

Thi Ngoc Diep Nguyen

# **EXPLORING AI FREE-FORM CONVERSATIONS AS A TOOL FOR KOREAN LANGUAGE LEARNING**

Faculty of Information Technology and Communication Sciences

Master of Science Thesis

Examiner: Aparajita Chowdhury

Examiner: Dr. Aino Ahtinen

February 2025

# ABSTRACT

Thi Ngoc Diep Nguyen: Exploring AI Free-Form Conversations as a Tool for Korean Language Learning  
M.Sc. Thesis  
Tampere University  
Master's Degree Program in Human-Technology Interaction  
February 2025

---

With the rise of Korean Wave (Hallyu)'s global popularity in the past decades, more and more adults in Europe and particularly in Finland are interested in learning the Korean language. Among various independent learning methods, chatbots driven by Artificial Intelligence (AI) have gained attention recently for their personalization, flexibility and positive effects on users' language proficiency. AI has facilitated language learners in enhancing their pronunciation, grammar, vocabulary, and speaking fluency. However, most studies concerning this AI-powered learning method have focused on students who learn languages in a monitored environment, following a predetermined curriculum rather than fully open-ended content. Little research has been done to explore how adult students interact in free-form Korean conversations with AI chatbots.

This master's thesis aims to understand adult learners' experience using AI-enhanced tools in their Korean language pedagogy. Three research questions concerning the Korean language learners' expectation (1), needs for guidance during AI-driven conversations (2), and motivation to use AI applications in their Korean education journey (3) were formulated. To shed more light on this subject, a three-step study was conducted with a total of nine participants residing in Finland (N=5) and South Korea (N=4), who had little to no experience learning Korean with AI-enhanced tools prior to the study. First, an online questionnaire focusing on learners' Korean education background and their expectations on AI as a Korean language learning tool was conducted. Next, participants experienced an AI-driven application called Universal independently for seven days, during which they also utilized a journal to note down their feedback and opinions. In the last step, participants shared their opinions regarding their experience learning with AI in audio-recorded semi-structured interviews.

The research findings indicate that learners have low expectations regarding the effectiveness of conversational AI in supporting their Korean language studies, with minimal differences between those living in South Korea and Finland. Both groups expected certain benefits, such as enhancing vocabulary, grammar, pronunciation, fluency, and speaking confidence, as well as access to language-related information. However, their expectations of AI chatbot's usefulness, their comfortability interacting with AI, and their perceptions of AI's human-like characteristics varied. Issues encountered during interactions with AI-powered chatbots underscored the users' need for certain guidance in several areas. These included the chatbot's unclear name and age, unsuited conversation topics, interruptions in conversation flow, insufficient vocabulary assistance, technical glitches and bugs, along with inadequate and unsatisfying feedback. Learners' motivation and willingness to practice Korean with AI chatbots were influenced by factors such as interaction quality, conversation content, feedback quality and frequency, AI's anthropomorphism level, as well as the severity of technical issues.

The thesis results add to existing studies by providing valuable information on the experience and design of AI chatbots taking into consideration the nature of the Korean language. The study shows that certain aspects of AI chatbot's design such as conversation topics and immediate feedback are appreciated by Korean learners, similar to the experience in other languages. Nevertheless, elements such as chatbots' character and pronouns are unique to the experience of Korean language students. This thesis project also provides six design implications to support future planning and implementation of AI chatbots for Korean language pedagogy. These implications include affordances, feedback and corrections, discussion topics, conversation interaction, chatbot character design, and chatbots' trustworthiness.

**Keywords:** Korean Language Education, Artificial Intelligence, AI, Chatbots, AI Chatbots, AI Applications, Human-Centered Design, User Experience.

The originality of this thesis has been verified using the Turnitin Originality Check service.

# USE OF AI IN THESIS

I have utilized AI tools in my thesis:

- No
- Yes

I acknowledge that I am fully responsible for the entire content of my thesis, including the parts generated by AI, and accept accountability for any violations of ethical standards in publications.

# PREFACE

***"If opportunity doesn't knock, build a door."***

*- Milton Berle*

Completing this master's thesis has been a challenging yet rewarding journey, shaped by persistence, flexibility, and a commitment to learning and growth. A quote by Milton Berle, "If opportunity doesn't knock, build a door," has been an inspiration that guided me throughout this process.

My journey was not without difficulties. At the start, my original research proposal faced unexpected challenges in participant recruitment and academic requirements. These setbacks made it necessary to rethink my plans and try new approaches. This process taught me the importance of adapting to new circumstances and finding creative solutions when things don't go as planned. While it was humbling at times, it also reminded me that obstacles can be opportunities to think differently and grow.

This project would not have been realizable without the help of many individuals. I am deeply thankful to those who participated in the empirical phase of this thesis. Your time, insights, and willingness to share your experiences have been crucial in shaping the findings of this thesis. I also want to sincerely thank my supervisors, Aparajita Chowdhury and Dr. Aino Ahtinen, for their guidance and support. Their feedback and encouragement helped me stay focused and keep going, even when things felt uncertain.

I hope this thesis motivates others to face challenges with courage and to create their own opportunities in life. It is a reminder that even in tough situations, progress is possible with determination and effort.

Tampere, 24 February 2025

Thi Ngoc Diep Nguyen

# Contents

1 Introduction	1
1.1 Motivation and Research Context	1
1.2 Research Goal and Research Questions	4
1.3 Thesis Structure	5
2 Literature Review	6
2.1 Artificial Intelligence in Education	6
2.1.1 Definition of Artificial Intelligence	6
2.1.2 AI-Powered Education Platforms	7
2.2 Chatbots in Language Pedagogy	7
2.2.1 Chatbots' Prevalence in Language Pedagogy	7
2.2.2 Text-Based and Voice-Based Chatbots	8
2.2.3 Chatbots' Usage Environment and Conversation Content	9
2.2.4 User Experience with Chatbots in Language Education	10
2.2.5 Learners' Motivation and Willingness to Communicate in a Second Language	12
2.3 Summary	13
3 Methodology	15
3.1 Research Approach and Research Phases	15
3.1.1 Research Approach	15
3.1.2 Research Phases	16
3.2 Empirical Study and Design Procedure	18
3.3 Data Collection and Analysis Methods	18
3.3.1 Data Collection Methods	18
3.3.2 Data Analysis Method	20
3.4 Chatbot Application for This Study	22
3.5 Participants and Research Ethics	24
3.5.1 Participants	24
3.5.2 Research Ethics	25

4 Findings	26
4.1 Expectation for Chatbots as a Korean Language Learning Tool	26
4.2 Guidance for Interaction with Chatbots in Korean	29
4.3 Motivations to Use Chatbots for Korean Language Education	33
5 Design Implications for Korean Language Education AI Chatbots	39
5.1 Affordance	39
5.2 Feedback and Corrections	40
5.3 Discussion Topics	41
5.4 Conversation Interaction	42
5.5 Chatbot Character Design	42
5.6 Chatbots' Trustworthiness	43
6 Discussion	44
6.1 Discussion on RQ1	44
6.2 Discussion on RQ2	45
6.3 Discussion on RQ3	46
6.4 Discussion on Design Implications	48
6.5 Limitations and Future Work	49
7 Conclusion	51
References	53
Appendices	60
Appendix 1	60
Appendix 2	61
Appendix 3	63
Appendix 4	65

# List of Figures

<b>Figure 1.</b> The Iterative Cycle of Human-Centered Design (source: The Design of Everyday Things - Norman, 2013, p222)	15
<b>Figure 2.</b> Excerpt from the current research's affinity diagram - Organized by themes	21
<b>Figure 3.</b> Univerbal application user interface (source: <a href="https://www.univerbal.app/">https://www.univerbal.app/</a> )	23
<b>Figure 4.</b> Korean conversation with chatbot on Univerbal (downloaded pdf version)	23
<b>Figure 5.</b> Expectations on AI applications' helpfulness (N=9)	28
<b>Figure 6.</b> Expectations on AI applications' human-like aspect (N=9)	28
<b>Figure 7.</b> Expectations on comfortability level in interacting with AI applications (N=9)	29
<b>Figure 8.</b> Univerbal's Create Your Own Topic setup	40

# List of Tables

<b>Table 1.</b> Current thesis project phases	16
<b>Table 2.</b> Semi-structured interview question sections	20

# 1 Introduction

This chapter discusses the motivation and context of this thesis work, followed by the definition of the research goal and research questions, as well as the thesis structure.

## 1.1 Motivation and Research Context

With the rise of the Korean Wave (also known as Hallyu)'s global popularity in recent years, more foreigners are interested in learning the Korean language. Indeed, Lee's (2014) research on beginner level foreign learners shows that K-pop and Korean pop culture are among their main motivations to learn Korean. This learning trend has spread over the world, from Asian countries like Malaysia (Nikitina & Furuoka, 2019), Vietnam (Han, 2023), to such Western countries as the United States (Lee, 2018). The methods used by students to learn Korean vary, yet over the past few years, one of the emerging educational tools for learning the Korean language as well as other languages is Artificial Intelligence (Kim & Su, 2024).

Although Artificial Intelligence has been around since the 20th century, the application of this technology in language pedagogy, especially speaking and listening practice, became prevalent only in the 2000s. The definition of Artificial Intelligence (AI) varies due to its several developments and discussions in diverse fields from anthropology, computer science, philosophy to linguistics and neuroscience (Luckin et al., 2016). This research work, however, focuses on the concept of AI chatbot in the context of language education. According to Clarizia et al. (2018), chatbot is a computer program capable of interacting with humans in dialogues to produce appropriate responses to humans' questions. Chatbots have been used increasingly in fields that require interaction, either verbally or in text form, with humans, such as travel and hospitality, financial services, healthcare, and education. With this ability to communicate with humans, AI-powered chatbots have gained more popularity in language learning in recent years, providing several benefits and increasing accessibility compared to the traditional classroom environment (Polakova & Klimova, 2024).

Several research has been conducted to investigate the advantages of AI chatbots in foreign language education. According to Haristiani (2019), AI chatbots are helpful as a language

learning tool in six ways. First, this technology provides an alternative to direct human contact, *making the interaction more relaxing*. Second, with chatbots, *the learning materials can be repeated to students for an unlimited number of times*. Third, AI chatbots *assist learners in practicing their reading and listening skills*. Fourth, studying with AI is *a novel and intriguing experience*. Fifth, this AI-powered learning method allows language learners *to utilize the grammars and vocabularies they have obtained*. Finally, with chatbots, students *receive immediate and useful feedback* on their orthography and language structures. These beneficial effects can be seen in learners of different ages, from elementary school to university, and different language abilities, from beginner to advanced levels (Lee & Lim, 2023; Lesia et al., 2022; Yang et al., 2022). Students across different backgrounds and cultures also have neutral to positive perceptions and attitudes towards the use of AI chatbots as a learning medium (Dahbi 2023; Lesia et al., 2022; Yang et al., 2022).

Nevertheless, little has been done in existing literature to understand Korean language learners' experience in using unstructured AI-driven conversational tools. These AI-powered applications provide free-form conversation simulations, similar to spontaneous dialogues with native speakers, in Korean as well as other languages. Although these tools have been available in the market and used by various learners, their use for Korean language education has rarely been studied. Most research conducted so far focuses on the effect of AI chatbots on English language learners. Moreover, the majority of AI chatbots were implemented and tested either in a classroom environment (Jeon, 2024; Yang et al., 2022), or in a research facility where students' interaction with the AI-enhanced tools were monitored (Divekar et al., 2022). There has been research regarding the use of chatbot applications off campus, where learners could utilize the tools in their own space and time to improve their English (Polakova & Klimova, 2024; Tai, 2024). Yet, such settings have not been studied on Korean language learners. Additionally, existing research mainly studied AI chatbots as an assisting tool for the curriculum students followed in class. As a result, the chatbots' educational contents and responses were semi-structured, predetermined based on the material students obtained from their language instructors at the time of the experiments (Lee & Jeon, 2024; Tai, 2024). An unstructured, more liberal form of interaction between AI chatbots and learners, where learners are free to use any vocabulary and grammar structures they want, and the content as well as the length of the conversation are not predetermined, has not been explored.

Additionally, while there has been research on the education of various languages such as Finnish, Russian, and German using smart technology in Finland, the exploration of Korean language training is still lacking. There is a significant demand for Korean language education in Europe in general (Carson & Do, 2013) and in Finland in particular. As a result, it is valuable to investigate learners' experience studying Korean using AI-powered applications in Finland. Korean studies have been provided in universities and institutions across Europe since 1947 (Carson & Do, 2013). In Finland, major universities (Helsinki University, Turku University) as well as other public organizations (Helsingin Aikuisopisto, Turun kesäyliopisto, Tampereen kesäyliopisto, KSI Tampere) and private institutions (Studentum, Kanjikaveri) offer Korean courses for various proficiency levels.<sup>1</sup> Besides the traditional classrooms, students in Finland learn languages through technology assisted platforms and applications. For instance, there have been studies in Finland about Russian language education in a virtual environment (Nikunlassi et al., 2021), Finnish morphology learning using mobile applications (Salmela et al., 2024), as well as English, Finnish, and German pedagogy enhanced by virtual reality technology (Kallioniemi et al., 2019). Nonetheless, Korean language training using smart technology has not been explored in Finland.

Given such a gap in current research and the demand for Korean education in Finland, it is worth looking into how Korean language learners utilize and interact with AI-powered tools. Thus, the aim of the study is to explore user experience in learning Korean with free-form conversational AI. The results from this study can contribute to existing literature regarding the application of AI in Korean language education. Specifically, the current study is expected to shed more light on Korean learners' experience in AI-powered free talking outside of the classroom environment. Furthermore, it will propose new ideas and add to the current design practices of AI chatbot as a conversational partner specifically in Korean language education.

---

<sup>1</sup> Universities:

<https://studies.helsinki.fi/instructions/article/language-centres-curriculum-korean>  
<https://www.utu.fi/en/university/faculty-of-social-sciences/ceas/language>

Public organizations:

<https://helao.fi/kurssit/kiellet/korean-kielikurssit/>  
<https://turunkesayliopisto.fi/koulutustarjonta/korean-kieli-ja-kulttuuri-6-5538/>  
<https://tampereenkesayliopisto.fi/en/oppiaineet/153/korean/>  
<https://www.vimmart.fi/ksitamper/>

Private institutions:

<https://www.studentum.fi/koulutushaku/korea-suomi/c830-d1607>  
<https://www.kanjikaveri.net/korea/ukk.php>

The empirical study of this research will be conducted online and outside of the classroom context. Such an environment is expected to maximize users' easement and freedom in their learning schedule and interaction with the AI chatbot. This method, which is flexible in terms of time and space, has been applied in previous research where users studied English with a chatbot outside of the classroom environment (Dahbi, 2023; Polakova & Klimova, 2024).

## 1.2 Research Goal and Research Questions

The overall goal of this thesis is *to understand Korean language adult learners' experience in learning with free-form conversational AI and their willingness to continue using this tool in their study journey*. This thesis work aims to answer the following research questions:

**RQ1:** What are the learners' expectations of conversational AI as a tool to improve their Korean language skills?

**RQ2:** What kind of guidance the learners need in free-form conversational practice with AI?

**RQ3:** What factors affect the learners' motivation and willingness in using AI free-form conversations to improve their Korean language skills?

An empirical study is designed and implemented to address these research questions. Particularly, a questionnaire will be delivered to the participants (N=9), where they respond to inquiries regarding their expectation and any prior experience using AI-enhanced conversational tools for learning purposes. Next, participants will experience an AI-powered conversational application independently for seven days, during which a journal will also be utilized. Following this testing period, an online, semi-structured interview will be conducted with each user to gain further understanding of their expectation, experience, and motivation using the chatbot tool. The data obtained from the questionnaire, the journal as well as the semi-structured interview aim to answer the research questions and provide insights into designing the experience of Korean language learners with AI.

### **1.3 Thesis Structure**

This thesis comprises seven chapters, starting with Chapter one, where the research motivation and context, research goal and research questions, as well as the structure of the research work are introduced. Chapter two contains related literature that supports this research, including the definition of AI and AI-enhanced chatbots in education, along with the growing presence of chatbots in language pedagogy. This chapter also covers chatbots' design for education, learners' experience, motivation and willingness to communicate in their target language using such AI-powered tools. Chapter three describes and explains the research participants and the mobile application used for the empirical study. In this chapter, the research approach, data collection and analysis methods selected for this study, along with the research phases, and experiment procedure will also be elaborated on.

