was a friend; we didn't just talk about topics relating to the grammar but other things too, which is realistic. It also really helped me improve my listening, speaking, and writing skills. It was nice to have an extra tool to help practice Korean.” These responses show that Sayromi helped increase the learners’ motivation by making the practice more fun and engaging, and by utilizing authentic materials that were familiar to them.

Regarding the effect that Sayromi had on the learners’ willingness-to-communicate, several responses showed the positive impact that Sayromi had on the learners in this regard. For example, one learner in Phase 1 wrote, “It is automated and it doesn't laugh at me when I struggle.” regarding what they liked about practicing Korean with Sayromi. And a learner in Phase 3 wrote “It's less nerve-wrecking and gives the students who learn at a slower pace to comprehend the listening, writing, etc. A good practice for the students who stutter like me.” These responses demonstrate that practicing Korean with Sayromi could be less stressful than practicing with a human partner, because the learners don’t have to worry about embarrassing themselves in front of another person and could take their time to respond to Sayromi. In terms of the learners’ answers to the numeric questions, the learners agreed in Phase 3 that Sayromi can help increase their confidence in Korean; the agreement score was 92% for both the text and voice modes. Therefore, the results show that a chatbot like Sayromi can meaningfully improve the learners’ motivation and willingness-to-communicate by offering them a comfortable and low-stress way to practice Korean with A.I.

## 6 Conclusion

In this study, an intelligent, multi-modal chatbot named Sayromi was developed for Korean language learning and used to teach Korean to three cohorts of Korean language learners across three phases that took place between September 2022 and August 2023. After each phase, the researcher analyzed the post-study questionnaire submissions to understand the educational effect of Sayromi and the rooms for improvement. By making appropriate improvements on Sayromi, the researcher was able to make Sayromi more effective and useful for language learning.

The positive results from Phase 3 in particular, in which Sayromi 2.1 was regularly used in an intermediate-level Korean language course throughout the semester, show how an intelligent, multimodal chatbot like Sayromi can be successfully and effectively incorporated into an in-person Korean language course. In this phase, the participants' responses to the numeric and open-ended questions show that they found Sayromi to be an engaging, useful, relevant, and effective way to practice Korean across multiple modalities like reading, writing, speaking, and listening.

While this study investigated and discovered the effect of an intelligent, multimodal chatbot on Korean language learners, it was not without limitations. First, it involved three phases that happened sequentially one after another, and did not involve control and experimental groups. Second, Sayromi may not understand the learner's utterance if not pronounced clearly. When the learner makes mistakes, Sayromi does not give specific feedback or correction. Third, since Sayromi uses a finite number of intents to handle the user input, it cannot respond to all kinds of inputs. Fourth, in the voice mode, the learner had a limited duration to think about their answer before speaking. Fifth, even though Sayromi made use of artificial intelligence features like intent detection, voice recognition, and text-to-speech, it did not make use of generative artificial intelligence to respond to user input.

Future studies can explore how an intelligent chatbot like Sayromi can be improved further by addressing these limitations. For example, they can try to use generative artificial intelligence and large language models to support a wider range of conversations and user inputs. However, when using such new artificial intelligence technologies, it's important to ensure that the chatbot's responses are not excessively advanced or challenging in relation to the learner's proficiency level. Otherwise, the educational effect of such an intelligent chatbot on the target learners can be affected by the difficulty experienced by the target learners. In this regard, the language instructor has an important role to play in designing and developing these intelligent chatbots, so that the conversations with these chatbots align well with the lessons taught in the Korean language classroom, and that the learners can practice with these chatbots effectively without issues.