Chapter four showcases the findings from the empirical study, namely the questionnaire, the usability testing, and the semi-structured interview. Based on the research findings, chapter five presents the design implications for out-of-class Korean language education chatbots. Chapter six considers relevant discussion points, the limitations of the research and possible future work. Chapter seven covers the conclusion of the thesis work.

## 2 Literature Review

This chapter covers relevant work that provides fundamental concepts and knowledge for the current study. Section 2.1 provides the definitions of Artificial Intelligence (AI), together with AI's emerging application in language education. Section 2.2 covers the definition of AI-enhanced chatbots, chatbots' application in language pedagogy, followed by learners' experience utilizing chatbots to study a second language. In this section, learners' motivation and willingness to communicate in their target language while using chatbots are also investigated. This chapter ends with Section 2.3, where a summary of the chapter content is provided.

### 2.1 Artificial Intelligence in Education

#### 2.1.1 Definition of Artificial Intelligence

The OECD, in their 2023 publication (p7), defined an AI system as “*a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.*”<sup>2</sup>

Since its first description by Alan Turing in 1950, AI has been continuously developed and applied in several fields such as customer service, medicine, computer science, linguistics, philosophy and education (Mariani & Borghi, 2024; Luckin et al., 2016; Ramesh et al., 2004). Its interfaces also vary, depending on the purposes and technological advancements. For instance, Eliza, one of the first AI-enhanced tools that could interact with humans in human language, was a text-based program run on a desktop computer (Weizenbaum, 1983). Apples' Siri and Amazon's Alexa, AI virtual assistants, can be accessed on smartphones or speaker devices using vocal commands (Hoy, 2018). On the other hand, social robots such as Pepper, with a humanoid physical appearance and installed AI-powered software, have been applied in the service industry, medical facilities, and educational institutions (Mishra et al., 2024).

---

<sup>2</sup> OECD, “Recommendation of the Council on Artificial Intelligence” at <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>

### **2.1.2 AI-Powered Education Platforms**

In the field of education specifically, AI has been widely studied and utilized to assist learners in several ways. Nadia Ahmad et al. (2024) research, by surveying medical undergraduates' experience and opinions on AI in medical education, found that AI has positive impacts on enhancing the students' information recall, time management and diminishing partiality in their assessment. In their systematic review, Papakostas et al. (2021) outlined both humanoid and non-humanoid social robots that have been applied in pedagogic programs for children aged 4 to 13 with special needs across the world. These AI-powered robots could play several roles such as a companion, social facilitator, a teaching assistant, and a rehabilitation tool for young learners with mental or physical disabilities.

In language education, AI has facilitated both language instructors and students. Specifically, AI contributed to teachers' syllabus design, educational method and assessment, study environment, as well as learners' personal experiences and outcomes (Semerikov et al., 2021). Although AI has numerous applications, Chatbots are among the most popular AI-driven systems utilized to enhance educational activities (Okonkwo & Ade-Ibijola, 2021).

## **2.2 Chatbots in Language Pedagogy**

### **2.2.1 Chatbots' Prevalence in Language Pedagogy**

Given their positive impacts on learners, chatbots have become a more and more dominant AI application in language education.

Jose and Jayaron Jose (2024, p5) defined chatbot as *“a computer program that employs artificial intelligence (AI) and natural language processing (NLP) to comprehend customer inquiries and automatically generate responses, mimicking human conversation.”* ELIZA and PARRY, the first chatbots, were developed respectively by Joseph Weizenbaum in the 1960s and Kenneth Colby in the 1970s (Weizenbaum, 1983; ZEMČÍK, 2019). Since then, chatbots have been growingly applied in second language (L2) pedagogy. For instance, in Yang et al.'s (2022) research, an AI chatbot called Ellie was implemented to help ten-to-eleven-year-old pupils practice English conversations in a restaurant, a shoe store, and a party room scenario. In another study, Divekar et al. (2022) combined virtual reality with chatbot to facilitate university students in learning beginner's level Chinese vocabulary through role-playing. In a Korean

language classroom, an AI-powered chatbot was integrated to assist learners aged 19 to 22 years old in interacting in Korean (Kim & Su, 2024). Indeed, the growing applications of chatbots in language education can be seen in the significant number of research articles on this subject over the past years. Haristiani (2019), for instance, took a closer look at five different chatbots utilized in language pedagogy, while four years later, Jeon et al. (2023) reviewed 37 studies on 39 speech-recognition chatbots in the same education field.

One reason chatbots are beneficial to language learners, compared to other forms of AI-enhanced educational technology, is their ability to provide an interactive learning environment, similar to the experience with a human instructor (Clarizia et al., 2018). Chatbots can communicate with learners in dialogues, adjusting their responses according to each student's inputs (Ciechanowski et al., 2019). This ability to maintain human-like interactions, provide instant feedback while being a non-human assisting tool, boost students' confidence in practicing conversations in their target language, which is one of the key components in language learning (Annamalai et al., 2023; Divekar et al., 2022).

Another advantage of chatbots is their convenience to use and personalization to individual students' needs. Whether in an offline classroom or an online platform such as websites and mobile applications, chatbots can be easily utilized by both learners and teachers (Jeon, 2024; Okonkwo & Ade-Ibijola, 2021). Off campus, learners can practice their target language with chatbots anytime and anywhere they wish, in different proficiency levels according to their own needs, and with unlimited number of uses (Annamalai et al., 2023; Haristiani, 2019; Lee & Lim, 2023). This flexibility enhances accessibility and increases users' practice frequency, which in turn have a positive effect on their learning progress (Annamalai et al., 2023; Dizon, 2017).

### **2.2.2 Text-Based and Voice-Based Chatbots**

Advancement in natural language processing (NLP) and computing capacity have allowed chatbots to understand and communicate with humans through speech, besides the text-based option (Lee, 2018; Wang, 2024). While text-based chatbots depend on texts or messages for interaction, voice-enabled chatbots allow learners to converse with them directly in oral form. For instance, in a study regarding voice-controlled conversational agents (VCAs) as a language partner, nine-year-old English language students used only speech to perform both small talk and specific task activities with a VCA (Lee & Jeon, 2024). With their capacity to orally interact with humans, AI-powered voice-enabled chatbots facilitate learners in practicing their speaking

and listening skills, two crucial components in language pedagogy. Indeed, Ali's (2020) systematic review on the application of AI in English teaching and learning indicates the role of AI in recognizing and evaluating human speech, as well as encouraging students' interaction and efficiency in a flipped learning classroom.

Human-like voice has been applied in chatbot's design for various language education projects. For instance, Tai and Chen (2024) applied OpenAI integrated CoolE Bot with a human-like voice to support 85 Taiwanese sixth graders in learning English with features such as English - Chinese translation, replay, and both vocal and text-formed communication. The result indicated that learners appreciated CoolE Bot's human-like voice, which made the interaction amusing and similar to that with native speakers. They also perceived the bot as a helpful interlocutor partner, whose captivating conversational style brought enjoyment to their English learning experience. In another study, a voice-controlled conversational agent (VCA) with a human-like voice was utilized to interact in English with 67 Korean students aged nine. After the experiment, the learners showed different perceptions of the VCA, ranging from VCA as a human, as an object, and a mix of the two. 71.6% of the participants assessed the VCA as a human-like interlocutor, whereas 28.4% regarded the VCA as a monodirectional machine resembling a dictionary (Lee & Jeon, 2024).

### **2.2.3 Chatbots' Usage Environment and Conversation Content**

AI-powered chatbots have been implemented both in the classroom and outside the classroom environment. In the classroom, chatbots are a part of the teaching material and tool, assisting the human instructor in language education (Jeon, 2024; Kim & Su, 2024; Yang et al., 2022). Outside the classroom, chatbots have been presented as a mobile or computer-based application, or integrated in other Internet of Things (IoT) devices (Dahbi, 2023; Lee & Jeon, 2024; Polakova & Klimova, 2024; Tai, 2024). In the out-of-class environment, students' interaction with chatbots can be either monitored by researchers in a research location (Divekar et al., 2022) or unmonitored in the students' own space and free time (Polakova & Klimova, 2024; Tai, 2024).

In most existing studies, the conversation contents with chatbots are predetermined based on the students' learning material in class or language proficiency level, regardless of the chatbot usage environment. For example, in Tai's (2024) study, despite the out-of-class environment and freedom to practice in their own time, B1-level English learners were given a list of fixed

topics to exchange with Google Assistant. Even in small talk activities, the conversational subjects have been selected and provided to learners in advance (Lee & Jeon, 2024). Little research has been done to understand users' communication with AI-powered chatbots about topics of free choice in an unsupervised environment.

#### **2.2.4 User Experience with Chatbots in Language Education**

The positive experience chatbots bring to L2 pedagogy have been studied and confirmed by students across different age groups and language proficiency levels.

First, language learners' pronunciation, speaking, and listening efficiency are enhanced through practice sessions with chatbots. For instance, Hsu et al. (2023) research showed how Amazon smart speakers strengthen English language learners' oral skills. In this study, the three authors conducted an experiment with 24 second-and-third-year-college learners from Taiwan, where they participated in small talks and acquired information by vocally interacting with Alexa. They then completed a mock TOEIC test, and their results were compared to those in the control group. According to the research findings, practice sessions with Alexa assisted students in recognizing their mistakes, adjusting their sentences, and enhancing their overall speaking skills. In a study by Tai (2024), 89 college students in Taiwan, aged 18-20 years old, were assigned into three groups to interact in English with either Google Assistant (GA) on mobile, L1 speakers, or L2 speakers. The study shows that by giving L2 learners constant feedback and frequent oral input, GA enhanced learners' pronunciation. GA also engaged learners in meaningful English conversations, resulting in their progress in speaking proficiency and fluency. Besides pronunciation and oral skills, listening skills were also exercised and enhanced through communication with AI chatbots (Annamalai et al., 2023; Tai, 2024).

Next, chatbots provide learners with opportunities to exercise their writing and reading skills, along with acquiring grammars and vocabulary. Specifically, chatbots assist L2 learners in recognizing and fixing their grammar mistakes, together with improving all four language skills (speaking, reading, listening, writing) through frequent interaction in an interesting and comfortable environment (Annamalai et al., 2023). This finding aligns with Kim & Su's (2024) implication regarding AI chatbot's role in helping L2 learners expand their language skills by utilizing correct grammar and diverse expressions during their interaction. Moreover, by giving new information and immediate feedback in multiple forms (written, visual, and audio), as well

as motivating self-directed learning, AI-powered tools like Google Assistant strengthened L2 students' vocabulary (Tai, 2024).

Additionally, practicing language skills with chatbots is flexible in schedule and personalized based on each learner's level, resulting in better accessibility and higher study frequency among learners. This benefit of chatbots was addressed in Annamalai et al.'s (2023) research where they surveyed Chatbot utilization in English language education among 360 Malaysian undergraduate learners. Specifically, participants stated that using chatbots, they had more independence and freedom in studying, able to arrange their own schedule regardless of their tutors' time frame. This flexibility motivated them to study as much as possible, and there was no limitation in the number of times chatbots respond to the users (2023, p5-6). In Lee and Lim's (2023) research, novice English language pupils were motivated to continue learning with an AI interactive agent designed to match the pupils' language proficiency level. With more time spent studying, their learning progress was influenced positively. Chatbots' availability to participate in L2 conversations with learners beyond the classroom schedule is also perceived as increasing learners' accessibility in L2 practice in Dizon's (2017) study with Japanese students who studied English with Alexa in Japan.

Furthermore, chatbots improve students' self-confidence and enjoyment in learning. Practicing a foreign language with AI, instead of humans, makes the experience less stressful, more pleasant, and helps students feel less embarrassed when making mistakes (Annamalai et al., 2023; Divekar et al., 2022). An undergraduate learner in a study conducted by Annamalai et al. (2023, p5) has stated *"I feel more relaxed when I am talking to a machine than a person and many will be reluctant to jump into a conversation with a stranger when they are still learning"*. Indeed, the tolerant and safe environment created by AI chatbots, coupled with their capacity to adapt to individual learner's language ability, allows learners to step out of their comfort zone and learn from experiment while promoting positive emotions such as engagement and satisfaction (Xiao et al., 2024).

Besides students, chatbots are also beneficial to L2 instructors. Gökçearsan et al. (2024) systematic analysis of published research in Social Science Citation Indexed Journals indicates that the use of AI chatbots contributes to teaching resources, decreases educators' workload while being financially affordable. Interviews with English as a Foreign Language (EFL) teachers regarding the role of AI in their classrooms show that they overall appreciated the benefits AI

brought. Besides sharing the teaching duties, AI assisted educators in ensuring the students understand their lessons through targeting specific learning material, be it writing, pronunciation, or grammar (Sumakul et al., 2022).

### **2.2.5 Learners' Motivation and Willingness to Communicate in a Second Language**

AI chatbots encourage second language (L2) learners to communicate in the target language, which is critical in achieving speaking proficiency. MacIntyre et al. defined Willingness to Communicate (WTC) in L2 as "*a readiness to enter into discourse at a particular time with a specific person or persons, using a L2*" (1998, p547). Kim and Su (2024) research regarding AI chatbot's effect on learners' WTC indicated that learners found it more relaxed sharing thoughts and practicing Korean with a chatbot, resulting in their increased WTC. This finding is in line with the EFL students' opinions about using chatbot in their classroom in Jeon's (2024) study. During the experiment, AI chatbot had a positive impact on learners' WTC, thanks to its ability to decrease their speaking anxiety. Some students were ready to interact in English with the chatbot, despite technical hindrance.

Several studies also stated that learners are motivated and engaged while utilizing AI chatbots for foreign language educational purposes. Okonkwo and Ade-Ibijola (2021) systematic review reported that in creating a relaxing and enjoyable studying environment, AI chatbots maintain students' engagement and motivation in interacting in the target language. Additionally, according to Sumakul et al. (2022), based on teachers' reports on students' happiness owing to the pleasure and learning autonomy in practicing English with chatbots, this AI-powered technology can boost learners' motivation in EFL classrooms.

Both motivation and willingness to communicate (WTC) in L2 contribute to L2 learning advancement. By conducting surveys on 59 ESL learners, aged 22 to 68, in New York, Lao concluded that there is a relation between students' motivation and their WTC in L2, along with perceived competence. Consequently, the author implied that boosting learners' confidence and self-perception would lead to more frequency of L2 communication, which in turn would result in L2 learning progress (2020).

## 2.3 Summary

AI-powered chatbots have been experimented and proved to be valuable in foreign language (L2) education. The benefits range from boosting learners' confidence in communicating in L2, improving their pronunciation, expanding their grammar and vocabulary, to enhancing all four skills of language proficiency (reading, writing, speaking, listening) (Annamalai et al., 2023; Tai, 2024; Xiao et al., 2024). In addition, the flexible and non-judgemental learning environment chatbots create has a positive impact on users' motivation and willingness to communicate in their target language (Jeon, 2024; Okonkwo & Ade-Ibijola, 2021). Previous studies have also indicated that chatbots are appreciated by both language instructors and students, and utilized both inside and outside a classroom environment (Dahbi, 2023; Sumakul et al., 2022; Tai, 2024).

With these proven capabilities in the education of multiple languages, chatbots are a potential tool that can also support Korean language learners. Lee (2014) indicated that K-pop and Korean pop cultures were the most influencing factors on learners' motivation to pursue Korean language education. Furthermore, among various learning contents, beginner's level students were most interested in "*Daily life in Korean society*" and "*Greetings/etiquette of Korean society and language*" (p211-212). For Korean learners with higher proficiency, the goal was more career oriented, focusing on their capacity to communicate in Korean in business settings (Han, 2023; 곽, 2019). Regardless of proficiency levels, the desire for Korean language communication, especially in verbal form, is undeniable. In Finland, spoken conversation is also a component of currently available Korean courses. Chatbots, with their conversation capacity, learning level versatility, and flexibility in terms of schedule and study environment, can be promising in assisting Korean language learners of different proficiency levels in their study journey.

Nevertheless, while several researchers have focused on English as the target language in their studies, little has been done to understand Korean language learners' experience with chatbots. Kim and Su (2024) project experimented with a Korean language chatbot, yet the interaction between students and the chatbot was set on the school campus, with the language instructor's presence and intervention when necessary. The chatbot's activities were designed based on the course syllabus, and the users' responses were predetermined in their course book. Consequently, although the research indicated chatbot's beneficial impacts on learners' WTC

and their positive experience using chatbots for Korean learning, it is unclear whether there would be similar outcomes when the chatbot is utilized off campus, without any supervision or any pre-structured content.

Additionally, Korean language students' experience learning with AI-powered chatbots has rarely been studied, if at all, in Finland. Here, advanced technological tools such as virtual reality and smartphone applications have been applied in the education of numerous languages like English, Russian, German, and Finnish (Kallioniemi et al., 2019; Nikunlassi et al., 2021; Salmela et al., 2024). However, when it comes to Korean, little research has been done. Given the significant number of Korean language learners in Finland, it is worth exploring chatbots as a medium for Korean language education.

Furthermore, although numerous studies discussed learners' perception on using AI chatbots for L2 education (Dahbi, 2023; Lee & Jeon, 2024; Polakova & Klimova, 2024; Yang et al., 2022), learners' expectations on AI chatbots as a language partner has not been addressed. Specifically, several studies indicate that learners of different backgrounds and age ranges, after practicing foreign languages with chatbots, have neutral to positive attitudes towards the experience. They perceived chatbots as human-like (Lee & Jeon, 2024), user-friendly, and pleasant to use (Polakova & Klimova, 2024). The students also appreciated that chatbots helped improve their English-speaking skills (Polakova & Klimova, 2024; Yang et al., 2022), explain grammar points and maintain their learning motivation (Dahbi, 2023). However, it is unknown whether these experiences met the students' expectations prior to the research experiments.

All reviewed relevant work contributed to the understanding of chatbots' advantages in L2 education and the users' experiences applying these AI-powered tools in their study. Previous research's findings also suggest that Korean language learners' interaction with chatbots in free-form conversations has not been thoroughly studied, hence worth further investigation.

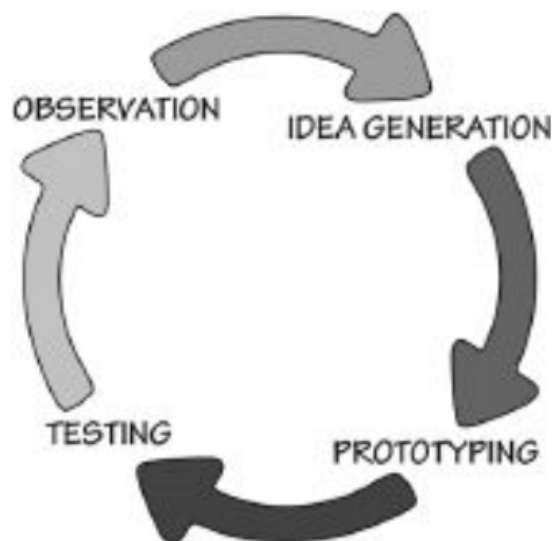
## 3 Methodology

In this chapter, details of the research approach and the research phases are showcased in Section 3.1. Section 3.2 demonstrates the empirical study design and procedure, followed by the data collection and analysis methods in Section 3.3. Section 3.4 describes the chatbot application utilized during the empirical process. Finally, Section 3.5 presents the study participants and explains the ethical principles applied in the current research.

### 3.1 Research Approach and Research Phases

#### 3.1.1 Research Approach

This study applies the *Human-Centered Design (HCD)* philosophy described by Don Norman in *The Design of Everyday Things* (2013). This approach emphasized on the creation of intuitive, efficient products that are aligned with users' needs and abilities. Norman stated, "*Design is concerned with how things work, how they are controlled, and the nature of the interaction between people and technology*" (Norman, 2013, p5). As a result, the design process starts with understanding the users' needs, followed by continuous testing and updating of design solutions until such needs are met. The author summarized this design process in four steps: Observation, Idea Generation (Ideation), Prototyping, and Testing.



**Figure 1.** The Iterative Cycle of Human-Centered Design (source: *The Design of Everyday Things* - Norman, 2013, p222)

In his book, Norman (2013) also discussed the core principles of HCD, which include *affordances*, *signifiers*, *feedback*, and *constraints*. Affordances refer to the qualities of an object that suggest its use, while signifiers communicate how users should interact with it. Feedback informs users about the results of their actions, helping them to understand whether they are progressing as intended. Constraints, on the other hand, limit possible actions, helping to prevent errors by narrowing choices to those that make sense for the task at hand.

The current study adopted the HCD approach by following the first two steps of the HCD process, namely Observation and Ideation. Specifically, the empirical study of this research investigated participants' needs in learning Korean and their experiences using an AI-enhanced tool for Korean conversation. The findings from the experiment were then the references to generate AI chatbot's design implications for Korean language learning. Additionally, the chatbot design ideation was formulated with respect to the HCD principles of affordances, signifiers, feedback, and constraints.

Additionally, the *Remote Unmoderated Usability Testing Method* was also applied during the current research. It was aimed to create a study environment similar to a real-life scenario, where the participant feels comfortable studying Korean in their own time and space. With this method, the researcher is not in the same location with the participant during the testing, and there is no facilitator (Bleecker & Okoroji, 2018). Testing guidelines (see [Appendix 3](#)) were prepared by the researcher and given to the participant in advance, who completes their tasks on their own. Each participant was free to utilize the AI-powered chatbot the way they wished for Korean language learning purposes, with at least one time practicing a conversation with a topic of their own choice.

### 3.1.2 Research Phases

The current study consists of five phases, presented in the table below.

**Table 1.** Current thesis project phases

Phase	Tasks	Tools/Methods	Results	Schedule
Phase 1 Literature	- Identifying existing work related to AI,	- Screening - Skimming	Research goals, approaches,	September 2024 -

Review	chatbot, chatbot's design & application in language education - Exploring research data collection & analysis methods - Organizing selected literature	- Note-taking - Organizing in excel sheet	methods, fundamental & relevant knowledge	March 2025
<b>Phase 2</b> Questionnaire	Preparing questionnaire based on related literature	- Consent form - Online questionnaire	Qualitative data (participants' background & expectations)	November 2024 - December 2024
<b>Phase 3</b> Usability Testing & Interview	- Testing app & designing experiment activities - Interviewing participants to collect qualitative data	- Consent form - Testing instruction - Univerbal app - Testing journal - Semi-structured interview - Audio recording	Qualitative data (participants' experience, expectation, motivation, feedback)	November 2024 - December 2024
<b>Phase 4</b> Result Evaluation & Findings	Organizing & analyzing collected data	- Transcription - Affinity diagram	Qualitative data to answer three RQs	December 2024 - January 2025
<b>Phase 5</b> Design Implications	Creating design guidance & suggestions based on research findings	- Organizing - Highlighting	Design guidelines for AI-enhanced Korean conversation practice app	January 2025 - February 2025

## 3.2 Empirical Study and Design Procedure

The empirical study was conducted entirely online to maximize the participants' freedom and comfortability in utilizing the chatbot application, as well as diminish restrictions in terms of space and time. Before the start of the experiment, *participants were provided with a consent form* and an information sheet detailing the research purpose, description, procedure, data collection, usage, storage and protection according to the EU's GDPR and Finland's law. The participants' rights and the project contact information were also provided in these documents (see [Section 3.5](#)).

After signing the consent form, each participant began the experiment by filling out *an online questionnaire about their background, experience and expectations for AI as a language learning tool*. Once the questionnaire was complete, each participant received instructions to *independently test an AI chatbot application* (see [Section 3.4](#)) during a period of seven days. The users were free to utilize the application the way they wish, regardless of usage location, frequency, and length. However, they were required to participate at least once in a free-topic conversation with the chatbot. Additionally, the participants were also instructed to keep a journal in which they wrote down their opinions and feedback regarding their experience with the chatbot (see [Section 4.3](#)). Finally, within one week after the application testing period, *each participant participated in one semi-structured interview* to discuss and elaborate on their expectation, experience, and motivation in learning Korean in general and learning Korean with the AI-enhanced chatbot in particular. During the interview, participants referred to the journal they kept and some also captured and mentioned screenshots of their conversation sessions. Besides a set of predetermined questions, each participant was also asked follow-up questions based on their responses to the initial inquiries. Each interview lasted between 20 to 60 minutes. All interviews were conducted online via Zoom and audio recorded.

## 3.3 Data Collection and Analysis Methods

### 3.3.1 Data Collection Methods

To collect data from participants, the following methods were utilized:

- **Questionnaire**

A questionnaire (see [Appendix 2](#)) consisting of eight inquiries was adapted based on previous literature to survey participants' expectations towards chatbots as a language learning partner. Specifically, 7-point Likert scale questions were utilized to understand participants' expectations towards chatbots' interface and interaction design (Berghe et al., 2021; Wiboolyasarin et al., 2024). Multiple choice and open-ended questions were also included to collect participants' age groups, their presumptions regarding chatbots' usefulness, and their comfortability in using these AI-powered tools (Dahbi, 2023; Kim & Su, 2024; Yang et al., 2022) (see [Section 3.5](#)). The participants were given the questionnaire online in Google Form and asked to respond to the questionnaire prior to their experience with chatbots.

- **Learning Journal**

During the seven-day unmoderated usability testing period, participants were encouraged to keep a learning journal. Suggestions for the learning journal content were prepared by the researcher and given to each participant in advance, together with the testing guidelines (see [Appendix 3](#)). The journal included the participant's opinions regarding the quality of the AI-driven conversations, application features, their motivations and preferences, as well as proposals on how to improve their experience with the AI chatbot. Participants were not requested to submit their journal so as to minimize any burden and restrictions on the writing format as well as the amount of time they had to spend on the journal. However, participants were requested to share and elaborate on the content of their journal during the semi-structured interview (see [Chapter 4](#)).

- **Audio-Recorded Semi-Structured Interview**

After the usability testing, a semi-structured interview (see [Appendix 4](#)) was conducted with each participant individually, with the goal of understanding their experience using a chatbot for Korean communication practice. The interview also shed more light on each participant's expectation in comparison to their experience. The interview method was selected as it is beneficial in collecting participants' self-reported experience and further elaboration on any initial feedback that was unclear to the researcher (Mackey & Gass, 2005). Follow-up questions were asked based on the participant's answer to the initial inquiry. The length of the interview varied from 20 to 60 minutes per participant.

The interview questions were divided into three main sections:

**Table 2.** Semi-structured interview question sections

<b>Section</b>	<b>Main Content</b>
General Korean learning experience & motivation	Questions regarding the participant's decision and motivation to learn Korean, their Korean learning methods and period, their general experience with learning Korean
Korean practice experience with AI chatbot	Questions regarding the participants' experience interacting with the AI chatbot, including their overall experience, description of a practice session, their expectations, feedback, emotions, progress ...
Motivation to continue using the AI chatbot tool	Questions regarding the participant's opinions about the AI chatbot as a language partner, their motivation to use the tool in the future

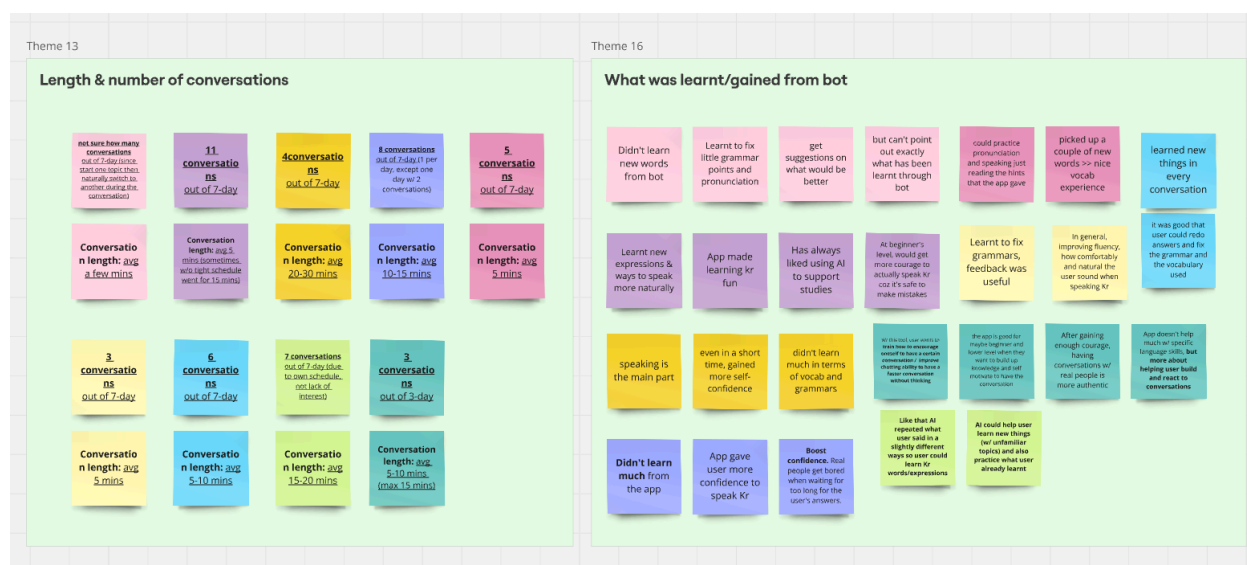
Each audio-recorded interview was then transcribed, where information was made anonymous and any personal information that could be used to identify the participants were removed from the transcripts. Since each interview was conducted between one researcher and one participant, without any other facilitator or note-taker, verbatim recording and transcription are reasonable mediums to ensure data was captured accurately and ready for the analysis phase.

### **3.3.2 Data Analysis Method**

The affinity diagram method was applied to analyze data gathered during the empirical process. By organizing research data, an affinity diagram, also known as the K-J method, arranges ideas into their relationships and discloses common themes. The affinity map is constructed upwards from each individual piece of information to categorize them into themes, without having any pre-determined structures. Information pieces are grouped together based on their similarities (Beyer & Holtzblatt, 1998).

With the affinity diagram method, this study aims to extract universal themes and patterns from participants' feedback and opinions gathered in both the questionnaire and the semi-structured interview. These themes and patterns would provide answers to the research questions and indications regarding the design of AI chatbots for foreign language education.

Conforming to the affinity diagram method, findings obtained from the research study were organized and encrypted. Specifically, participants' background information, Korean learning experience, opinions, and feedback collected from all steps of the empirical phase, namely the pre-test questionnaire, the individual journal, and the audio recorded semi-structured interview, were coded and formulated into notes. Each participant was coded in a different color note. Participants living in Finland were represented as **F1, F2, F3, F4** and **F5**, while those living in South Korea were marked as **K1, K2, K3** and **K4**. The affinity diagram was constructed using [Miro](#) (see [Appendix 1](#)), where all gathered data were organized by participants. Sticky notes with related contents were then grouped together and recurring themes were identified. There were a total of 789 sticky notes, 4 main groups divided into 32 themes. Some of the main themes discovered from the collected data include expected benefits from the AI chatbot, its' level of usefulness, learners' comfortability, conversation content, AI chatbot's feedback, AI chatbot's character, interaction difficulties, as well as conversation motivating and discouraging factors.



**Figure 2.** Excerpt from the current research's affinity diagram - Organized by themes

### 3.4 Chatbot Application for This Study

An AI-powered chatbot mobile application called Univerbal was chosen for participants to experience during the seven-day period.

Univerbal<sup>3</sup> was developed by Quazel Inc. headquartered in Switzerland. The application features an AI tutor that facilitates speaking practice through vocal conversations in several languages, including Korean. The application delivers a range of conversation themes based on each user's interests and their language proficiency level. During a typical conversation with AI, users are provided with instant feedback, translation tools available in different languages, and exercises relevant to the discussion topic. The tool also presents a "Create your own conversation" mode where users can decide what role the AI chatbot plays (for example, a friend, a job interviewer, a fictional character, a celebrity...), in what environment the conversation takes place (i.e., over the phone, in a coffee shop, in a garden, in a Renaissance fair...) and what topic to practice.