## References

- Alm, A. & Nkomo, L. M. (2020). Chatbot Experiences of Informal Language Learners: A Sentiment Analysis. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, 10(4), 51–65. <http://doi.org/10.4018/IJCALLT.2020100104>
- Bashori, M., van Hout, R., Strik, H., & Cucchiaroni, C. (2022). Web-based language learning and speaking anxiety. *Computer Assisted Language Learning*, 35(5–6), 1058–1089. <https://doi.org/10.1080/09588221.2020.1770293>
- Calabrich, S. L. (2016). Learners' Perceptions of the Use of Mobile Technology in a Task-Based Language Teaching Experience. *International Education Studies*, 9(12), 120-. <https://doi.org/10.5539/ies.v9n12p120>
- Choi, J.H. (2023, Jan. 2). A snapshot of Korean language now. *The Korean Herald*. <https://www.koreaherald.com/view.php?ud=20230102000432>
- Dörnyei, Z., Macintyre, P. D., & Henry, A. (2015). *Motivational dynamics in language learning*. Multilingual Matters.
- Fryer, L.K., & Carpenter, R. (2006). Bots as language learning tools. *Language Learning & Technology*, 10(3), 8–14.
- Fryer, L. K., 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>
- García Botero, G., Questier, F., & Zhu, C. (2019). Self-directed language learning in a mobile-assisted, out-of-class context: do students walk the talk? *Computer Assisted Language Learning*, 32(1–2), 71–97. <https://doi.org/10.1080/09588221.2018.1485707>
- Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: a review of technology types and their effectiveness. *Computer Assisted Language Learning*, 27(1), 70–105. <https://doi.org/10.1080/09588221.2012.700315>
- Hsu, M.-H., Chen, P.-S., & Yu, C.-S. (2021). Proposing a task-oriented chatbot system for EFL learners speaking practice. *Interactive Learning Environments*, 1–12. <https://doi.org/10.1080/10494820.2021.1960864>
- Hwang, W.-Y., Guo, B.-C., Hoang, A., Chang, C.-C., & Wu, N.-T. (2022). Facilitating authentic contextual EFL speaking and conversation with smart mechanisms and investigating its influence on learning achievements. *Computer Assisted Language Learning, ahead-of-print*(ahead-of-print), 1–27. <https://doi.org/10.1080/09588221.2022.2095406>
- 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>
- Joyes, G., Gray, L., & Hartnell-Young, E. (2010). Effective practice with e-portfolios: How can the UK experience inform implementation? *Australasian Journal of Educational Technology*, 26(1). <https://doi.org/10.14742/ajet.1099>
- Klímová, B. (2018). Mobile phones and/or smartphones and their apps for teaching English as a foreign language. *Education and Information Technologies*, 23(3), 1091–1099. <https://doi.org/10.1007/s10639-017-9655-5>
- Refat, N., Kassim, H., Rahman, M. A., & Razali, R. bin. (2020). Measuring student motivation on the use of a mobile assisted grammar learning tool. *PLOS ONE*, 15(8), e0236862. <https://doi.org/10.1371/journal.pone.0236862>

- Satar, H. M., & Özdener, N. (2008). The Effects of Synchronous CMC on Speaking Proficiency and Anxiety: Text Versus Voice Chat. *The Modern Language Journal*, 92(4), 595–613. <https://doi.org/10.1111/j.1540-4781.2008.00789.x>
- Sung, M.-C. (2020). Pre-Service Primary English Teachers' AI Chatbots. *Language Research*, 56(1), 97–115. <https://doi.org/10.30961/lr.2020.56.1.97>
- Sykes, J. M. (2013). Synchronous CMC and Pragmatic Development: Effects of Oral and Written Chat. *CALICO Journal*, 22(3), 399–431. <https://doi.org/10.1558/cj.v22i3.399431>
- Tai, T.-Y., & Chen, H. H.-J. (2023). The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments*, 1–18. <https://doi.org/10.1080/10494820.2020.1841801>
- Weizenbaum, J. (1966). ELIZA---a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45. <https://doi.org/10.1145/365153.365168>
- Yang, H.J., Kim, H.Y., Shin, D. & Lee J.H. (2019). A study on adopting AI-based chatbot in elementary English-speaking classes. *Multimedia-Assisted Language Learning*, 22(4), 184–205.

### Author's information

*Sujung Nam*

*University of Hawai'i at Manoa*

[nams@hawaii.edu](mailto:nams@hawaii.edu)