A conversation with an AI chatbot in Univerbal is unstructured, where the interaction between the chatbot and the user is unscripted and develops based on the user's input. Specifically, the AI chatbot asks questions related to the selected topic, to which the user is free to apply any grammar, vocabulary and expressions to reply in the targeted language. Based on the user's answer, the chatbot would respond and ask follow-up questions. Although the chatbot gives the user the choice to stop or continue the conversation after an appropriate amount of time based on the user's performance, there is no fixed timeframe, and the conversation can last for as long as the user wants.

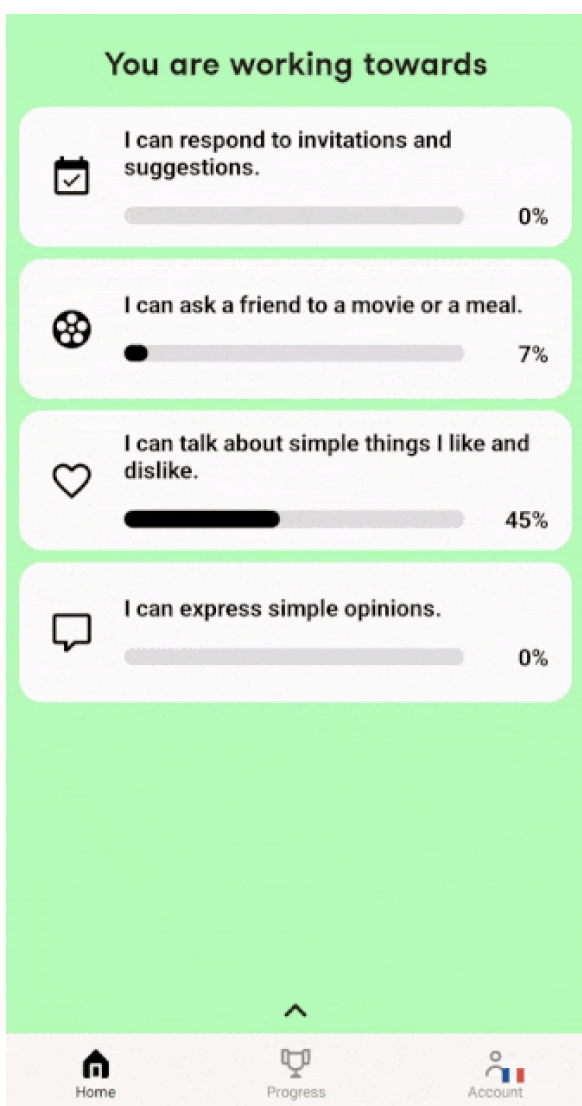
After comparing various voice-enabled chatbots for educational purposes available on the market (Talkpal, Tutor Lily, TalkOn AI, KSI Korean AI Tutor, Lango Talk, and Univerbal), Univerbal was considered the most competent and fitted application for this research, as its voice recognition for the Korean language is the most advanced, with both informal and formal spoken language options available depending on the AI's role. Both text-based and voice-enabled communication content are presented to assist users' interaction with the chatbot. In addition, there is no limitation in the length of the conversations, regardless of the

---

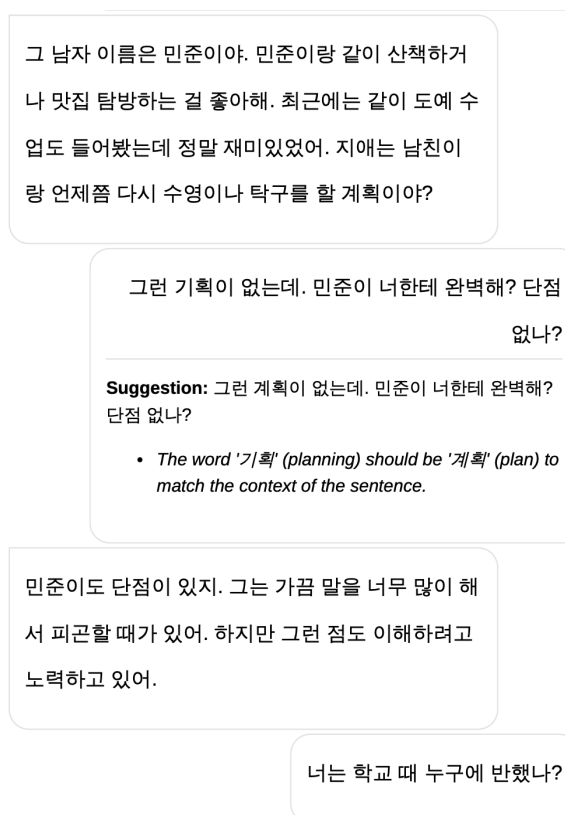
<sup>3</sup> Univerbal official website: <https://www.univerbal.app/>

selected topics. The application is also free for access on iOS and Android devices during a trial period in both Korea and Finland.

Figure 3 shows the Univerbal application's interface. Here, the application shows suggested goals and conversation topics based on the user's selected proficiency level and interests. The learner's study progress and learning streak are also presented. Figure 4 demonstrates a snippet of a free-choice-topic Korean conversation with the AI tutor, who played the role of a friend during the interaction. In this example, the AI tutor interacted with the user in Korean and gave suggestions and corrections in English.



**Figure 3.** Univerbal application user interface (source: <https://www.univerbal.app/>)



**Figure 4.** Korean conversation with chatbot on Univerbal (downloaded pdf version)

Before the seven-day experiment phase, the mobile application was tested, on both iOS and Android devices, in South Korea and Finland to confirm its availability and uniformity in these two countries.

## **3.5 Participants and Research Ethics**

### **3.5.1 Participants**

The participants were nine adult Korean language learners, aged 19 to 64 years old, living in either Finland (five learners) or South Korea (four learners). The nine learners come from different backgrounds with different mother tongues, and Korean is not their first language. They all participated in Korean classes before, but not at the time of this research experiment. Their Korean proficiency levels vary from A2 (upper beginner) to B2 (upper intermediate), divided equally between the South Korea and Finland locations. All participants speak fluent English and learn Korean as a foreign language. Since age, gender, and nationality can be considered as sensitive information, and the focus of the current study was on learners' experience with an AI-enhanced chatbot, such data was not strictly identified or included. All participants were recruited online and participated voluntarily in all empirical steps of this project, namely answering an online questionnaire, testing an AI-powered application independently, and participating in a Zoom semi-structured interview.

Regarding participants' experience of using AI, either text-based or vocal form, three out nine participants have used AI-powered tools for Korean language learning purposes, but none has used Univerbal prior to the current experiment. One participant living in South Korea has used ChatGPT for vocabulary translation and grammar explanation in English as an assisting tool during their Korean courses. One participant living in Finland has tried conversing in Korean with AI, yet AI's responses in Japanese made the experience unsuccessful. Another participant living in Finland used the recently released free vocal command feature of ChatGPT to assist their Korean learning, yet the experience was confusing and lacked conversation transcription. Nevertheless, all participants were familiar with the use of mobile applications in general, and eight out of nine have Korean language learning applications on their smart devices.

### 3.5.2 Research Ethics

This research was conducted in regard to the EU's GDPR and Finland's law to ensure the rights and welfare of the research participants. Prior to the experiment, participants were provided with a consent form (see [Appendix 1](#)) and an information sheet in Tampere University format (see [Appendix 1](#)).

The information sheet detailed the research purpose; description of the research process; procedure for collecting research data; potential risks and benefits of participation; data confidentiality, processing and storage; participants' rights and privacy protection policy; together with the researcher's contact information. Recorded interview audios from research participants were transcribed, where all identifying information was removed. Once transcribed, all recorded audios were immediately destroyed. Collected data, with no identifying information, is stored in Tampere University's drive, secured with passwords, for up to two years after the thesis is over. After this period, the collected data will be destroyed. The consent form contained similar information with the information sheet, formatted online and sent to the participants via email. Each participant had the research information explained to them in writing and the opportunity to raise any questions or requests prior to the start of the experiment. The experiment only started after each participant had signed the consent form, stating their understanding of all relevant information, and their consent to their participation in all steps during the empirical phase of the research. The signed consent form was sent to each participant for future reference and contact. The participants have the rights to withdraw from the study at any time during their participation and refuse to answer any questions without any consequences. Furthermore, before the beginning of the Zoom interview, the final step of the experiment, each participant was provided again with the description of the interview, their rights, and the opportunity to raise any questions or concerns. The audio recording of the interview was announced to each participant before it was recorded.

## 4 Findings

In this chapter, the research findings were organized in line with the affinity diagram method (Beyer & Holtzblatt, 1998) and presented following each research question.

### 4.1 Expectation for Chatbots as a Korean Language Learning Tool

As described in the previous chapter, participants responded to an online questionnaire at the beginning of the empirical study, which explored their background and expectations for AI-powered application as a Korean language learning tool. The participants' responses to this questionnaire provided valuable data that help answer the first research question of the current study. Furthermore, the semi-structured interviews conducted at the end of the experiment elaborated on the learners' expectation for the AI chatbot. The collected data were organized into the following themes: *overall expectations for the chatbot*, *factors influencing the participants' expectations*, *expected chatbots' benefits*, *expected levels of helpfulness and human-likeness*, and *expected level of comfortability using chatbots*.

*Overall expectations.* Overall, the participants had little to no expectation on AI as a Korean language educational tool. Three out of nine participants (**K1**, **K3**, **K4**), including those who had utilized AI tools such as ChatGPT for their Korean study, stated that they didn't know what to expect from Univerbal, the tool used in this study. **F3** set their expectation low, due to previous experience with AI-powered applications. Three people (**K4**, **F4**, **F5**) recalled that they were "*pleasantly surprised*" and "*positively surprised*" by Univerbal, and that the experience exceeded their expectations. **K3** claimed that after the experiment, their skepticism towards AI decreased, while **F3** shared that "*I feel like I would trust 70% if my trust earlier was around 20.*" when talking about their trust towards Univerbal compared to previous AI applications.

*Factors influencing the participants' expectations.* Their expectation was influenced by the technology they used before, whether it was for Korean language learning or for other purposes. For example, participant **F1**, familiar with using search engines such as Google to look for information, expected the AI-powered chatbot to have common knowledge and be able to correctly respond to questions that could be answered with a simple internet search. Two participants (**K1**, **K2**) were worried that the interaction with AI in Korean would be strange due to

their previous experience with other AI-powered applications. **K2** stated, *“At the beginning, I was a little bit scared that it's gonna feel unnatural to talk with AI, but it didn't.”* The reason was because K2 had a preconception of AI being either too robotic or overly human, making the conversation awkward on the mobile application. For one participant (**K3**), the notion of AI made them skeptical about its authenticity. They felt less enthusiastic and tried to avoid anything AI-related, saying *“Just because of the AI aspect of it. It's a word that has a lot of baggage. And so I was skeptical in the sense that I wasn't sure how interesting it would be.”* Regarding negative experience with AI, **F3** had tried to converse in Korean with an AI application prior to the current experiment, only to be replied in Japanese. Since then, their expectation for AI had been low, assuming there would be multiple illogicalities and glitches during the interactions. Similarly, **F2**, despite having little experience with AI tools, expected the AI application to be robotic, lacking human factors, such as age, name, and facial expressions.

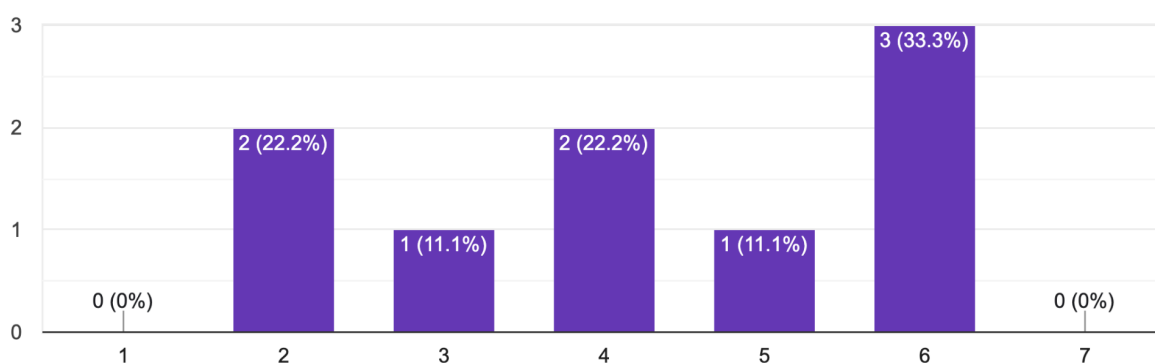
Besides previous experience with technology, participants' various methods of studying Korean also affected their expectation towards Univerbal. Specifically, since all participants have been learning actively through several means such as websites and mobile applications, social media, K-drama and K-pop, besides Korean language books and classes, they expected the AI-enhanced tool to be a mere extra means to practice what they have learnt elsewhere, rather than a main method to gain new knowledge of Korean. All nine users expected to utilize AI applications at home by themselves; while one participant expected to also use them with friends and family members at home, and another participant would use them in class as well. Additionally, as **K1** had experienced difficulty practicing Korean face-to-face with native speakers due to their low Korean proficiency, they expected the interaction with AI would be equally boring and weird, as there would not be anything to talk about.

*Expected chatbots' benefits.* The benefits AI chatbots could bring vary, according to the participants' views prior to the testing period. Four participants (**K2, K3, F1, F4**) expected enhancement in their speaking fluency and responsiveness in conversations through practicing with AI. Specifically, **K3** expected to gain *“Natural question-and-answer interaction”*, while **F4** wished *“to learn to be more eloquent, not only using simple words or simple language.”* Other possible advantages include improvement in pronunciation, vocabulary, grammar, speed, and confidence in using Korean. One participant (**F2**) thought of AI as a tool for *“gathering and searching data instead of Google.”*

*Expected levels of helpfulness, human-likeness, and comfortability.* The participants' opinion regarding how helpful and human-like AI should be, as well as their comfortability interacting with AI also vary. The participants were asked to grade AI applications on a scale of 1-7 (1 being the least, 7 being the most of each characteristic) for its helpfulness, human-like aspect and their comfortability interacting with AI. For helpfulness, the participants' responses ranged from 2 to 6 (see Figure 5), while the human-like characteristic ranged from 2 to 7 (see Figure 6). Regarding how at ease people are when using AI-enhanced applications, the answers went from 2 to 7 (see Figure 7).

AI 적용된 앱은 한국어 회화 능력에 얼마나 도움이 된다고 생각하시나요? How helpful do you think AI applications are for your Korean conversational skills?

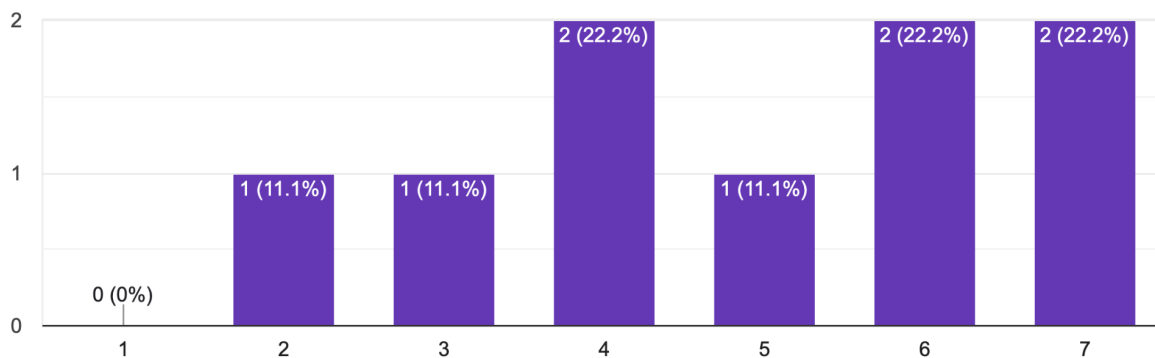
9 responses



**Figure 5.** Expectations on AI applications' helpfulness (N=9)

한국어 학습 대화에서 AI는 얼마나 인간 같아야 한다고 생각하시나요? How human-like do you think AI should be in a Korean practice conversation?

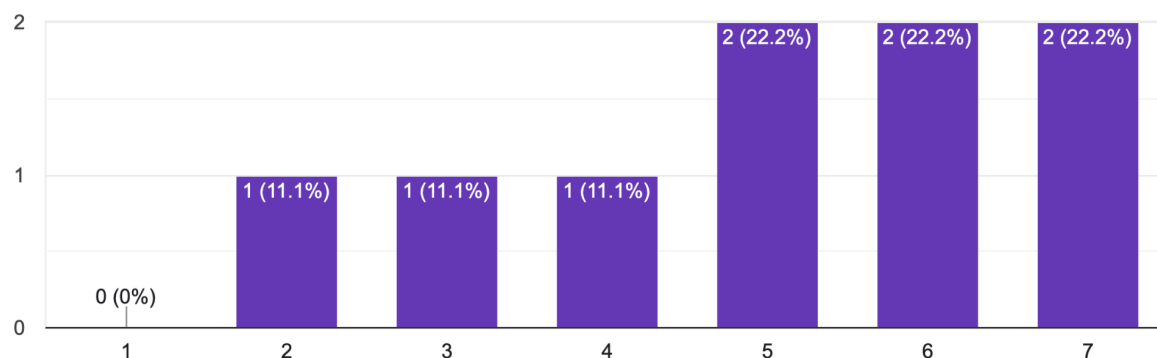
9 responses



**Figure 6.** Expectations on AI applications' human-like aspect (N=9)

AI 대화형 앱과 상호작용하는 것이 얼마나 편안하게 느끼시나요? How comfortable are you in interacting with AI conversational applications?

9 responses



**Figure 7.** Expectations on comfortability level in interacting with AI applications (N=9)

In short, participants had generally low expectations towards the use of conversational AI as an assisting tool for their Korean study. There was no clear distinction between the expectations of participants living in South Korea compared to those living in Finland, as both groups had little expectations and some levels of pre-perceptions of AI-powered tools. In addition, the two groups had similar opinions about what AI could help them in their Korean learning, including improving vocabulary, grammar, pronunciation, fluency and confidence in speaking, as well as searching for Korean language related information. The views of both groups regarding AI's usefulness, human factors, and their comfortability interacting with AI vary greatly on the scale of 1-7.

## 4.2 Guidance for Interaction with Chatbots in Korean

Through semi-structured interviews and journals shared by the participants during the interviews, valuable findings regarding the participants' experience, difficulties, and needs during their interaction with Univerbal AI chatbot were gathered as follows. These findings are categorized into the following themes: *the chatbot character's must-have information, freedom in creating conversation topics, vocabulary support, conversation flow, feedback quantity and quality, other conversational supporting features, conversation styles, learning format, as well as user interface design and affordances.*

*Chatbot character's must-have information.* Due to the nature of the Korean language, learners wished to know certain aspects of the AI conversation partner such as name and age. To decide what speech (formal, informal, casual, or polite) to use and how to properly address a counterpart in a Korean conversation, factors such as name and age are crucial. Univerbal chatbot, having no indication of such factors and giving no self-introduction in any conversation, made some participants confused and find it challenging to carry on the discussion. For example, in a discussion suggested by the application about butterflies with a dragon, **F5** reflected *"The one thing that I could personally use some help with is the formality level. So I don't know how to address a dragon. How polite should I be?"* **F2** and **F4** also raised similar concerns, saying it would be easier if chatbot was given a name. Especially for **F2**, this inhuman aspect of the chatbot made it robotic and the interaction with it unnatural.

*Freedom in creating conversation topics.* Regarding conversation themes and contents, participants had different opinions whether they would like to propose their own topics. **K2** and **F4**, for instance, had no difficulty knowing what they wanted to discuss. **K2** preferred to select their own topics, saying *"I could focus more on what I want to practice."* **F4** remarked *"I think this is very fun and the fact that you can choose what you want to talk about and with whom you can choose, even a celebrity if you want to, or like a deceased person that you want to talk with. I think it's very fun and you can use your imagination as much as you want to."* **F1** appreciated the topic suggestions from the application, yet they wanted the final decision to be theirs, since some proposed discussions, such as workplace themes, were not interesting. On a different angle, **F5** considered the application's topics to be original, inspiring but challenging for their Korean proficiency levels. Similarly, **F3** and **K3** found some of the presented discussions unfit for their levels, yet would like to modify them to make them more suitable. **K3** stated *"I like the chance that I would get something, at least in terms of a learner, that I would get something that I wouldn't think of myself, even though, like I said, I chose some easy topics in the beginning."* **K1** and **K4** favoured AI's available discussions, since they didn't know what to talk about. **K1** was also concerned about their ability to maintain the conversations, saying *"I feel like if I chose a topic, it would go in too deep that I cannot say anything anymore."*

*Vocabulary support.* Although most participants had difficulty finding the right vocabulary to continue at some point during the conversation, they didn't think of seeking help directly from the application. All users stated that they were not sure why they didn't ask for direct assistance

from the chatbot, and that they were not prompted to do so. **F3** commented on the reason why they didn't request any support from the chatbot "*I could have done that. It didn't cross my mind.*" Instead, participants referred to external sources such as Google to find the vocabulary they needed or tried to express themselves in other ways. From a different angle, **F5** believed it might be good to get help from a tutor character within the application, yet it was not entirely necessary. They recalled "*it's actually good that I do need to think about my answers and I use the translator only if I am completely stuck.*"

*Conversation flow.* Participants had dissimilar feedback regarding the flow of the conversations with Univerbal chatbot. During the discussions, all participants answered the chatbot's questions most of the time, instead of asking questions back and forth. Participants like **K2**, **K4**, and **F4** enjoyed the natural, human-like flow of the interaction, and were motivated to continue the conversations by AI's constant questions. They didn't mind that the chatbot was the main inquirer. **K4** commented "*the conversation was very flowing*" and "*it felt like a very smooth conversation, the reply seemed very natural from the AI as well.*" On the contrary, **F3** disliked the one-sided format of the conversation, saying "*I feel like the conversation structure was not very natural in that it would encourage me to ask questions, it was like, it felt more, I noted that this felt more like an exam than a conversation.*" **F1** also grew tired of the conversations quickly, since the chatbot kept asking similar questions without sharing about itself like a human being. Similarly, as AI always replied to each user's answers with another question, **K3** was not prompted to give any inquiries themselves. **K3** also mentioned "*in terms of the flow, it felt like a river, right, it kept moving, instead of stopping to look around at the scenery. I mentioned some disappointments, and I feel like I wish it had pushed me more deeply to talk about certain things, instead of just going on to the next thing quickly.*" In **F2**'s case, the initial reason for not asking the chatbot any questions was their uncertainty on how to address this counterpart. However, **F2** shared "*I think the bigger reason behind is maybe that, I may lack some kind of, you know, ability or experience using these kinds of apps, how to effectively use this.*"

*Feedback quantity and quality.* Additionally, the feedback and suggestion feature was considered useful by some participants, while for some others, it was either not obvious or not helpful. Six out of nine participants found AI's comments and corrections useful. Thanks to the chatbot's feedback, they learnt new vocabulary and more natural expressions, realized their mistakes regarding grammar, spelling, pronunciation, and use of words. **K3** shared "*I was glad with the suggestions, and they were all good suggestions*", while **F4** commented "*it does*

*actually correct your answers very effectively.*” On the other hand, **F1**, realizing that AI’s corrections on missing particles were repetitive, stopped checking them after a while. **F3** found the constant indication of little mistakes unnecessary, saying *“In the free conversation, the feedback wasn’t really helpful. And it was even a bit, I wouldn’t say discouraging, but it just felt like it was pointing out details instead of bigger things that I would more like to have feedback on.”*

*Other conversational supporting features.* Furthermore, the hints and missions presented during the conversations could assist users in maintaining the conversation. **K4** shared that the tasks given during each discussion gave them a clue on what to talk about. With this feature, **K4** felt that they could also test their vocabulary and conversation skills. Similarly, when **F3** found it challenging to reply to AI’s question, they relied on the hints provided by the app to continue the conversation. **F3** stated *“it would give me three hints and I could try saying those out loud. Then that made it more comfortable.”*

*Conversation styles and learning format.* Regarding the free conversation style, some learners found it useful and fun, while some others preferred a more structured format of learning. Participants such as **K2**, **K4**, **F4**, and **F5** were pleased with the free-style discussion format. When talking about the most engaging aspects of the app, **K2** stated *“I could choose a topic which I like, and then the conversation felt natural. And also, it was reactive, based on what I say. So if I answer, the bot’s reacting to my answer and following up from there.”* **F1** also appreciated the talking feature, saying *“the conversation feature was the most helpful.”* Similarly, **F2** found the unstructured conversation style enjoyable, as it gave *“this freedom to express yourself a little bit more freely using whatever grammar and words you want to use.”* On the contrary, **K1** wished to be given a set of vocabulary to practice on, instead of having to come up with the terminology on their own to maintain the conversation. **F3** appreciated the pre-built lesson option, where *“there was a certain chunk of stuff that you were supposed to learn and I kind of enjoyed that aspect. It was more structured.”*

*User interface design and affordances.* There was a need for clearer prompts to present features that support the conversations with AI. For instance, **K3**, not realizing they could communicate with AI in both text-based and voice-based forms, ended up texting the chatbot during the entire experiment period. **K3** commented *“I was not aware that I had the option to speak to the AI.”* **F2** requested feedback for their performance during the conversations, even

though this feature was already available on the application. **F5** didn't know they could decide and type in their own topics for each discussion, instead of merely shuffling a list of suggested conversations.

In summary, there are variations in the experience conversing with AI chatbot among nine participants, who encountered different challenges and appreciated different supporting features available during the AI-enhanced interaction. The concerns that participants had include the lack of chatbot's name and age, the unsuitable discussion topics, the conversation flow and structure, the lack of vocabulary support, the quality of AI's feedback and corrections, and the unclear presentation of available features in the application. The idea of learning Korean through freestyle conversation with AI also received mixed opinions, where some participants enjoyed it, while others preferred a more guided and structured learning format. Additionally, there was no clear difference in the experience of participants living in Finland compared to those living in South Korea, except that the learners in Finland concerned more about how to address the AI chatbot during their interaction.

### **4.3 Motivations to Use Chatbots for Korean Language Education**

During the semi-structured interviews, participants shared insights on what motivated them to learn Korean in general and to converse with AI chatbot as a means to support their study specifically. The collected data was grouped into the following themes: *the participants' own motivation to study Korean, their learning methods, their difficulties in learning Korean, their perception of the chatbot's advantages, conversation flow, the chatbot's interaction quality, the chatbot's human factors, discouraging elements of the AI-enhanced application, the participants' willingness to continue using the chatbot, and their opinions regarding their Korean learning motivations.*

*The participants' own motivation and Korean learning methods.* Prior to this research, all nine participants were highly motivated in their Korean study and utilized various methods to assist their learning, which greatly affected how they experienced the experimented AI-enhanced application. All participants had taken Korean courses and seven had face-to-face language exchange with native Korean speakers. Moreover, all have been actively learning Korean via K-drama and K-pop, social media, as well as websites and applications. The reasons each

participant decided to learn Korean slightly varied, yet they were all stimulated by K-drama and Korean culture. **F2**, **F3**, and **F4** loved the sounds of the language, while **K1**, **K2**, and **K4**'s main study purpose was for communication with Korean speakers, especially family and friends. **K3**, **F1** and **F5** learned Korean with travelling and communication goals in mind.

*The participants' difficulties in studying Korean.* There are several difficulties that participants face when learning Korean, including memorizing vocabulary, challenging pronunciation, lacking opportunities to practice speaking, and lacking confidence in discussing various topics. Confidence in conversation and challenging vocabulary were among the most common struggles. **K2** recalled "*I'm afraid of not being understood or having to repeat myself and then making mistakes. And this fear stops me from speaking mostly. And then I take a long time to think. So yeah, this stops me from speaking.*" In addition, **F5** mentioned culture as one factor that brought friction to their Korean learning, "*This is my first Asian language. So how the culture is included in the language was a surprise to me. The Korean way of addressing people is different. The politeness forms and how to say people's names and so on. It was a surprise to me.*" These motivations, goals, and learning difficulties influenced what features and aspects of the AI chatbot each participant drew attention to, together with how they were encouraged to utilize it.

*The learners' perception of the chatbot's advantages.* Generally, participants found AI-powered conversations helpful in different ways, some more than the others, which affected their consideration of adding the chatbot tool into their learning toolbox and schedule. **K1**, **F1** and **F4** gained more confidence in speaking Korean, while **K2**, **K3**, and **F2** enhanced their fluency, responsiveness and conversational skills through practicing with AI. **F4** reflected "*I would get more courage to actually speak it, because, of course, you are, in a way, safe to make mistakes.*" **K3** and **F5** could improve their grammar and vocabulary, as well as pick up natural expressions in Korean. **F5** shared "*I learned new things in every conversation. So it was good that I could redo my answers and fix the grammar and the vocabulary that I used.*" For **F3**, the chatbot helped with the practice of pronunciation and speaking, together with new vocabulary acquisition. Overall, **K2**, **K3**, **K4**, **F4**, and **F5** were satisfied with the experience and benefits that the AI chatbot brought compared to their other available learning methods. On the other hand, **K1**, **F1**, **F2**, and **F3** had a more critical view on this learning tool.

*The chatbot's conversation flow.* Considering the interaction and supporting features available during the AI-powered conversations, participants pointed out several factors that encouraged them to keep engaging with the chatbot. **F1** felt motivated to maintain the discussion with chatbot when it made suggestions relevant to the topic at hand without being asked by the participant. In a similar way, the chatbot encouraged **K2**, **K3** and **F2** to engage in the conversations by continuously asking follow-up questions based on the participant's response. **F2** indicated *"if the chatbot is asking me a question, of course it motivates me to, okay, now I have to answer the question. Let's do it that way. Let's see what comes next."* **K4**, on the other hand, found the application's proposed topics and missions given during the conversation stimulating, as they could practice various areas of Korean which they rarely used in their class or daily life.

*The chatbot's interaction quality.* Moreover, the interactive and eloquent aspect of the AI-enhanced application makes it more intriguing and stimulating compared to various Korean language pedagogy tools. **K1** considered Univerbal one of the best in terms of interaction and natural language, compared to Duolingo, Memrise and similar Korean learning applications available on the market. **K3** stated *"The app that I used for this project, I guess, was much more interesting than something like the apps that just practice grammar with filling in the blank and such. I found that this app, speaking with an AI instructor, was very interesting in how it made me want to be more active and creative in what I was writing."* **K1** and **F4** preferred this application to ChatGPT, as ChatGPT was sometimes confusing while using too long sentences and complicated language. In contrast to Duolingo, **K2**, **F2**, and **F5** found Univerbal more logical, useful, and fluent. Especially for **F2**, *"it's more enjoyable than a tool, that, let's say Duolingo-type tool, that tells you exactly this way. And if you don't do it this way, then it's a mistake."* Additionally, **F5** recalled that Univerbal could utilize both formal and informal Korean speeches, unlike Google Translator which normally uses the informal language.

*The chatbot's human factors.* The AI chatbot's human factors and human-like levels also enormously affected participants' enjoyment and willingness to continue conversing with AI. Eight out of nine participants were generally satisfied with the chatbot's voice. *"I think it was very good, and I had a female and a male as well, like multiple different people, and you can also choose who the person is but I think it was very good and very understandable. The accent and the pronunciation were very good, and native-like."* (**F4**) Regarding the AI voice, being understandable is the most crucial factor to most participants. In addition, **K2** and **F4**

appreciated that the voice was slightly robotic but human enough. Regarding how human-like AI was, **K2** remarked *“For the voice itself, not too much, but on one hand, I also liked that. Maybe if it felt too human, I would have felt weird. I don't know. So I liked that. And then the conversation felt human.”* **F2** also considered AI's voice as acceptable, but emphasized the lack of other human elements in the AI character, such as name, age, and facial expressions, as the most disengaging point of the conversations. On the matter of how human-like the chatbot should be, **F4** and **K2** appreciated both its human and inhuman aspects. For **K2**, *“it was a perfect mix of human and not too human”*, while for **F4**, *“it tries to be human, but it doesn't try to be human in a way. And I think it's very good, it doesn't concern me too much. I think it's human-like enough for me to actually keep on the conversation.”* For **K3**, it was the structure of AI's responses that made it robotic. Specifically, the formats AI used in its answers were repetitive, limited and *“exactly felt like a teacher talking to a child, which also feels unnatural.”* The participant commented that it would be more enjoyable if the chatbot used a variety of sentence structures to engage in conversations.

*Discouraging elements.* Nevertheless, there were elements that discouraged learners from conversing with the chatbot. For instance, the chatbot asking the same questions in circles made the conversations tedious and caused **F1** to quit. Moreover, certain glitches and bugs, such as a sudden switch from Korean to English or French, a reverse in AI's roleplay, an inconsistency in AI's use of formal and informal languages, AI's misunderstanding of users' speeches and inaccurate suggestions in terms of Korean grammar and vocabulary, resulted in participants' confusion and frustration, which sometimes led to their exit from the conversation. **F3** pointed out two occasions where such a blip occurred: *“So I was having a conversation with the app, the AI of the app forgot their role and suddenly they were speaking as if our roles were reversed. So it felt a bit shady in a way.”* **K1** had a serious incident where they got stuck in a never-ending conversation with the chatbot for an hour, which only ended when **K1** closed the application. The experience was described by **K1** as *“It's very exhausting.”* Furthermore, the constant feedback and corrections given by the chatbot could be burdensome. **F5** shared *“sometimes when I make a lot of mistakes, it's overwhelming when every single sentence that I say is like, hey, here's a correction for you, and I'm like, again? Sometimes the correction is actually the thing that I was either trying to say or said, but the AI didn't pick up.”* Regarding the same subject, **F3** found it discouraging when AI gave unhelpful feedback, identifying trivial details instead of focusing on more meaningful corrections.

*The participants' willingness to continue using the chatbot.* During the study, the participants had three to eleven conversations out of seven days using the AI-powered application. Each conversation generally lasted 5 to 10 minutes, while some conversations could go up to 15 to 20 minutes. After this 7-day period, *eight out of nine participants* were willing to continue using the application for their Korean study. Specifically, five wanted to continue exploring the application and engaging in conversations with the AI chatbot at its current state; two were keen on taking advantage of the application if the subscription were cheaper or free; and one participant would be happy to pay for the subscription when the conversation quality was improved. The reason one participant was not interested in using the AI-enhanced tool was because they had limited time dedicated to Korean and better methods to learn the language at hand, such as Korean classes and one-on-one practice sessions with a native teacher. They stated, *"If I didn't have any other conversation partners to learn Korean with, then I would use this chatbot, I think."* Most participants would interact with the application at home, by themselves to practice conversations, while one participant (**K4**) would additionally use it in their Korean class to find information related to grammar and vocabulary.

*The participants' opinions regarding their Korean learning motivations.* While some participants thought motivation to learn Korean would come elsewhere and not from the application itself, some others did find the AI application a stimulation in their Korean study journey. **K2**, **F4** and **F5** were motivated to continue studying Korean by the application. **F4** stated *"It is very fun and as I said, it's very human-like and you can do it whenever you need to, whenever you have time and wherever you would just have your phone with yourself. So I think it's very useful and convenient, especially."* **F1** reflected that the AI chatbot could possibly encourage them to study Korean if used for a long period of time. On the contrary, **K1**, **F2** and **F3** commented that the AI application could be a supporting means to study Korean, but not the source of motivation. **F2** indicated *"It's a tool that gives you additional help on your learning journey. But the motivation itself needs to come somewhere else."*

Concisely, various factors influence participants' willingness and motivations to learn Korean through AI-enhanced free-form conversations. These factors include interaction quality, conversation contents, feedback and correction quality and frequency, AI's human-like level, together with severity of glitches and bugs. There were mixed opinions whether the AI chatbot encouraged learners to study Korean itself. Whether willing to continue using the application on their learning journey or not, all participants agreed that this type of AI-enhanced conversational

tool is beneficial in some ways and could be an assisting application to support their study. Additionally, there is no evident difference between the opinions of learners living in Finland and those living in South Korea.

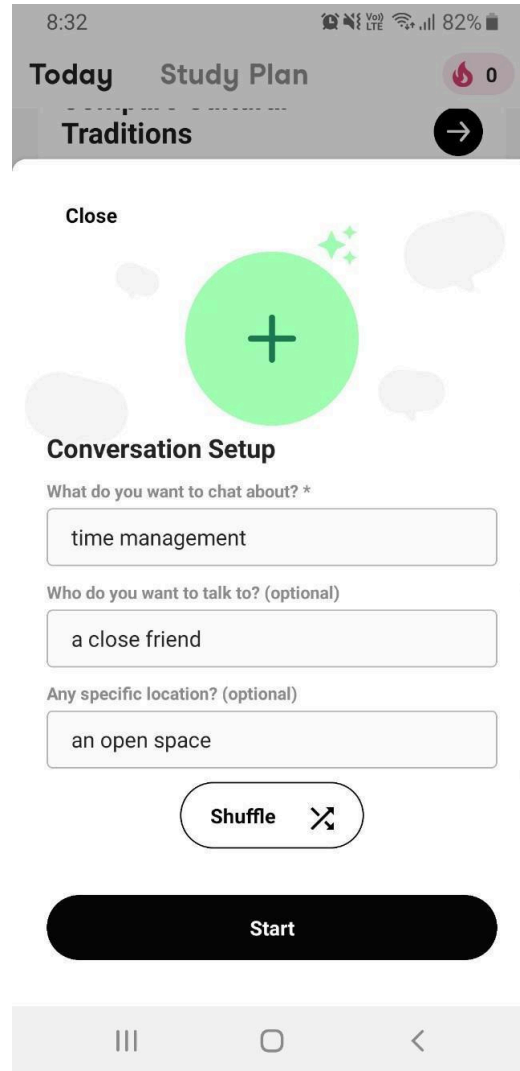
## 5 Design Implications for Korean Language Education AI Chatbots

Based on the research findings, this chapter presents six design implications for considerations in future design and implementation of AI-enhanced chatbots specifically for Korean language pedagogy.

### 5.1 Affordance

It is crucial to provide suitable affordances to help users recognize and understand how to use both main and supporting features during the conversations. Specifically, AI chatbots can demonstrate the availability of voice-based and text-based communication by using both voice and text as default to start a new conversation. This way, users are notified of the available modes to engage in the discussion. Moreover, as they are not always obvious for every user, supporting features such as feedback, and conversation transcript review should be introduced to users through an onboarding mechanism which does not obstruct the discussion flow, yet noticeable enough for users to be acknowledged.

Better affordance should also be considered in discussion topic settings. It is important that users are acknowledged about what they can decide about the topics, including what, where, and to whom the discussion will occur. Additionally, options specifically about the Korean language such as type of speech and pronoun should be available (refer to [Section 5.3](#) and [Section 5.5](#)). *Figure 8* shows the current user interface of the Create Your Own Topic feature, in which some participants during the empirical phase either didn't know that the options were editable or didn't clearly understand what it meant by "location".



**Figure 8.** Univerbal's Create Your Own Topic setup

## 5.2 Feedback and Corrections

The quality and frequency of corrections given to learners during conversations with AI have a significant impact on their experience. Receiving constant feedback can be overwhelming and frustrating, especially if the feedback is repetitive and focuses on trivial errors such as mispronunciation and missing optional elements of sentences. Providing corrections and suggestions for every single response can also break the flow of the conversation. Indeed, when feedback is presented only in written format, users have to stop their flow of thoughts to check the feedback, while AI waits for them to return to the conversation. On the other hand,

when comments are polite, accurate, and given at a proper rate, users find them useful and get motivated to continue learning.

As a result, it is important to maintain the conversation flow, while concentrating on more meaningful and constructive feedback. Specifically, to keep the conversation smooth and natural, feedback can sometimes be integrated into the chatbot's replies and given in both written and spoken formats. This way, users can return to the discussion transcript to check the feedback, but also listen and comprehend the feedback during the conversation. Since each user has a different learning style and purpose, giving them the option to select how frequently they wish to receive corrections and what types of comments they wish to review can be beneficial. For instance, a user who focused on practicing their fluency would care more about enriching their vocabulary and expressions rather than every typo or mispronunciation they make. In this case, an auto-correction feature would be efficient, which can still provide learners with the accurate spelling in a less intervening form.

### **5.3 Discussion Topics**

An unfit topic can discourage and hinder learners from participating in conversations with AI. While a low-level topic is dull and inefficient to improve the users' language skills, a difficult topic can be challenging and cause users to disengage from the conversations. Some users can also be demotivated by topics outside of their interests, while some others have difficulty coming up with their own discussion topics. Moreover, it is also frustrating to be forced to engage in conversation themes in a certain order, where users cannot skip what they are not keen on practicing.

Consequently, it is favorable to provide users with a wide range of theme suggestions, including the option to create their own topics. By giving learners various lists of topics divided by slightly different proficiency levels, AI applications can provide them with the opportunity to work on suitable topics while encouraging them to explore unfamiliar themes whenever they feel ready. Moreover, there should be both original, imaginative topics (such as talking about the meaning of life with Descartes on university campus) and familiar but practical ones (such as talking about travelling with a friend at a coffee shop). For the option of setting one's own discussion, there

could be additional selections for type of speech (informal 반말, formal 존댓말, honorifics 높임말) and conversation length (limited or unlimited).

## 5.4 Conversation Interaction

Various response structures and diverse vocabulary can make the conversation more natural, while motivating language students to express themselves and engage in the discussion topic. Besides Yes-No and What questions, by inquiring How and Why, AI chatbots can motivate learners to use more complex sentences and a wide range of vocabulary to elaborate their answers. Additionally, instead of always following up with a new question, AI chatbots can occasionally leave rooms for learners to ask their own questions and develop the conversation in the direction they wish.

Moreover, to aid learners in extending the conversation, it is useful for AI chatbots to recognize hindrance and voluntarily provide support. For instance, users' long pauses can be a signal for chatbots to rephrase the questions or offer vocabulary hints. Additionally, users' sudden exit from the conversation can arise from challenging or tedious topics. In this case, it can be helpful for learners to be provided with a list of relevant grammar and expressions to practice on, instead of having to come up with the material themselves.

Integrating feedback and new expressions into the conversations can also be advantageous. Pausing to check each correction in the middle of a discussion can break the conversation flow. To prevent this intervention, AI chatbots can combine critical corrections and introduce new expressions in their response. Learners are then also given the chance to request further explanations for the new vocabulary or grammar points if necessary.

## 5.5 Chatbot Character Design

Due to the nature of the Korean language, it is critical for the chatbot to have a name (for instance, Descarte 데카르트, Eunkyeong 은경...) or a pronoun (for example, Grandfather 할아버지, CEO 대표님...) to ensure learners know how to address the AI counterpart during the conversations. This detail can be vocally introduced to the users by the chatbot at the start of the conversation or presented in the chatbot's name tag. An AI-generated profile picture suited

for each conversation character would also be beneficial in creating a more natural atmosphere and help users acknowledge the proper way to address the AI chatbot during the conversation.

While a certain level of chatbots' human-like characteristics has a positive effect on the conversation flow, overly embellished anthropomorphism can make learners uncomfortable and the interaction awkward. These characteristics include the chatbots' voice, backstory, and response formats. Consequently, it is beneficial to balance these factors and focus on the human-like nature of the conversation flow, instead of building an overly sophisticated chatbot character.

## **5.6 Chatbots' Trustworthiness**

Various factors during the interaction with chatbots can influence how much users trust the AI-generated contents. These factors include the accuracy of the information and feedback provided by chatbots, the frequency and intensity of bugs and glitches, as well as the conversation contents. For example, when chatbots give incorrect information that otherwise would be easily confirmed by simple internet search, or when chatbots suddenly speak in a different language, users find them less trustworthy and, in some cases, feel discouraged to continue the conversation.

Consequently, it is important to minimize AI chatbots' technical issues. When bugs occur, users should be provided with a way to move on from the incident, whether by continuing or exiting the conversation. Allowing learners to report bug incidents would also contribute to the improvement of chatbot quality. Regarding conversation contents, non-repetitive responses and diverse sentence structures make chatbots more authentic and the conversation flow more natural. Additionally, keeping the chatbots' data pool up-to-date makes them more reliable. It would also be helpful to clarify in the application the data source and year, so as users are acknowledged about the relevance and accuracy level of information provided by chatbots.

## 6 Discussion

This chapter covers the research discussion, the limitations of the study, and possible future work.

### 6.1 Discussion on RQ1

During the empirical phase of the current research, nine Korean language learners in Finland and South Korea participated in the research, engaging with an AI chatbot in three steps. First, they completed a questionnaire on their background and expectations for AI in Korean language learning. Next, they experienced the AI-enhanced app Universal independently for seven days, recording their feedback in journals. Finally, they shared their experiences in semi-structured interviews.

Pertaining to **RQ1**: *What are the learners' expectations of conversational AI as a tool to improve their Korean language skills?*, the study findings show that learners have little and low expectations, due to their prior perception of AI. There was no evident difference in expectations between participants residing in South Korea and those living in Finland; both groups demonstrated similarly modest expectations, alongside preconceived notions regarding AI-powered tools. Furthermore, both cohorts expressed *shared opinions on the potential applications of AI in their Korean language education, including enhancing vocabulary, grammar, pronunciation, fluency, and speaking confidence, as well as facilitating access to language-related information*. However, the participants' assessments of AI's usefulness, the role of human factors, and their comfort levels with AI interaction displayed considerable variation along a 1-7 scale.

The learners' little and low expectations towards AI-driven Korean education chatbots might be due to their prior perception of AI. For learners, the notion of AI means unnatural conversations and awkwardness due to AI being either too human or too robotic. Indeed, the level at which language students perceive chatbots' anthropomorphism can positively or negatively influence their emotions towards these AI-powered tools (Gyeong & Chang, 2025). In terms of anticipation about their comfortability interacting with AI, findings show varied opinions among Korean learners. The anticipated uncomfortable interaction with AI chatbots might have resulted

from their previous unsuccessful experience using AI, whether it was to learn languages or to serve other purposes (participants **K3**, **F3**). Despite the overall low expectation, students still expect AI chatbots to be able to assist them in certain aspects of Korean language learning, such as pronunciation, vocabulary, grammar, fluency, and communication confidence. This is in line with the perception of AI chatbots and their benefits by English language learners across several studies (Dahbi, 2023; Polakova & Klimova, 2024; Yang et al., 2022).

## 6.2 Discussion on RQ2

Concerning **RQ2**: *What kind of guidance the learners need in free-form conversational practice with AI?*, the experiences of the nine participants varied significantly. These participants faced different obstacles and valued distinct support features embedded within the AI-enhanced learning process. Notable concerns included *the absence of the chatbot's name and age, inappropriate conversation topics, issues with conversation flow and structure, inadequate vocabulary support, suboptimal feedback and corrections from the AI, and the unclear presentation of the available features within the application*. Opinions on the use of free-form AI conversations for Korean language learning were also divided, with some participants appreciating this format, while others favored a more structured, guided approach. In addition, no substantial differences were observed between the experiences of participants in Finland and those in South Korea, although learners in Finland expressed greater concern regarding the appropriate form of address when interacting with the AI chatbot.

These findings from the current research provide notable remarks. Since several features of the AI-powered application that support conversations with chatbots were not obvious to learners, it is plausible to say that certain onboarding and usage instructions are necessary to ensure users can utilize chatbots to their benefit. This suggestion is in accordance with various existing research on language education using chatbots, where students were interacting with chatbots under the guidance and instructions of either human teachers or research moderators (Divekar et al., 2022; Kim & Su, 2024; Yang et al., 2022). Even in Polakova and Klimova (2024) study where participants used a chatbot independently multiple times outside of the classroom environment, they interacted with the chatbot in class during the first session to familiarize themselves with the tool and raise questions regarding its usage.

Furthermore, the current research reveals that conversation topics matching the learner's proficiency level coupled with vocabulary suggestions and sample expressions is essential, especially for those with lower levels of Korean. This view is shared in Xia et al. (2023) study on 323 middle schoolers' English learning experience with chatbot in Hong Kong SAR. Specifically, the research indicated that while advanced students are more capable of generating ideas and modulating their interaction with chatbots, lower proficiency level learners would need assistance, including comprehensive instructions, conversation samples, and expression recommendations.

Another notion regarding conversation guidance is that real-time interaction and feedback are helpful and highly appreciated by Korean learners. This finding is in line with Wiboolyasarin et al. (2024) study on elements influencing the design of Thai language-teaching chatbots. While Wiboolyasarin and their team did not mention details on how feedback should be designed and integrated into the interaction with chatbots, the current research suggests that comments from chatbots should focus on meaningful grammar and vocabulary points, while occurring at a reasonable frequency so as not to overwhelm learners and obstruct the flow of the conversations. When considering the nature of the Korean language, the current research provides insight into the importance of a chatbot character with a pronoun and age indication in supporting a straight-forward and natural conversation practice. This factor has not been explored in existing literature, hence can be useful in future research and design of AI chatbots for Korean language pedagogy.

### **6.3 Discussion on RQ3**

With regards to **RQ3: *What factors affect the learners' motivation and willingness in using AI free-form conversations to improve their Korean language skills?***, multiple factors influenced participants' willingness and motivation to engage in Korean language learning through AI-driven free-form conversations. These factors included *the quality of interactions, the relevance of conversation content, the frequency and quality of feedback and corrections, the AI's human-like characteristics, and the extent of technical glitches or malfunctions*. Participants' opinions on whether the AI chatbot itself effectively motivated them to study Korean varied. Nevertheless, regardless of their willingness to continue using the application after the 7-day period, all participants agreed that AI-assisted conversational tools could offer valuable support

for language learning. Additionally, there was no apparent difference in the opinions of learners residing in Finland compared to those in South Korea.

Among various elements influencing learners' drive and engagement, AI chatbots' interaction quality is crucial in capturing the learners' attention and willingness to converse in Korean. Particularly, chatbots' questions and replies should be natural and instantaneous, while users' utterances should be accurately understood by chatbots (Polakova & Klimova, 2024; Yang et al., 2022). In addition, instead of a flawless human mimicry, certain non-human elements can reduce students' social anxiety of making errors hence motivating them to practice their targeted language (Annamalai et al., 2023; Divekar et al., 2022). The present research findings indicate that learners do not consider AI chatbots as a replacement for human-human interaction, but rather a supplemental learning tool to help them improve their Korean skills and get ready for future communications with Korean native speakers. As a result, learners accept the chatbot's non-human factors such as its robotic voice and are more concerned with the content and natural flow of the conversations. This topic of AI-human vs human-human interaction has been systematically reviewed in Jeon and Lee (2024) paper. Here, the authors suggested that both human-like and mechanical aspects of chatbots can be positively perceived by students, and recommended considering chatbots as unique conversational partners, rather than attempting to build them as human replicas.

Moreover, suitable and interesting topics together with various question structures keep the AI-powered conversations engaging. Captivating conversation contents help learners enjoy practicing the targeted language and acquire new knowledge, while monotonous inquiries from chatbots make the experience tedious (Polakova & Klimova, 2024). On the other hand, technical issues, including both chatbots' mistakes and user interface design, arising during the interaction negatively affect users' learning experience hence discouraging them from continuing the interaction with AI chatbots. Recurrent technical problems and the chatbot's errors result in learners' frustration and negative perception of the AI-enhanced tool (Polakova & Klimova, 2024).

## 6.4 Discussion on Design Implications

The current research findings resulted in six design implications for the future development and implementation of AI-powered chatbots for Korean language education. Three of these considerations, namely affordances, feedback and conversation topics, have been discussed by Wiboolyasarin et al. (2024) regarding chatbots for Thai language pedagogy and extended by the current study. Specifically, AI-driven applications should be easy to use for learners of different proficiency levels. It is also beneficial for AI chatbots' feedback to be instant, accurate while remaining positive to help students relax and maintain their confidence in practicing their target language. Additionally, stimulating conversation contents and various response formats increase learners' interest and engagement (Wiboolyasarin et al., 2024).

As AI chatbots' level of trustworthiness affects how much users accept these technologically advanced applications, maximizing the users' trust would increase their comfort and enhance their interaction with AI (Guo et al., 2022). In Guo et al. (2022, p476) study, "*five design semantics categories including controllability, transparency, intelligence, intimacy and engagement*" were proposed to guide the design of reliable chatbots. The current research provides design suggestions in accordance with these semantics, such as AI's ability to give polite quality comments (Intimacy), as well as provide users with the options to create their own conversation content (Engagement). However, while the need to visualize information security and show chatbots' understanding towards users based on users' personal data were presented in Guo et al. (2022) research, none of the participants in the current study expressed any anxiety regarding their privacy during their interaction with AI chatbots. Indeed, their trust towards chatbots concerns more with the quality of AI-generated contents and the conversation authenticity level.

In addition, this master's thesis provides implications regarding AI chatbot character design unique to Korean language learning. Specifically, while original conversation topics can motivate learners to practice their targeted language and gain new knowledge (Polakova & Klimova, 2024), sometimes these topics come with unusual AI conversation partners (for example, a dragon, Tarzan...) that make students perplex how to address the character and what level of politeness they should apply in their Korean speech. As a result, it is critical in such cases to provide students with extra information such as the chatbot character's name, pronoun, and age to ensure a smooth and natural conversation.

## 6.5 Limitations and Future Work

The current study's limitations include the utilization of readily accessible Univerbal during the empirical phase, the research time constraint, as well as the focus on AI chatbots in Korean language education without any generalization for other languages. The number of research participants, along with their demographic information and residing location were also limited in this thesis.

The use of the readily available AI-enhanced mobile application Univerbal during the research empirical phase might have affected the participants' experience hence the findings of the study. Specifically, since the application was not solely developed for Korean language education, the nature of the Korean language regarding pronouns and types of speeches was not thoroughly considered. The original topic suggestions, such as discussing philosophy with Decartes and paintings with Leonardo da Vinci, were also more relevant to western languages and cultures. Furthermore, during the empirical phase, the application was updated once with changes in the user interface and available conversation topics. This system upgrade might have influenced the different experiences of participants engaging with the old version (**K1**, **F1**) versus those interacting with the new version.

The seven-day time constraint for participants to experience conversations with the AI chatbot could have influenced their exploration of the mobile application, along with their experience and motivations regarding the use of this AI-enhanced tool. Indeed, during the interview, some participants mentioned their lack of knowledge about available features in the application and expressed their willingness to explore and learn to use the tool to their advantage. Additionally, all learners commented that one week was a relatively short time to come to any concrete conclusions about their study progress. As a result, in future studies, it would be beneficial to follow and study users' experience and motivation in studying Korean using AI-powered applications over a longer period of time.

There was also a limitation regarding the research participants' demographic information and living location. Given that age, gender, and nationality can be regarded as sensitive data, and since the primary focus of this study was on learners' interactions with an AI-enhanced chatbot,

these demographic details were neither collected nor specifically addressed. Moreover, to extend the participant pool and ensure a more comprehensive dataset, individuals living in both South Korea and Finland, with different backgrounds and different mother tongues, whose first language is not Korean, were recruited. This expanded recruitment strategy aimed to increase the overall number of participants, thus strengthening the research findings.

Furthermore, the current research relied solely on qualitative data analysis, focusing on non-numerical information to derive insights. As there were only nine learners of various age ranges and Korean proficiency levels participating in the experiment, the study did not provide any qualitative conclusions regarding any specific Korean learner demographic groups. Nevertheless, small sample sizes are often adequate in qualitative research when the objective is to deeply explore experiences and perspectives, rather than to generalize the findings to a broader population (Tenny et al., 2017). In addition, this thesis work did not include quantitative measures to assess success or failure rates, limiting the ability to provide measurable conclusions. Future research could tackle this limitation by including quantitative data collection and analysis, which would enhance the findings and offer a more comprehensive understanding through statistical evidence.

Additionally, this thesis only concentrates on the experience of Korean language learners, without generalizing any of its design implications for the education of other languages. Due to the focus on Korean language education during the empirical phase, there was not enough evidence to suggest whether the design implications regarding Korean education AI chatbots in this thesis work could be applied to other languages. Whether these design suggestions for Korean language pedagogy could be generalized to other languages would require further investigation in future research.

## 7 Conclusion

The current research aims to understand Korean learners' expectation, experience and motivations in using AI-driven chatbots to support their Korean language education. The findings show that learners hold low expectations regarding the usefulness of conversational AI in supporting their Korean language studies, with no significant differences between those residing in South Korea and Finland. Both groups identified potential benefits, such as improving vocabulary, grammar, pronunciation, fluency, and speaking confidence, as well as accessing language-related information. However, their evaluations of AI chatbot's usefulness, comfort levels, and perceptions of human factors varied. Challenges during interactions with AI-powered chatbots highlight the need for additional guidance in several areas. These included the unclear identity of the chatbot, unsuited conversation topics, disruptions in conversation flow, insufficient vocabulary support, chatbot's technical issues, and inefficient feedback. Factors influencing learners' motivations and willingness to practice Korean with AI-powered chatbots include interaction quality, relevant content, feedback frequency, human-like AI characteristics, and technical issues.

The research findings contribute to existing literature by providing a closer look at how learners' interaction with AI chatbots in Korean resembles and differs from the experience in other languages. The current paper discovery is similar to data in existing research regarding learners' engagement thanks to real-time feedback, original conversation topics, as well as their perceived benefits of chatbots. On the other hand, it provides additional values such as how to integrate feedback into conversations and how to design discussions that are effective considering the nature of the Korean language. It is also noteworthy that in terms of AI chatbots' anthropomorphism, the Korean language learners in this study prioritize the human-like quality of the conversation flow but appreciate certain AI's non-human factors. Furthermore, they do not consider chatbots as an alternative to human-human interaction, but rather an assisting application to prepare them for conversations with Korean native speakers in the future.

Based on learners' experience and feedback during the research experiment, this master's thesis provides six design implications for future design and realization of AI chatbots as a tool for Korean language conversation practice. These implications include affordances, feedback and corrections, discussion topics, conversation interaction, chatbot character design, and

chatbots' trustworthiness. For future research, it would be beneficial to study how Korean language learners' interaction and motivation to practice Korean with AI chatbots evolve over a long period of time. Quantitative data and analysis would also strengthen the findings and provide a deeper insight in this topic.

## References

Ali, Z. (2020). Artificial Intelligence (AI): A Review of its Uses in Language Teaching and Learning. *IOP Conference Series. Materials Science and Engineering*, 769(1), 12043-. <https://doi.org/10.1088/1757-899X/769/1/012043>

Annamalai, N., Rashid, R. A., Munir Hashmi, U., Mohamed, M., Harb Alqaryouti, M., & Eddin Sadeq, A. (2023). Using chatbots for English language learning in higher education. *Computers and Education. Artificial Intelligence*, 5, 100153-. <https://doi.org/10.1016/j.caeai.2023.100153>

Berghe, R., Haas, M., Oudgenoeg-Paz, O., Krahmer, E., Verhagen, J., Vogt, P., Willemsen, B., Wit, J., & Leseman, P. (2021). A toy or a friend? Children's anthropomorphic beliefs about robots and how these relate to second-language word learning. *Journal of Computer Assisted Learning*, 37(2), 396–410. <https://doi.org/10.1111/jcal.12497>

Beyer, H., & Holtzblatt, K. (1998). *Contextual design: defining customer-centered systems*. Morgan Kaufmann.

Bleecker, I. D., & Okoroji, R. (2018). *Remote usability testing: actionable insights in user behavior across geographies and time zones* (1st edition). Packt Publishing Ltd.

Carson, L., & Do, E. (2013). Establishing a Korean language programme in a European Higher Education context: Rationale, curriculum and assessment procedures. *Language Learning in Higher Education (Berlin, Germany)*, 3(1), 151–171. <https://doi.org/10.1515/cercles-2013-0008>

Ciechanowski, L., Przegalinska, A., Magnuski, M., & Gloor, P. (2019). In the shades of the uncanny valley: An experimental study of human–chatbot interaction. *Future Generation Computer Systems*, 92, 539–548. <https://doi.org/10.1016/j.future.2018.01.055>

Clarizia, F., Colace, F., Lombardi, M., Pascale, F., & Santaniello, D. (2018). Chatbot: An Education Support System for Student. *Cyberspace Safety and Security*, 291–302. [https://doi.org/10.1007/978-3-030-01689-0\\_23](https://doi.org/10.1007/978-3-030-01689-0_23)

Dahbi, M. (2023). Integrating an Intelligent Language Tutoring System in Teaching English Grammar. *Arab World English Journal*, 14(4), 189–196. <https://doi.org/10.24093/awej/vol14no4.11>

Divekar, R. R., Drozdal, J., Chabot, S., Zhou, Y., Su, H., Chen, Y., Zhu, H., Hendler, J. A., & Braasch, J. (2022). Foreign language acquisition via artificial intelligence and extended reality: design and evaluation. *Computer Assisted Language Learning*, 35(9), 2332–2360. <https://doi.org/10.1080/09588221.2021.1879162>

Dizon, G. (2017). Using Intelligent Personal Assistants for Second Language Learning: A Case Study of Alexa. *TESOL Journal*, 8(4), 811–830. <https://doi.org/10.1002/tesj.353>

Dong Bae Lee. (2014). Motivations of Learning Korean and Their Influence on Cultural Content: Korean (Popular/K-pop) Culture for Beginner Korean Learners. *Kugŏ Kyoyukhak Yŏn'gu*, 49(4), 191–218. <https://doi.org/10.20880/kler.2014.49.4.191>

Gökçearsan, Ş., Tosun, C., & Erdemir, Z. G. (2024). Benefits, Challenges, and Methods of Artificial Intelligence (AI) Chatbots in Education: A Systematic Literature Review. *International Journal of Technology in Education*, 7(1), 19–39. <https://doi.org/10.46328/ijte.600>

Guo, Y., Wang, J., Wu, R., Li, Z., & Sun, L. (2022). Designing for trust: a set of design principles to increase trust in chatbot. *CCF Transactions on Pervasive Computing and Interaction (Online)*, 4(4), 474–481. <https://doi.org/10.1007/s42486-022-00106-5>

Gyeong Kim, M., & Chang Lee, K. (2025). Proposing the “Digital Agenticity Theory” to analyze user engagement in conversational AI chatbot. *Journal of Business Research*, 189, 115162-. <https://doi.org/10.1016/j.jbusres.2024.115162>

Han, Y. (2023). Motivations for Learning Korean in Vietnam: L2 Selves and Regulatory Focus Perspectives. *Journal of Language, Identity, and Education*, 22(6), 559–573. <https://doi.org/10.1080/15348458.2021.1935961>

Haristiani, N. (2019). Artificial Intelligence (AI) Chatbot as Language Learning Medium: An inquiry. *Journal of Physics. Conference Series*, 1387(1), 12020-. <https://doi.org/10.1088/1742-6596/1387/1/012020>

Hoy, M. B. (2018). Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. *Medical Reference Services Quarterly*, 37(1), 81–88. <https://doi.org/10.1080/02763869.2018.1404391>

Hsu, H.-L., Chen, H. H.-J., & Todd, A. G. (2023). Investigating the impact of the Amazon Alexa on the development of L2 listening and speaking skills. *Interactive Learning Environments, ahead-of-print*(ahead-of-print), 1–14. <https://doi.org/10.1080/10494820.2021.2016864>

Jeon, J. (2024). Exploring AI chatbot affordances in the EFL classroom: young learners' experiences and perspectives. *Computer Assisted Language Learning*, 37(1–2), 1–26. <https://doi.org/10.1080/09588221.2021.2021241>

Jeon, J., & Lee, S. (2024). Can learners benefit from chatbots instead of humans? A systematic review of human-chatbot comparison research in language education. *Education and Information Technologies*, 29(17), 23329–23360. <https://doi.org/10.1007/s10639-024-12725-9>

Jeon, J., Lee, S., & Choe, H. (2023). Beyond ChatGPT: A conceptual framework and systematic review of speech-recognition chatbots for language learning. *Computers and Education*, 206, 104898-. <https://doi.org/10.1016/j.compedu.2023.104898>

Jose, J., & Jayaron Jose, B. (2024). Educators' Academic Insights on Artificial Intelligence: Challenges and Opportunities. *Electronic Journal of E-Learning*, 22(2), 59-. <https://doi.org/10.34190/ejel.21.5.3272>

Kallioniemi, P., Ronkainen, K., Karhu, J., Sharma, S., Hakulinen, J., & Turunen, M. (2019). *CityCompass VR - A Collaborative Virtual Language Learning Environment*.

Kim, A., & Su, Y. (2024). How implementing an AI chatbot impacts Korean as a foreign language learners' willingness to communicate in Korean. *System (Linköping)*, 122, 103256-. <https://doi.org/10.1016/j.system.2024.103256>

곽부모. (2019). Korean Language Education and Korean Studies in Central Europe : Focusing on Korean Studies Program for the Degree Students in the Czech Republic. *Kugŏ Kyoyukhak Yŏn'gu*, 54(5), 5–22. <https://doi.org/10.20880/kler.2019.54.5.5>

Lee, H. (2018). *Voice user interface projects : build voice-enabled applications using Dialogflow for Google Home and Alexa Skills Kit for Amazon Echo* (1st edition). Packt.

Lee, I. (2018). Effects of Contact with Korean Popular Culture on KFL Learners' Motivation. *The Korean Language in America*, 22(1), 25–45. <https://doi.org/10.5325/korelangamer.22.1.0025>

Lee, K.-A., & Lim, S.-B. (2023). Designing a Leveled Conversational Teachable Agent for English Language Learners. *Applied Sciences*, 13(11), 6541-. <https://doi.org/10.3390/app13116541>

Lee, S., & Jeon, J. (2024). Visualizing a disembodied agent: young EFL learners' perceptions of voice-controlled conversational agents as language partners. *Computer Assisted Language Learning*, 37(5–6), 1048–1073. <https://doi.org/10.1080/09588221.2022.2067182>

Lesia Viktorivna, K., Andrii Oleksandrovych, V., Iryna Oleksandrivna, K., & Nadia Oleksandrivna, K. (2022). Artificial Intelligence in Language Learning: What Are We Afraid of. *Arab World English Journal*, 8, 262–273. <https://doi.org/10.24093/awej/call8.18>

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson. <http://discovery.ucl.ac.uk/1475756/>

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal (Boulder, Colo.)*, 82(4), 545–562. <https://doi.org/10.1111/j.1540-4781.1998.tb05543.x>

Mackey, A., & Gass, S. M. (2005). *Second language research methodology and design*. Lawrence Erlbaum.

Mariani, M. M., & Borghi, M. (2024). Artificial intelligence in service industries: customers'

assessment of service production and resilient service operations. *International Journal of Production Research*, 62(15), 5400–5416. <https://doi.org/10.1080/00207543.2022.2160027>

Mishra, D., Romero, G. A., Pande, A., Nachenahalli Bhuthegowda, B., Chaskopoulos, D., & Shrestha, B. (2024). An Exploration of the Pepper Robot's Capabilities: Unveiling Its Potential. *Applied Sciences*, 14(1), 110-. <https://doi.org/10.3390/app14010110>

Nadia Ahmad, Hirok Chakraborty, & Ratnesh Sinha. (2024). Exploring medical student's outlook on use of artificial intelligence in medical education: a multicentric cross-sectional study in Jharkhand, India. *Cogent Education*, 11(1). <https://doi.org/10.1080/2331186X.2024.2396704>

Nikitina, L., & Furuoka, F. (2019). Language learners' mental images of Korea: insights for the teaching of culture in the language classroom. *Journal of Multilingual and Multicultural Development*, 40(9), 774–786. <https://doi.org/10.1080/01434632.2018.1561704>

Nikunlassi, A. O., Loginova, A. D., Ufimtsev, G. V., & Protassova, E. Y. (2021). Designing online educational Russian language courses in Finland. *Rusistika (Moskva. Online)*, 19(1), 66–84. <https://doi.org/10.22363/2618-8163-2021-19-1-66-84>

Norman, D. (2013). Human-Centered Design. In *The Design of Everyday Things*. Basic Books.

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

Okonkwo, C. W., & Ade-Ibijola, A. (2021). Python-bot: A chatbot for teaching python programming. *Engineering Letters*, 29(1), 25–34.

Papakostas, G. A., Sidiropoulos, G. K., Papadopoulou, C. I., Vrochidou, E., Kaburlasos, V. G., Papadopoulou, M. T., Holeva, V., Nikopoulou, V. A., & Dalivigkas, N. (2021). Social robots in special education: A systematic review. *Electronics (Basel)*, 10(12), 1398-. <https://doi.org/10.3390/electronics10121398>

Petra Polakova, & Blanka Klimova. (2024). Implementation of AI-driven technology into

education – a pilot study on the use of chatbots in foreign language learning. *Cogent Education*, 11(1). <https://doi.org/10.1080/2331186X.2024.2355385>

Ramesh, A. N., Kambhampati, C., Monson, J. R. T., & Drew, P. J. (2004). Artificial intelligence in medicine. *Annals of the Royal College of Surgeons of England*, 86(5), 334–338. <https://doi.org/10.1308/147870804290>

Salmela, R., Lehtonen, M., Koykka, J., Garusi, S., Laine, M., & Bertram, R. (2024). Training Finnish morphology with a smartphone application in adult beginner level learners. *Computer Assisted Language Learning*, 1–33. <https://doi.org/10.1080/09588221.2024.2365664>

Semerikov, S. O., Striuk, A. M., & Shalatska, H. M. (2021). AI-assisted language education: critical review. *Osvitnij Vimir (Online)*, 4, 1–7. <https://doi.org/10.31812/ed.623>

Sumakul, D. T. Y. G., Hamied, F. A., & Sukyadi, D. (2022). Artificial Intelligence in EFL Classrooms: Friend or Foe? *LEARN Journal: Language Education and Acquisition Research Network*, 15(1), 232-256.

Tai, T.-Y. (2024). Effects of intelligent personal assistants on EFL learners' oral proficiency outside the classroom. *Computer Assisted Language Learning*, 37(5–6), 1281–1310. <https://doi.org/10.1080/09588221.2022.2075013>

Tai, T.-Y., & Chen, H. H.-J. (2024). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions. *Computers and Education*, 220, 105112-. <https://doi.org/10.1016/j.compedu.2024.105112>

Tenny, S., Brannan, J. M., & Brannan, G. D. (2017). Qualitative study.

Tiffany Laiyin Lao. (2020). The Relationship Between ESL Learners' Motivation, Willingness to Communicate, Perceived Competence, and Frequency of L2 Use. *Studies in Applied Linguistics & TESOL*, 20(2).

Wang, K. (2024). From ELIZA to ChatGPT: A brief history of chatbots and their evolution. *Applied and Computational Engineering*, 39(1), 57–62.

<https://doi.org/10.54254/2755-2721/39/20230579>

Weizenbaum, J. (1983). ELIZA—A Computer Program For the Study of Natural Language Communication Between Man And Machine. In *Communications of the ACM* (Vol. 26, Number 1, pp. 23–28). Association for Computing Machinery. <https://doi.org/10.1145/357980.357991>

Wiboolyasarín, W., Wiboolyasarín, K., Tiranant, P., Boonyakitanont, P., & Jinowat, N. (2024). Designing chatbots in language classrooms: an empirical investigation from user learning experience. *Smart Learning Environments*, 11(1), 32–25. <https://doi.org/10.1186/s40561-024-00319-4>

Xia, Q., Chiu, T. K. F., Chai, C. S., & Xie, K. (2023). The mediating effects of needs satisfaction on the relationships between prior knowledge and self-regulated learning through artificial intelligence chatbot. *British Journal of Educational Technology*, 54(4), 967–986. <https://doi.org/10.1111/bjet.13305>

Xiao, Y., Zhang, T., & He, J. (2024). The promises and challenges of AI-based chatbots in language education through the lens of learner emotions. *Heliyon*, 10(18), e37238-. <https://doi.org/10.1016/j.heliyon.2024.e37238>

Yang, H., Kim, H., Lee, J. H., & Shin, D. (2022). Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. *ReCALL (Cambridge, England)*, 34(3), 327–343. <https://doi.org/10.1017/S0958344022000039>

ZEMČÍK, Mgr. T. (2019). A Brief History of Chatbots. *DEStech Transactions on Computer Science and Engineering, aicae*. <https://doi.org/10.12783/dtcse/aicae2019/31439>

# Appendices

## Appendix 1

### Important Links

- [Research Information Sheet](#)
- [Research Participation Consent Form](#)
- [Affinity Diagram of Research Findings](#)

## Appendix 2

### Pre-testing Questionnaire

1. 이름을 알려주세요. **Please let me know your first name.\***

---

2. 나이가 어떻게 되세요? **What is your age?\***

- 19세 미만 / Under 19 years old
- 19세 - 24세 / 19 - 24 years old
- 25세 - 34세 / 25 - 34 years old
- 35세 - 44세 / 35 - 44 years old
- 45세 - 64세 / 45 - 64 years old
- 64세 이상 / Over 64 years old

3. 한국어 수업 시간 외에 한국어 학습을 지원하기 위해 어떤 도구를 사용하시나요?\*

**What tool(s) do you use to support your Korean studies outside class?**

(해당하는 항목을 모두 선택해 주세요. *Please select all applicable answers.*)

- 웹사이트 또는 모바일 앱 / Websites and/or mobile applications
- 소셜 미디어 (인스타그램, 페이스북 등) / Social media (Instagram, Facebook,...)
- 한국 미디어 (음악, 영화, TV 프로그램, 드라마 등) / Korean media (music, movies, TV shows, dramas...)
- 원어민과의 대면 언어 교환 / Face-to-face language exchange with native speakers
- 아무것도 안함 / Nothing
- Other

언어학습 AI앱에 관한 질문 / **Questions regarding AI language learning apps**

전에 언어학습 AI앱을 사용해보신 분은 경험을 토대로 작성해주시고 사용경험이 없으면 자신이 생각하는대로 느낌을 작성해주세요. If you have used an AI language learning app before, please answer based on your experience. If you have no experience, please answer based on your opinions.

4. 주로 AI와 한국어 대화 연습을 어디서 하시겠습니까?\*

**Where would you practice Korean conversations with AI?**

(해당하는 항목을 모두 선택해 주세요. *Please select all applicable answers.*)

- 자택에서 혼자서 함 / At home, by myself
- 자택에서 가족이나 친구와 함 / At home, with friends or family members
- 공공장소 (카페, 한국어 행사 등)에서 함 / At a public place (a cafe, a Korean language event,...)
- 한국어 수업에서 학습 보조 도구로 함 / In Korean class, as an assisting learning tool
- Other

**5. AI 적용된 앱은 한국어 회화 능력에 얼마나 도움이 된다고 생각하시나요?\***

**How helpful do you think AI applications are for your Korean conversational skills?**

도움이 되지않음 / Not helpful      1 2 3 4 5 6 7      도움이 됨 / Very helpful

**6. AI와 대화해보면서 한국어 학습의 어떤 부분에 도움이 되었다고 생각하시나요?\***

**What aspects of your Korean skills do you think could be improved through conversations with AI?**

---



---

**7. AI 대화형 앱과 상호작용하는 것이 얼마나 편안하게 느끼시나요?\***

**How comfortable are you in interacting with AI conversational applications?**

매우 불편함 / Very uncomfortable      1 2 3 4 5 6 7      매우 편함 / Very comfortable

**8. 한국어 학습 대화에서 AI는 얼마나 인간 같아야 한다고 생각하시나요?\***

**How human-like do you think AI should be in a Korean practice conversation?**

인간 같지 않음 / Not human-like      1 2 3 4 5 6 7      인간 같음 / Very human-like

**9. 학습 목적으로 AI와 소통하는 것에 대한 경험 또는 개선 되었음 하는 다른 의견이 있습니까?  
Any other feedback regarding your experience and/or expectation about communicating with AI for learning purposes?**

---



---

## Appendix 3

### 7-Day AI Conversation Practice Instruction

During 7 days, use the mobile application to practice conversations in Korean as much as you wish, at any time and in any way you wish.

However, please include at least one session of free-topic conversation with AI (detailed instruction in step 3)

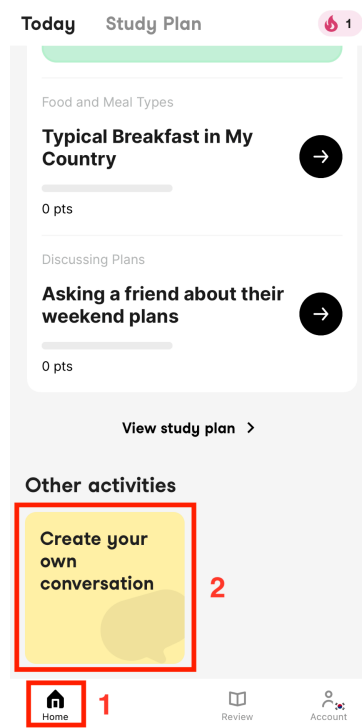
#### Step 1: Download the Univerbal app for [Android](#) or [iOS](#) devices

#### Step 2: Sign up for a free 7-day trial account

- 1- To sign up, you need a valid email address.
- 2- Once you sign up, a verification code will be sent to your email address (It might take a few minutes for the code to come, so please be patient)
- 3- Verify your account with the code to get access to the app.

#### Step 3: Have a conversation with AI on a topic of your choice:

- 1- Go to the Homepage on the app
- 2- Go to “Create your own conversation” section:
  - Enter a topic and the AI role (a friend, a sibling, a job interviewer, a co-worker...) of your choice
  - Start and complete the conversation



**NOTE:**

During 7 days, please note down any comments and feedback you have regarding the experience of conversing with AI. The note can include, for example:

- What you think about the quality of the conversation with AI
- What you like and don't like about the AI during the conversation
- How useful you think the conversation is to your Korean studies
- What you like to change or add to the conversation / application
- What factors motivate you to learn Korean through this app in the long term
- Other thoughts and feedback

This note will help you remember and describe your experience during the interview session, which will be scheduled after this 7-day period.

I will contact you shortly after the 7-day period to set up an online interview via Zoom.

Thank you and happy studying!

## Appendix 4

### Semi-Structured Interview Questions

Section	Question
General experience in Korean language learning	<ul style="list-style-type: none"> <li>-Why did you decide to learn Korean?</li> <li>-What are your goals for learning Korean?</li> <li>-How long have you been learning Korean?</li> <li>-What are the difficulties you face when learning Korean (in general)?</li> <li>-What motivates you to keep learning Korean?</li> </ul>
Overall experience	<ul style="list-style-type: none"> <li>-How is your overall experience practicing Korean conversations with the chatbot?</li> <li>-Did the experience meet your expectations? Why/why not?</li> </ul>
Detailed experience	<ul style="list-style-type: none"> <li>-Can you describe a free topic conversation you had with the chatbot?</li> <li>-What did you learn through the conversation?</li> <li>-How does using the chatbot make you feel about learning Korean?</li> <li>-Do you feel comfortable making mistakes when interacting with the chatbot?</li> <li>-What aspects/features of the chatbot do you find helpful for conversation practice?</li> </ul>

	<p>-What difficulties did you have during the conversation with the chatbot?</p> <p>-Is there anything you would like to change/add to make the conversation practice easier and more effective?</p> <p>-How does interacting with the chatbot compare to other language learning methods you've tried?</p>
Motivation	<p>-What aspects of using the chatbot do you find most engaging or enjoyable?</p> <p>-What do you think about chatbot as a conversation partner? (expectation on how human-like chatbot should be)</p> <p>-How many conversation sessions did you have with the chatbot during one week?</p> <p>-How long was each conversation? When and why did you decide to stop the conversation?</p> <p>-In what ways do you think the chatbot has motivated you to improve your language skills?</p> <p>-Do you feel like you are making progress in your language learning by using the chatbot?</p> <p>-How do you measure your progress when using the chatbot? How do you feel about the progress?</p> <p>-Do you feel like the chatbot helps you stay motivated to learn?</p>

	<p>-Do you want to continue using the chatbot to learn Korean? Why / Why not?</p> <p>-What motivates you to continue using the chatbot?</p> <p>-What discourages you from continuing using the chatbot?</p>
Others	<p>-Any other feedback or opinions you would like to share regarding your experience with the chatbot?</p>